On weighted Poisson distributions and processes, with associated inference and applications

by
Philip Albert Mijburgh 10084704 (UP), 400194849 (McMaster)

Submitted in fulfillment of the degree PhD Mathematical Statistics

In the Faculty of Natural and Agricultural Sciences
University of Pretoria
Pretoria

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

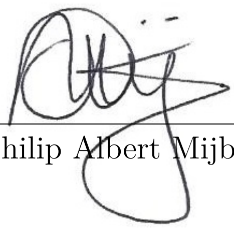
In the Faculty of Science
McMaster University
Hamilton

McMaster University

26 November 2020

Declaration

I, Philip Albert Mijburgh, declare that this thesis, which I hereby submit for the degree *PhD Mathematical Statistics* at the University of Pretoria in cotutelle with McMaster University, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.



Philip Albert Mijburgh

24/08/2020

Date

This thesis is dedicated to my father, Philip, and wife, Marianne.

To my father: Thank you for raising me the way you did and for encouraging me to take my studies as far as possible. Without your support, this thesis would never have been started.

To my wife: Thank you for always believing in me and for encouraging me through many long nights of work. Without your support, this thesis would never have been completed.

Acknowledgments

I wish to express my sincere appreciation for the following persons, without whom this thesis would not be possible:

- My supervisors, Dr. Jaco Visagie and Prof. Narayanaswamy Balakrishnan. Dr. Visagie is the most supportive supervisor I could have hoped for, who often went way beyond the call of duty in supporting my research. Prof. Balakrishnan for his truly vast knowledge and ability to guide my efforts into fruitful directions.
- Prof. Andriette Bekker, who mentored and taught me how to perform a proper research study. Her guidance and support during my postgraduate studies has been invaluable.
- Everyone involved in establishing the cotutelle program between the University of Pretoria and McMaster University.
- My supervisory committee and examiners for their invaluable feedback and recommendations.
- My family and friends for their continued support. Specifically Johnnie, who walked this long road with me.

Lastly, I would like to acknowledge the financial assistance I received:

- UP Postgraduate Research Support Bursary.
- UP Postgraduate Study Abroad Bursary.
- STATOMET Bursary.
- SASA/NRF Academic Statistics Bursary.

Note that all results, opinions and conclusions reached in this thesis are those of the author and should not necessarily be ascribed to those funding or supporting this research,

Abstract

In this thesis, weighted Poisson distributions and processes are investigated, as alternatives to Poisson distributions and processes, for the modelling of discrete data.

In order to determine whether the use of a weighted Poisson distribution can be theoretically justified over the Poisson, goodness-of-fit tests for Poissonity are examined. In addition to this research providing an overarching review of the current Poisson goodness-of-fit tests, it is also examined how these tests perform when the alternative distribution is indeed realised from a weighted Poisson distribution. Similarly, a series of tests are discussed which can be used to determine whether a sample path is realised from a homogeneous Poisson process.

While weighted Poisson distributions and processes have received some attention in the literature, the list of potential weight functions with which they can be augmented is limited. In this thesis 26 new weight functions are presented and their statistical properties are derived in closed-form, both in terms of distributions and processes. These new weights allow, what were already very flexible models, to be applied to a range of new practical situations.

In the application sections of the thesis, the new weighted Poisson models are applied to many different discrete datasets. The datasets originate from a wide range of industries and situations. It is shown that the new weight functions lead to weighted Poisson distributions and processes that perform favourably in comparison to the majority of current modelling methodologies. It is demonstrated that the weighted Poisson distribution can not only model data from Poisson, binomial and negative binomial distributions, but also some more complex distributions like the generalised Poisson and COM-Poisson.

Contents

Chapter 1

Introduction

Since its initial publication nearly 200 years ago, the Poisson distribution has been widely used to model count data in a range of different disciplines and scientific fields of study. Consequently, a large number of researchers have studied its properties, proposed generalisations, additions and modifications in attempts to augment the initial distribution proposed by Siméon Denis Poisson [110]. (See Haight [63], Patil and Joshi [106] and Johnson and Kotz [74] for extensive lists and discussions of Poisson generalisations.) One of the properties of the Poisson distribution is equidispersion, meaning that the distribution's mean is equal to its variance. This fact, although useful in specific settings, often limits the distribution's ability to accurately model observed data.

In practical situations data that are overdispersed (where the variance exceeds the mean) are common (see Hinde et al. [70] and Lawless [86] for two examples), but underdispersed situations (where the mean exceeds the variance) are not unheard of (see Wimmer et al. [143] for an example). Böhning [71] noted that overdispersion is often the result of latent heterogeneity in data, which means that a sample of variables consists of different sub-groups that have not been distinguished from each other. This occurrence is often referred to as "heterogeneity and aggregation" in the literature. Ideally, all subpopulations would be treated and modelled separately; however, after data collection, it might not always be possible to distinguish these subgroups from each other. It could also be possible that segmenting the sample into subgroups would lead to such small sample sizes for the subgroups as to make modelling them impractical. In such situations attempting to include the underlying heterogeneity in the modelling of the data would be preferable to merely ignoring it.

Another factor that could lead to non-equidispersion is correlated data. Bosch and Ryan [16] noted that natural situations that lead to competition (or alternatively stated, those that have a negative correlation structure inherent in the data) result in underdispersed data, whereas overdispersed situations result from a positive correlation structure. They anecdotally demonstrated this by considering two previous papers. The first concerned the number of sea-urchin eggs fertilised in various time intervals, where sperm were "competing" to fertilise eggs, which led to an underdispersed data set. The second paper recorded the number of patient visits to doctors in Canada, where one visit would increase the probability of a repeat visit (is it postulated these repeat visits are either due to the patients not recovering fully, or to doctors prescribing follow-up visits). This positive correlation in the data leads to the number of visits being overdispersed.

To illustrate the shortcomings of the Poisson distribution to model data that are not realised from an equidispersed distribution, a dataset of weekly sales figures of 800 items from a store is used. This dataset can be found at `https://archive.ics.uci.edu/ml/datasets/Sales_Transactions_Dataset_Weekly`, and was initially analysed by Tan and San Lau [136]. In Figure 1.1 below, the empirical probability mass function (Definition 10.2) of the data is compared to the fitted probability mass function of the Poisson distribution (Definition 10.3), using maximum likelihood parameter estimation. In Figure 1.1, the empirical values are represented by blue dots with vertical lines and the theoretical values by black dots (which are connected to make interpretation clearer). For illustrative purposes, only one extreme underdispersed (item 726) and overdispersed (item 409) graph are shown out of the possible 800 items.

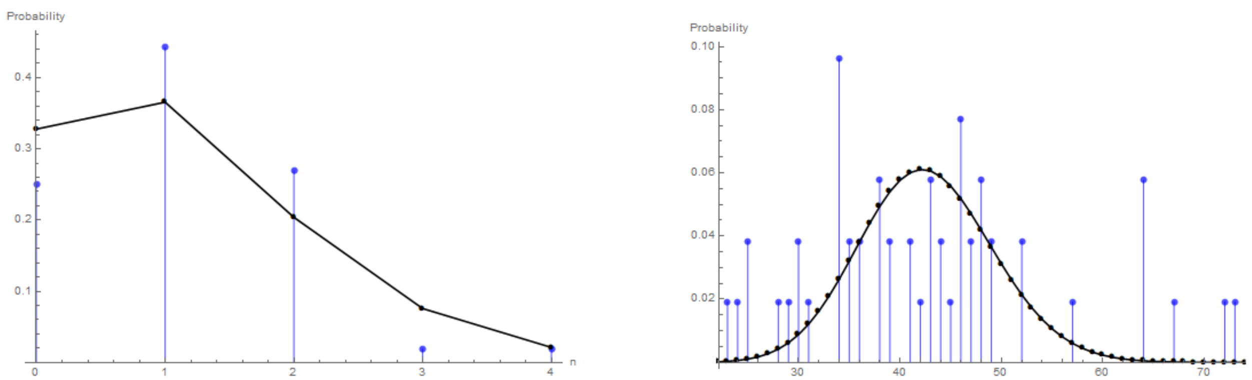


Figure 1.1: Underdispersion and overdispersion against fitted Poisson

As can be seen from the plots, the Poisson distribution does not provide a good fit to the data, whether under or overdispersed. In the (first) underdispersed case, the Poisson distribution underestimates probabilities close to the mean and overestimates those in the tails. In contrast, in the (second) overdispersed case, the values close to the mean are overestimated, and those in the tails tend to be underestimated.

It could be argued that the data could be better modelled by using a time series approach, as was proposed by Tan and San Lau [136]; however, using the Poisson distribution is not a wholly unsound idea since the data will likely adhere relatively closely to most of the assumptions of the Poisson distribution. These being that:

1. The number of items sold during the week must be an integer amount ranging from 0 to infinity. (The data is "count" data.)

2. The event of an item being bought does not affect the likelihood that more or less of the same item will be purchased. (Events are independent.)

3. The average rate at which purchases occur from week to week is constant. (The rate of events occurring is constant.)

The third restriction seems least likely to be met, since seasonal fluctuations may affect the number of weekly sales.

To overcome the limitations of the Poisson distribution, several authors have developed or proposed various alternative methods and distributions to model count data that are not equidispersed. Many of these techniques are briefly discussed in Chapter 2. This thesis proposes to use the weighted Poisson framework to deal with data that are not equidispersed (and consequently do not follow the Poisson distribution). The thesis is laid out in such a way that the order of the chapters largely follows the same progression that someone who wishes to implement the proposed models would have to follow.

In Chapter 3, goodness-of-fit testing for the Poisson distribution will be discussed, where the performance of classical tests of historical importance is compared to that of more recent tests for Poissonity. The powers of these tests against weighted Poisson alternatives are considered for the first time in the statistical literature in this study. In Chapter 4, a range of weight functions, the vast majority which are novel, will be investigated. In this chapter, statistical properties of the resulting weighted Poisson distributions will be derived, and plots will be presented to demonstrate the potential shapes that the distributions can assume. Chapter 5 will demonstrate the application of the weighted Poisson methodologies to various observed datasets to demonstrate the flexibility and wide range of applications of the proposed models. The stochastic process equivalents of these topics will be discussed in Chapters 6, 7 and 8 respectively. This progression is summarised in Figure 1.2.

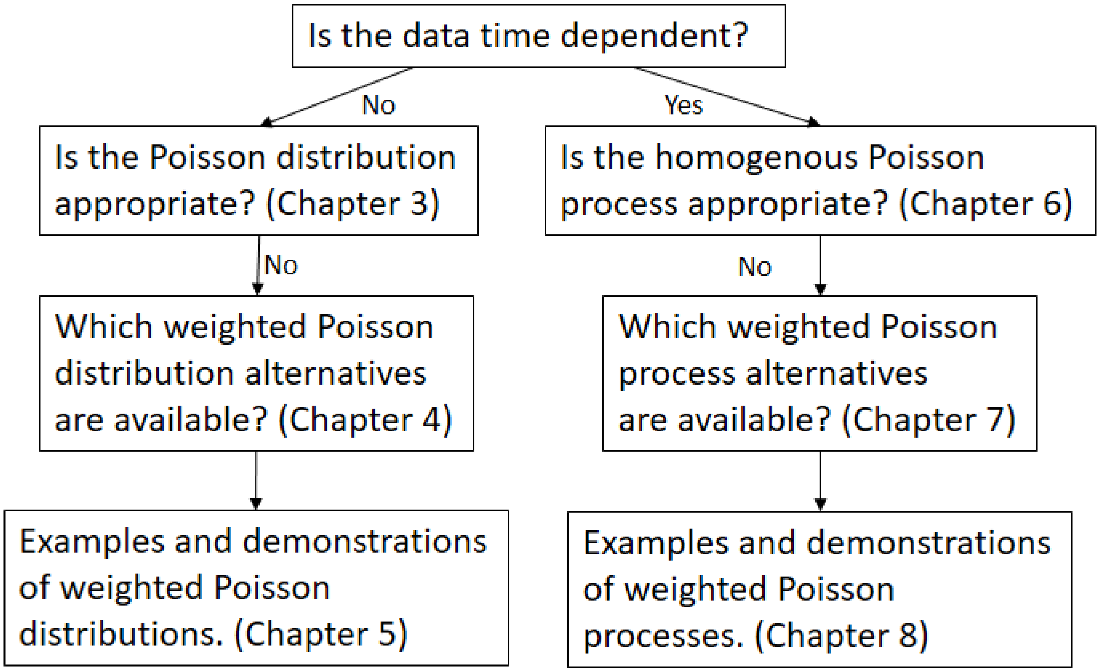


Figure 1.2: Flow of the thesis

While the main focus of this thesis is concerned with the modelling of non-equidispersed data, the weighted Poisson framework also offers some novel implementation possibilities for specific practical problems. These will be discussed in more detail in later chapters; however, they include concepts like the "zero-inflated" Poisson (which is often used to model churn rate or insurance claim numbers), the modelling of truncated distributions, (for instance the number of vehicles involved in car accidents is strictly one or more), as well as bi-modal (and potentially multi-modal) distributions. As will be seen in Chapter 5, the weighted Poisson distribution is also capable of modelling data which come from more traditional distributions like the Poisson, binomial and negative binomial.

Chapter 2

Literature Review

This chapter is focused on alternatives to the Poisson distribution that can be used for the modelling of non-equidispersed count data. These methods have received much attention, and consequently, a few overview papers have been written regarding them (see Kokonendji [82] and Sellers and Morris [124].) In Section 2.1, an overview of the methods currently available in the literature is presented. However, this is in no way a definitive collection of the vast body of research dedicated to the modelling of data exhibiting non-equidispersion. In Section 2.2, the weighted Poisson distribution is formally defined, and the literature that forms the foundation of this thesis is discussed.

Note that many of the techniques discussed in this chapter were initially presented in a multivariate or regression-based settings. To standardise the notation, as well as to keep the theme of the thesis coherent, these models will be reduced to the univariate cases. While it may be possible to extend the weighted Poisson distributions discussed in this thesis to the multivariate and regression realms, this lies outside the scope of this thesis.

2.1 Available methods

Before presenting the methodologies that are commonly used to model non-equidispersed data, it is useful to define a measure of the dispersion of a random variable, so that a coherent notion of what is meant by under, over and equidispersion can be conveyed.

While many different measures of dispersion exist (see Boos and Brownie [15]), the one that is most commonly used in the relevant literature is the "Fisher index" (FI) defined below.

Definition 2.1.
Suppose that X is some random variable with expected value $E(X)$ and variance $Var(X)$. Then the Fisher index is given by

$$FI(X) = \frac{Var(X)}{E(X)}. \tag{2.1}$$

If $FI(X) = 1$, X is said to be equidispersed. If $FI(X) > 1$, X is overdispersed, and if $FI(X) < 1$, X is underdispersed.

Establishing the relationship between the expected value and the variance of a random variable is a key feature noted in most relevant research papers, since it not only establishes if a random variable is over or underdispersed, but it also determines the types of data structures that can be modelled with specific distributions.

The remainder of this section is segmented by the types of data that the various models can accommodate. Section 2.1.1 will discuss models that are overdispersed, Section 2.1.2 will discuss models that are underdispersed, and Section 2.1.3 will discuss models that can be both under and overdispersed.

2.1.1 Overdispersion

Below two models are considered that are used in the modelling of overdispersed data.

Negative binomial modelling

The most common approach to dealing with overdispersed data is to use the negative binomial distribution. A random variable is said to follow the negative binomial distribution if it has the probability mass function as given in Definition 10.6. A negative binomial random variable, X, has expected value $E(X) = \frac{pr}{1-p}$ and variance $Var(X) = \frac{pr}{(1-p)^2} = E(X)\frac{1}{1-p}$. Since $0 < p < 1$, it follows that $0 < 1 - p < 1$, and consequently, the variance of a negative binomial distribution is larger than its expected value.

There is a host of literature available which discusses the many ways in which the negative binomial distribution can be justified theoretically as well as practically as an alternative to the Poisson distribution when modelling count data. Arguably the most comprehensive reference source for negative binomial modelling and regression is Hilbe [69] in which many facets and applications of this method are discussed.

There are many ways in which the negative binomial distribution can be obtained from the Poisson distribution. One example is using mixed Poisson modelling, discussed below. It should also be noted that the Poisson distribution is a case distribution attained of the negative binomial distribution when $r \to \infty$ and $\lambda = \frac{rp}{1-p}$.

Mixed Poisson modelling

Many authors have investigated the possibility of using mixed Poisson distributions to model count data. In essence, X is said to follow a mixed Poisson distribution if X is $Poisson(\lambda)$ distributed. However, instead of λ being a parameter, λ is itself assumed to be a random variable that has an underlying distribution. This method provides a rather intuitive approach to deal with aggregation and heterogeneity in data, since, as long as the distribution of λ adequately models the distribution of the heterogeneity, the resulting $Poisson\left(\lambda\right)$ distribution should provide an adequate model for the overall population.

Feller [47] investigated the properties of this distribution and showed, among other things, that mixed Poisson distributions always have a higher probability of observing zeros relative to the Poisson distribution, given that they have the same mean. Shared [125] showed that mixed Poisson distributions also have thicker tails. Karlis and Xekalaki [77] published an overview paper discussing the general properties of mixed Poisson distributions. They also derived the relationship between the mean and variance of the mixture distribution: $E\left(X\right) = E\left(\lambda\right)$ and $Var\left(X\right) = E\left(\lambda\right) + Var\left(\lambda\right)$. Since $Var\left(\lambda\right) \geq 0$ mixed Poisson distributions are overdispersed (or equidispersed in the trivial case that the random variable λ has no variability, in which case the distribution reduces to the Poisson).

Greenwood and Yule [60], Cameron and Trivedi [18], Johnson et al. [74], Sichel [127, 128], Willmot [141], Sankaran [97, 96] , Karlis and Xekalaki [77], and many others have discussed a range of potential distributions that λ can follow. Greenwood and Yule [60], as well as Cameron and Trivedi [18] specifically proposed using a gamma distribution for λ. Assuming that λ follows a gamma distribution ensures that a random variable follows a negative binomial distribution. For more details on the use of the multivariate case, see Hausman et al. [64] and McCullagh and Nelder [93].

Lawless [86] discussed mixed Poisson distributions and negative binomial modelling of count data in a similar vein to that of Cameron and Trivedi [18]. However, the discussion focused more on the specifics of how parameters could be estimated as well as the robustness of the achieved estimates. In the paper, special attention was paid to dealing with overdispersed data; however, the author noted that with only a slight relaxing of the constraints placed on their distribution, it would also be possible to model underdispersed data through the use of binomial or generalised binomial models (See Olkin et al. [104]).

2.1.2 Underdispersion

Below two models are considered that are used in the modelling of underdispersed data.

Binomial modelling

A frequently used method to deal with underdispersed data is to use the binomial distribution. A random variable is said to follow the binomial distribution if it has the probability mass function given in Definition 10.5. A binomial random variable, X, has expected value $E\left(X\right) = np$ and variance $Var\left(X\right) = np\left(1-p\right) = E\left(X\right)\left(1-p\right)$. Since $0 < p < 1$, it follows that $0 < 1-p < 1$, and consequently, the variance of a binomial distribution is smaller than its expected value. Due to the fact that the binomial distribution is extremely prevalent in statistical implementations, its univariate properties have been very well researched. As a result, any research that has been published regarding it in the last decades have all concerned multivariate or regression settings. Its univariate form, however, is still considered a useful distribution to model underdispersed data.

Similar to the negative binomial distribution, a connection exists between the binomial and Poisson distributions. The Poisson distribution is a limiting case of the binomial distribution, obtained when $n \to \infty$ and $\lambda = np$.

Condensed Poisson modelling

If X is a Poisson process (see Definition 10.25), then the inter-arrival times (see Definition 10.31) of events follow an exponential distribution (see Definition 10.7). However, if the inter-arrival times of a process are defined to be Erlang$(2, \lambda)$ distributed (see Definition 10.8) then the resulting random number of events per unit time, say X^*, is said to follow a condensed Poisson distribution.

The condensed Poisson distribution has probability mass function given by

$$f\left(x^*; \lambda\right) = \begin{cases} e^{-\lambda}\left(1 + \frac{\lambda}{2}\right) & x^* = 0 \\ e^{-\lambda}\left(\frac{\lambda^{2n^*-1}}{2(2n^*-1)!} + \frac{\lambda^{2n^*}}{2(2n^*)!} + \frac{\lambda^{2n^*+1}}{2(2n^*+1)!}\right) & x^* = 1, 2, ... \end{cases} \quad (2.2)$$

The condensed Poisson distribution has expected value $E\left(X^*\right) = \frac{\lambda}{2}$ and variance $Var\left(X^*\right) = \frac{\lambda + e^{-\lambda} sinh(\lambda)}{4}$ where $sinh\left(\lambda\right)$ denotes the hyperbolic sin function. The Fisher index of this distribution is strictly less than 1. Originally this distribution was named the "asynchronous counting distribution" by Haight [63], but was renamed to the condensed Poisson distribution by Chatfield and Goodhardt [20]. Chatfield and Goodhardt [20] also used the distribution to model consumer purchasing data and showed that it provided an adequate representation of the observed datasets.

2.1.3 Under and overdispersion

Below seven models are considered that are used in the modelling of over and underdispersed data.

The generalised Poisson distribution

In some of the relevant literature, this distribution has also been called the Lagrange-Poisson distribution.

One of the most prolific researchers into the modelling of non-equidispersed data is Prem Consul. Consul and Jain [28, 29, 30] developed a generalised, two-parameter Poisson distribution as a limiting form of the generalised negative binomial distribution. A random variable, X, is said to follow the generalised Poisson distribution if it has the probability mass function as given in Definition 10.4. This distribution allows for under, over and equidispersion since

$$E(X) = \frac{\lambda_1}{1-\lambda_2} \quad \text{and} \quad Var(X) = \frac{\lambda_1}{(1-\lambda_2)^3} = E(X)\frac{1}{(1-\lambda_2)^2}.$$

Thus the mean of the generalised Poisson distribution will be smaller than, equal to, or greater than the variance if the value of λ_2 is positive, zero, or negative, respectively. Consul and Jain [29] also applied their proposed model to different datasets, traditionally assumed in the literature to follow the Poisson, binomial, and negative binomial distributions. They showed that the use of the two-parameter generalised Poisson model resulted in satisfactory measures of fit, even when considering that the datasets were traditionally assumed to be realised from other distributions.

In the years after the derivation of the distribution, much research has been published investigating the various properties and applications of the distribution. See Consul [26, 27] as well as Consul and Shoukri [31] to name a few.

A semi-parametric approach - Generalising the relationship between the mean and the variance

Unlike in most of the literature, Hinde [70] attempted to model overdispersion in data not by finding or deriving distributions that fit the data, but rather by explicitly specifying the relationships between the expected value and variance in a model. It was noted in the paper that in most binary response cases it was assumed that count data was binomially distributed, but that data would often exhibit a greater level of overdispersion than the binomial model allowed for. That is to say: assuming that $X \sim Bin(n,p)$,

$$E(X) = np \quad \text{and} \quad Var(X) = np(1-p) = E(X)(1-p)$$

would underestimate the level of dispersion. In response to this problem, four alternative variance formulations were proposed: The constant overdispersion model

$$Var(X) = E(X)\phi(1-p),$$

where ϕ is a constant. The beta-binomial overdispersion model

$$Var(X) = E(X)(1-p)(1+\phi(n-1)),$$

(this equation can be modified to account for excess underdispersion by replacing ϕ with the correlation coefficient of data). The logistic normal model

$$Var(X) = E(X)(1-p)\left(1+\sigma^2(n-1)p(1-p)\right),$$

and the general variance model

$$Var(X) = E(X)(1-p)\left(1+\phi(n-1)^{\alpha}(p(1-p))^{\beta}\right),$$

which encapsulates the other three models (when α and/or β are set to either zero or one).

Using a similar methodology to the binomial distribution, Hinde [70] also investigated how the variance-mean relationship of the Poisson distribution could be altered. Once again four alternative models were proposed: the constant overdispersion model

$$Var\left(X\right) = E\left(X\right)\phi,$$

the negative binomial type variance

$$Var\left(X\right) = E\left(X\right) + \tfrac{E(X)}{k},$$

(where k is a parameter of the gamma function), the Poisson-normal type models

$$Var\left(X\right) = E\left(X\right)\left(1 + E\left(X\right)k\right),$$

and a generalised variance model where

$$Var\left(X\right) = E\left(X\right)\left(1 + \phi E\left(X\right)^{\alpha}\right).$$

Two important facts should be noted about Hinde's [70] methodology. First, assuming the desired form of the relationship between the model's variance and expected value often results in the models not being associated with any known probability distribution. This has the disadvantage of not allowing for maximum likelihood estimation of parameters, but rather requiring that some ad hoc, quasi-likelihood methods need to be implemented. Additionally, many other important statistical properties are merely ignored when using this approach. The second fact is that while probability distributions do not exist for most models, in some particular cases, they do, especially when an approach similar to that applied earlier in this section by Cameron and Trivedi [18] is used.

Ver Hoef and Beoveng [139] compared the modelling of overdispersed data using the negative binomial distribution and the quasi-Poisson approach (which is similar to the constant Poisson overdispersion model discussed in Hinde [70]). While they did not make general statements as to which method performs better, they noted that in the example they considered, the quasi-Poisson method provided a better fit to the data.

The double Poisson distribution

Efron [42] proposed the double Poisson distribution as a special case of the double exponential family. A random variable, X, is said to follow a double Poisson distribution with parameters $\alpha, \lambda > 0$ if it has the following probability mass function:

$$f\left(x;\alpha,\lambda\right) = c\left(\alpha,\lambda\right)\left(\tfrac{e^{-\alpha\lambda}}{\sqrt{\alpha}}\right)\left(\tfrac{e^{-x}x^{x}}{x!}\right)\left(\tfrac{e\lambda}{x}\right)^{\alpha\lambda}, \quad x = 0,1,2,..., \tag{2.3}$$

where $c(\alpha, \lambda)$ a normalising constant, and is given by

$$\frac{1}{c(\alpha,\lambda)} = \sum_{j=0}^{\infty} \left(\frac{e^{-\alpha\lambda}}{\sqrt{\alpha}}\right)\left(\frac{e^{-j}j^j}{j!}\right)\left(\frac{e\lambda}{j}\right)^{\alpha\lambda} \approx 1 + \frac{1-\alpha}{12\lambda\alpha}\left(1+\frac{1}{\lambda\alpha}\right).$$

This distribution has expected value $E(X) = \lambda$ and variance $Var(X) = \frac{\lambda}{\alpha} = \frac{E(X)}{\alpha}$, and consequently is overdispersed for $0 < \alpha < 1$, underdispersed when $\alpha > 1$, and if $\alpha = 1$ the double Poisson distribution simplifies to the Poisson distribution.

The Conway-Maxwel-Poisson distribution

The Conway-Maxwel-Poisson distribution (more commonly referred to as the COM-Poisson distribution) was initially proposed by Conway and Maxwell [32] in the context of queuing theory. This two-parameter distribution possesses many properties which make it very attractive for practical implementation. For example, it is a member of the exponential family, which, in turn, enables simple maximum likelihood parameter estimation. It also contains the binomial, negative binomial and Poisson distributions as special cases of the family, which implies that the COM-Poisson distribution can accommodate both under and overdispersion.

Shmueli et al. [126] noted, however, that "no probabilistic or statistical characterisations of this distribution and extremely few applications appear in the literature." Consequently, Shmueli et al. [126] derived many of the properties of the distribution as well as proposing three methods by which parameter estimates of the distribution could be calculated. Furthermore, their theoretical results were used in the analysis to two datasets. One contained the quarterly sales of a specific paper of clothing at stores of a large national retailer (which were overdispersed), the other contained the lengths of words in a Hungarian dictionary (which were underdispersed). The authors showed that in both situations, the COM-Poisson distribution provided a substantially better fit to the data than was the case for the Poisson distribution. It should be noted that it is only after the Shmueli et al. [126] paper that the COM-Poisson became a commonly used distribution to model non-equidispersed data.

A random variable, X, is said to follow a COM-Poisson distribution with parameters $\lambda > 0, \alpha \geq 0$, if it has the following probability mass function:

$$f(x;\alpha,\lambda) = \frac{\lambda^x}{(x!)^{\alpha} Z(\alpha,\lambda)}, \quad x = 0, 1, 2, ..., \tag{2.4}$$

where $Z(\alpha, \lambda)$ a the normalising constant given by

$$Z(\alpha,\lambda) = \sum_{j=0}^{\infty} \frac{\lambda^j}{(j!)^{\alpha}}.$$

Closed-form expressions for the moments of the COM-Poisson distribution do not exist, but Sellers et al. [123] derived the following approximations based on the moment generating function:

$$E(X) \approx \lambda^{\frac{1}{\alpha}} - \frac{\alpha-1}{2\alpha} \quad \text{and} \quad Var(X) \approx \frac{1}{\alpha}\lambda^{\frac{1}{\alpha}}$$

which hold when $\alpha < 1$ or $\lambda > 10^{\alpha}$. The parameter α is often called the "dispersion parameter" since it dictates whether the distribution will be over, under or equidispersed. If $\alpha > 1$ the distribution is underdispersed, $0 < \alpha < 1$ implies overdispersion, and $\alpha = 1$ will result in the distribution simplifying to the Poisson. It should also be noted that if $\alpha \to \infty$, then the COM-Poisson tends to the $Bernoulli\left(\frac{\lambda}{\lambda+1}\right)$ distribution and if $\alpha = 0, \lambda < 1$ the distribution is $geometric\,(\lambda)$.

Changing birth rate distribution

Bosch and Ryan [16], developed the 'changing birth rate distribution', which allows for the modelling of over and underdispersion. The distribution originated through the use of Markov processes, and although closed-form expressions do not exist in general for the distribution, there are special cases in which these expressions exist. The methodology applied by Bosch and Ryan [16] is similar to that used by Consul [27, 26] when they proposed the two-parameter generalised Poisson distribution.

The gamma count distribution

Winkelmann [144] proposed a distribution to model count data based on the difference between two lower-incomplete gamma functions. X is said to follow a gamma count distribution with parameters $\alpha, \lambda > 0$ if it has the following probability mass function

$$f(x; \lambda) = \frac{1}{\Gamma(\alpha x)}\gamma(\alpha x, \lambda) - \frac{1}{\Gamma(\alpha x + \alpha)}\gamma(\alpha x + \alpha, \lambda) \tag{2.5}$$

where $\gamma(.)$ is the lower incomplete gamma function as given in Definition 10.11 and $\Gamma(.)$ is the gamma function in Definition 10.9.

This distribution is overdispersed when $0 < \alpha < 1$ and underdispersed if $\alpha > 1$. Closed-form expressions for the moments of the distribution do not exist.

Compound Poisson modelling

Another class of distributions which can be used to model non-equidispersed data is the compound Poisson distribution. Suppose that X is a Poisson random variable and that $Z_1, ..., Z_X$ are independent and identically distributed (i.i.d.) random count variables, independent of X. Then Y is called a compound Poisson distribution if

$$Y = \sum_{j=1}^{X} Z_j. \tag{2.6}$$

The properties of compound distributions are well known and can be found in most introductory statistics textbooks (see Bain and Engelhardt [4] and Johnson et al. [74]). Zhang and Li [146] and Zhang et al. [147] studied the characterisation, properties and application of this model to risk theory. It should be noted that the over or underdispersion of the compound Poisson model is completely defined by the summands Z_j, since $E(Y) = E(X)E(Z)$, $Var(Y) = E(X)E(Z^2)$, and consequently $FI(Y) = \frac{E(Z^2)}{E(Z)}$.

2.2 The weighted Poisson distribution

Although many of the methods mentioned in Section 2.1 have been applied in various specific practical situations, most of them have certain drawbacks that limit their application in more general situations. Some can exclusively accommodate either over or underdispersion, some do not have closed-form expressions for the probability mass functions or moments, and some require nonstandard parameter estimation methods in order to fit these models to observed data.

However, in 1934, Fisher [50] introduced the concept of weighted distributions through the method of ascertainment. This concept has been used extensively since then as a method to augment standard probability mass functions, the goal being to define more flexible distributions which allow users to better fit data and to facilitate the selection of appropriate models for observed data. (It should be noted that while many authors cite Rao [113] when discussing the method of ascertainment, it was, in fact, Fisher [50] who initially proposed the idea.) Although some papers have been published that use this method in conjunction with the Poisson distribution with specific weight functions, they are limited in number.

In addition to detailing the papers that discuss the weighted Poisson distributions in this section, some initial definitions are given that will be used throughout the remainder of the thesis.

Definition 2.2.
Let N be a random variable with probability mass function $f(n) = P(N = n)$. Suppose that when the event $N = n$ occurs, the probability of ascertaining it is $w(n)$. The observed value, n, is then a realisation of the random variable N^w, which is said to be the weighted version of N. The probability mass function of N^w is given by

$$f_w(n) = P(N^w = n) = \frac{w(n)f(n)}{E(w(N))} \quad , \quad n = 0, 1, 2, ..., \tag{2.7}$$

where $E(w(N)) = \sum_{k=0}^{\infty} w(k)f(k) < \infty$ and $w(n)$ is a non-negative function on $\mathbb{N}_0$.

See Definition 10.1 for the definitions of $\mathbb{N}_0$ and $\mathbb{N}_1$.

Another, possibly more natural, way of thinking about the method of ascertainment is to think of $w(n)$ as some non-negative "weight function" that is multiplied with the probability mass function of N to give some new "weighted" variable N^w. $E(w(N))$ can be interpreted as the normalising constant. Suppose that N in Definition 2.2 is a Poisson random variable, then the resulting random variable, N^w, is said to follow a weighted Poisson distribution with weight function $w(n)$. Exploring a wide range of these weight functions and the properties of the resulting weighted Poisson distributions is at the core of this thesis.

An important point must be made about the notation of the weight function. Often in the literature, the weight function is merely written as $w(n)$. However, for the remainder of this thesis, distinction will be made between weight functions that depend only on n, denoted by $w(n)$, functions that also include a (possibly vector-valued) free parameter, ϕ, denoted by $w(n;\phi)$, and weight functions that depend on ϕ as well as the Poisson rate parameter λ, $w(n;\lambda,\phi)$. These distinctions are important since some of the results available in the literature are classified according to these different forms of weight functions (see Balakrishnan and Kozubowski [6] for examples where this is the case).

For the remainder of this chapter, general statistical properties of the weighted Poisson distribution, as well as various weight functions that are currently available in the literature, will be discussed. Derivations showing some of the properties of these weight functions will be included in Chapter 4.

If $w(.) = c\ \forall n \in \mathbb{N}_0$ where c is some constant, then the weighted Poisson distribution reduces to the Poisson distribution. This result is easy to verify since this simplification in Equation 2.7 is apparent:

$$f_w(n) = \frac{w(n) f(n)}{E(w(N))} = \frac{cf(n)}{\sum_{k=0}^{\infty} cf(k)} = \frac{cf(n)}{c\sum_{k=0}^{\infty} f(k)} = f(n),$$

since $f(k)$ is the Poisson probability mass function and thus sums to 1.

If $w(n) = n$, the resulting weighted Poisson distribution is known as the "size-based" Poisson or "1-translated" Poisson distribution. This specific distribution was discussed by Patil and Rao [107] and by Kokonendji et al. [83]. The size-based weight function leads to an underdispersed weighted Poisson distribution.

If $w(n;\phi) = (n!)^{1-\alpha}$, the weighted Poisson distribution is equal in distribution to the COM-Poisson distribution discussed in Section 2.1. This equality has been discussed in many papers, one of which is Kokonendji et al. [83]. As has already been mentioned, the COM-Poisson can be under, over or equidispersed depending on the value of α.

The weight function $w(n;\phi) = (n+a)^r$ where $a > 0$ is a displacement parameter and $r \in \mathbb{R}$ was discussed in depth by Castillo and Perez [38]. This weighted Poisson distribution is overdispersed, underdispersed and equidispersed if and only if $r < 0, r > 0$ and $r = 0$ respectively. As such, the parameter r can be interpreted as the dispersion parameter. Castillo and Perez [38] derived many properties of the weighted Poisson distribution when

$w(n;\phi) = (n+a)^r$. They derived expressions for the stochastic order of many different parameterisations of their proposed weighted Poisson distribution. (Stochastic order quantifies the notion of one random variables being "larger" than another. In essence it is the stochastic equivalent of the traditional mathematical order operators like $<, \leq, >, \geq$). They also showed that if two weighted random variables exist with the same form of weight function, containing different parameter values, it is possible to consider one as a weighted version of the other. Specifically, suppose that N^w is a weighted Poisson random variable with rate parameter λ_1 and weight function $w(n;\phi) = (n+a)^{r_1}$, then if N^w is weighted again with respect to weight function $w_1(n;\phi,\lambda_2) = \left(\frac{\lambda_2}{\lambda_1}\right)^{r_1}$, the resulting random variable N^{w_1} is equal in distribution to a weighted Poisson distribution with rate parameter λ_2 and weight function $w(n;\phi) = (n+a)^{r_1}$. Similarly the weights $w_2(n;\phi) = (n+a_1)^{r_1-r_2}$ and $w_3(n;\phi) = \left(\frac{n+a_2}{n+a_1}\right)^{r_1}$ result in N^{w_2} being weighted Poisson distributed with rate parameter λ_1 and weight function $w(n;\phi) = (n+a_1)^{r_2}$ and N^{w_3} following a weighted Poisson distribution with rate parameter λ_1 and weight function $w(n;\phi) = (n+a_2)^{r_1}$.

Castillo and Perez [39] also investigated the properties of the weighted Poisson distribution if the weight function is given by $w(n;\phi) = e^{rt(n;\lambda,\phi)}$, where $t(n;\lambda,\phi)$ is a real valued function depending on n, and which may but need not depend on λ or ϕ. The majority of their paper focuses on the various forms that $t(n;\lambda,\phi)$ can assume, and how these various forms affect the overall properties of the distribution. Specifically they noted that if $t(n;\lambda,\phi)$ is statistically bounded of order n (in other words if $t(n;\lambda,\phi) = O(n)$) then their weighted Poisson distribution would be a member of the regular exponential family. Many potential forms of this function were given: $t(n;\phi) = n^a, 0 < a \leq 1$, $t(n;\phi) = ln(n+a), a > 0$, $t(n) = \sqrt{n}ln(n+1)$, $t(n) = e^{-n}$ and $t(n;\phi) = \frac{an+b}{cn+d}$ where c and/or d are not equal to 0. However, if $t(n;\lambda,\phi)$ is not statistically bounded of order n the resulting weighted Poisson distributions are likely not members of the regular exponential family, but may still be valid probability mass functions. Some of these cases were again presented: $t(n;\phi) = n^{a+1}, a > 0$ (Gelfand and Dalal [54]), $t(n) = e^n - 1$ (Lindsay [91]) and $t(n) = nln(n) - n$ (Efron [42]). This second group of functions is only defined when $r < 1$. The relationship between the parameter r, the convexity of $t(n;\lambda,\phi)$ and the dispersion of the distribution was derived and it was shown that for a convex $t(n;\lambda,\phi)$, $r < 0$ implies overdispersion and $r > 0$ implies underdispersion.

Other weight functions have been investigated that are of the same form as those proposed by Castillo and Perez [39]. Two of these were discussed by Kokonendi et al. [83]. They occur when $w(n;\lambda,\phi) = e^{r|n-\lambda|}$ and $w(n;\lambda,\phi) = \left(\frac{n}{\lambda}\right)^{rn} = e^{rI(n;\lambda)}$, where $I(n;\lambda)$ is the Kullback-Leibler distance (also known as the relative entropy). See Ridout and Besbeas [117] for these cases as well as the case where

$$w(n;\lambda,\phi) = \begin{cases} e^{-\beta_1(\lambda-n)} & n \leq \lambda, \\ e^{-\beta_2(n-\lambda)} & n > \lambda. \end{cases} \tag{2.8}$$

The origins of using the Kullback-Leibler distance as a weight function can be traced back to Efron [42]. As was shown in Kokonendi et al. [83], both of these weight functions result in distributions which can be over, under or equidispersed depending on whether $r > 0, r < 0$ or $r = 0$ respectively.

In addition to discussing some specific weight functions already mentioned in this section, Kokonendi et al. [83] also investigated many general properties of the weighted Poisson distribution. Specifically, they investigated how the shape of the weight function is related to the dispersion of the resulting weighted Poisson distribution. Their results can be seen as a generalisation of the Castillo and Perez [39] paper since the theorems in Kokonendi et al. [83] allow for arbitrary forms of weight functions. It is easiest to describe how the theorems in their paper relate to each other by using the following flow chart:

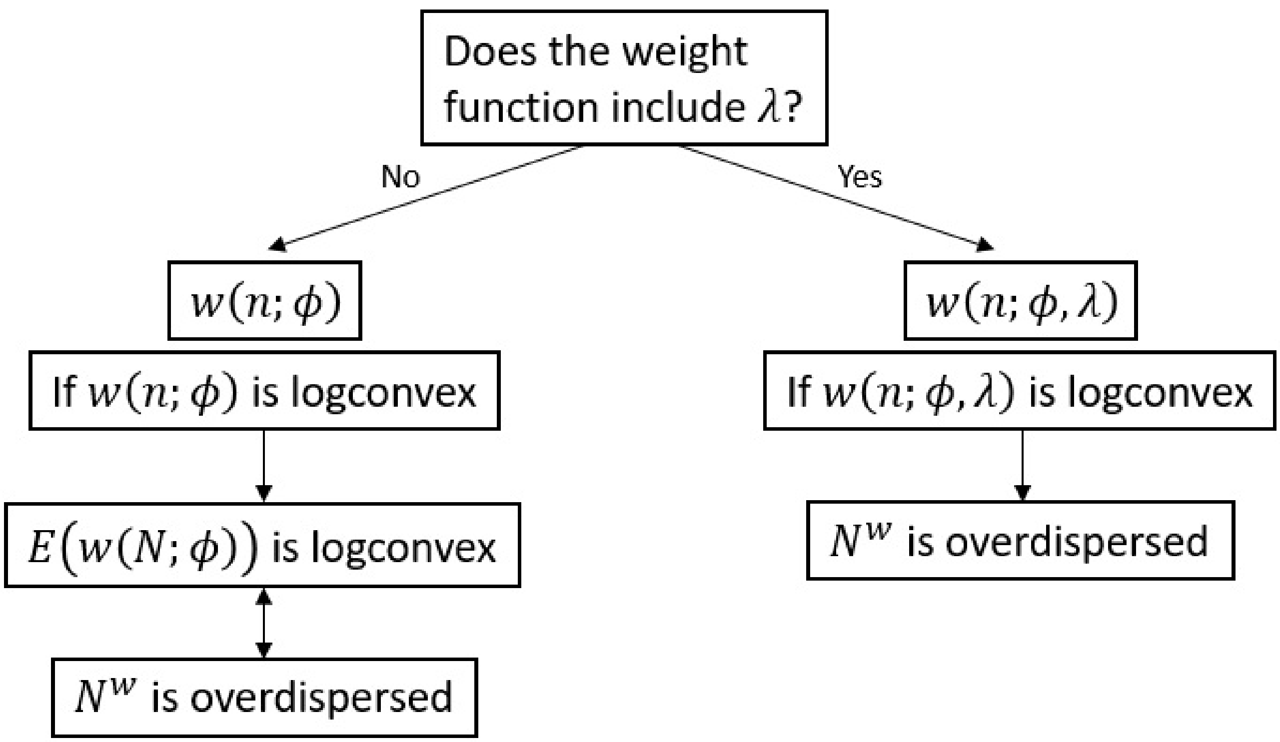


Figure 2.1: Relation between the weight function and dispersion

Note that, in the above theorems, if logconvexity is replaced with logconcavity, the theorems still hold, but with the resulting weighted Poisson distributions being underdispersed. Logconcavity and logconvexity are defined in Definition 10.13. One other important concept discussed by Kokonendi et al. [83] is the notion of "dual" weighted Poisson distributions. Two weighted Poisson distributions with weight functions $w_1(n)$ and $w_2(n)$ are said to be dual if $w_1(n)\, w_2(n) = 1 \ \forall n \in \mathbb{N}_0$. In essence this means that if one of the weighted Poisson distributions is overdispersed, the other will be equally underdispersed. Consequently, having a form of weight function that possesses a dual partner will guarantee that the resulting weighted Poisson distributions are able to model both under and overdispersion.

An important conclusion drawn from the above papers is that the dispersion of the weighted Poisson random variable is directly linked to the logconcavity/logconvexity of the relevant weight function.

Chapter 3

Poisson Distribution: Goodness-of-fit Testing

In this chapter, goodness-of-fit tests for the Poisson distribution will be researched. In Section 3.1, many of the tests currently available in the literate will be discussed. This is by no means a exhaustive list of all possible tests, but rather, focuses on some of the most commonly used and efficient methods as described in recent papers. Two of the most recent overview papers regarding Poisson goodness-of-fit are Gürtler and Henze [62] and Karlis and Xekalaki [76]. In Section 3.2, many of the different methods listed in Section 3.1 will be tested against a range of different non Poisson simulated datasets to determine the cases in which specific methods will perform the best and in which cases their performance may not be ideal. In order to test the applicability of the different tests in as many different scenarios as possible, the underlying data will be simulated out of a wide variety of discrete distributions, each with varying parameters. In addition to simulating out of commonly used distributions like the negative binomial and compound Poisson distributions, data will also be simulated out of various weighted Poisson distributions since no literature exists describing how Poisson goodness-of-fit tests fare against weighted Poisson alternatives.

Generally, goodness-of-fit tests can broadly be classified into one of three categories: The first can be referred to as "standard" goodness-of-fit tests. These tests can be applied to general distributions, and do not depend on the properties of a specific distribution. The majority of these tests are non-parametric (also called distribution-free), and as a result, are often very robust with respect to the underlying distribution of the data. An example of a commonly used "standard" goodness-of-fit test is the Pearson chi-squared test, which can be applied to any discrete distribution. The second is the group of tests that depend on a particular property of the assumed underlying distribution under the null hypothesis. This category of test is often more powerful than the first but does require that the underlying distribution of the data be known or assumed, which, if incorrect, can lead to faulty conclusions from the tests. The third category consists of tests where both the distributions under the null as well as the alternative hypotheses are assumed. These tests can often be the most powerful, but as a drawback also require that the most assumptions be made regarding the distributions. If either (or both) of the distributions are misidentified, these tests can give unreliable results.

In general, as the number of distributional assumptions increase, so too does the power of the test; but so too does the error of the test when the assumptions are violated. It should be noted that these three categories are not always as clear cut as they are presented above, and that tests do exist that do not fit nicely into a single category. For example, Lee [88] and Ghahfarokhi et al. [1] presented tests where the alternative hypothesis is not a singular distribution, but rather a family of distributions, which would technically fall between the second and third category as defined in this thesis.

3.1 Current tests

The process of conducting goodness-of-fit tests can be described as follows: A statistic is created based on a sequence of random variables. The complexity, underlying methodology and composition of these statistics can vary greatly, but in essence they all rely in some way on data. Once the statistic has been defined, its distribution must be established under the null hypothesis. This can be done in three different ways:

1. The theoretical distribution of the statistic can be derived - This process is the most theoretically taxing, but if expressions for the distribution of the statistic can be found the level of computation needed when testing the goodness-of-fit of a dataset may be significantly reduced. Actually deriving the distribution of a statistic may also lead to other insights regarding the performance and potential pitfalls of a certain statistic.

2. The asymptotic distribution of the statistic may be derived - In many practical cases small sample sizes are not a concern for researchers or practitioners. This has become especially true in the last decades as the sizes of datasets have increased dramatically. In these cases merely deriving the asymptotic distribution of a statistic may provide negligible differences from the theoretically derived distributions.

3. The values in the tails of the distribution can be simulated - This method requires the most initial processing power since tables of values have to be generated for varying percentiles with differing sample sizes and parameter choices. However, once these values have been calculated, applying the actual tests are rudimentary.

The method chosen can depend on many factors, both practical and theoretical. In this thesis the third method has been opted for. Once the percentiles in the tails of the distributions have been obtained (under the null hypothesis), the appropriate statistic can be calculated from a sequence of observations. If this calculated statistic falls outside of the relevant percentiles in the relevant table, the hypothesis that the data is distributed as assumed under the null hypothesis can be rejected.

What follows is a description of various different methods available in the literature that can be used to test for Poissonity.

3.1.1 Probability mass function tests

The chi-square test is arguably the most well known, non-parametric goodness-of-fit test that exists. The test's statistic is given by

$$\chi^2 \;=\; \sum_{j=0}^{k} \left(\frac{(O_j - E_j)^2}{E_j} \right) \tag{3.1}$$

where O_j is the number of actual observations equal to j, E_j is the expected number of observations equal to j under the null hypothesis, and k is the largest "bin" of the statistic. The chi-square statistic usually requires that observations be grouped into bins if the frequency of observations is very low, especially in the tails of the distribution. The choice of the size of the bins is usually the most problematic part of the chi-square test, where different choices of k can lead to varying conclusions of the test. Other problems include the fact that while the statistic is known to follow a chi-square distribution asymptotically, this is only true for large samples, and where each E_j value is larger than 5. Additionally, the test gives very little insight into what the actual distribution is if the null distribution does get rejected, although this is a problem with the vast majority of goodness-of-fit tests.

To overcome the problem of grouping data into bins Nass [99] proposed an alternate statistic

$$N \;=\; \frac{\sum_{j=0}^{m} \frac{O_j^2}{E_j} - n - m + 1}{\sqrt{\frac{m-1}{m}\left(2m - \frac{(m-1)^2 + 2m}{n} + \sum_{j=0}^{m} \frac{1}{E_j}\right)}} \tag{3.2}$$

which is asymptotically normally distributed. (Where n is the sample size and m is the largest observation.) While this statistic does overcome the binning problem, it does so at the cost of having a larger variance relative to the traditional chi-square test.

A range of similar tests exists in the literature. To name a few, these are the Neyman chi-square test [102], the Freeman-Tukey test [52] and the likelihood ratio test [73]. Read and Cressie [116] proposed a general test which includes all of these tests as special cases, their test statistic is given by

$$I^\lambda \;= \frac{1}{\lambda(\lambda+1)} \;\; \sum_{j=0}^{m} E_j \left(\left(\frac{O_j}{E_j} \right)^{\lambda+1} - 1 \right). \tag{3.3}$$

Beltran-Beltran and O'Reilly[12] proposed a ratio test based, not only on the number of observations O_j but also conditional on the sum of all the observations T. Their proposed statistic is given by

$$\Delta = \frac{\frac{T!}{\prod_{j=0}^{m}(j!)^{O_j}} \left(\frac{1}{n}\right)^T}{\frac{\prod_{j=0}^{m}(O_j!)}{n!}}. \tag{3.4}$$

Gonzalez-Barrios et al. [58] proposed a novel test based on the conditional probability mass function of the Poisson distribution. In essence their test considers the number of different ways in which a given sum of observations could have occurred, then calculates the probability that a specific realisation would come out of a Poisson distribution. Suppose that the realisation of a collection of random variables, $X_1, ..., X_n$ is ordered: $x_1 \leq ... \leq x_n$, and that $\sum_{j=1}^{n} x_j = t$. Then

$$\begin{aligned} P(x_1, ..., x_n|t) &= \text{(Number of permuations)} \begin{pmatrix} t \\ x_n \end{pmatrix} \begin{pmatrix} t - x_n \\ x_{n-1} \end{pmatrix} \\ &\times \begin{pmatrix} t - x_n - x_{n-1} \\ x_{n-2} \end{pmatrix} ... \begin{pmatrix} t - x_n - ... - x_2 \\ x_1 \end{pmatrix} \left(\tfrac{1}{n}\right)^t . \end{aligned} \tag{3.5}$$

By calculating all the probabilities of possible permutations such that $\sum_{j=1}^{n} x_j = t$, it is possible to determine which permutations have such a low likelihood of occurring that their presence would indicate likely deviance from the Poisson distribution. It should be noted that while this test works for small sums of observations, for high totals, this test becomes computationally impractical to implement.

3.1.2 Cumulative distribution function tests

Another commonly used goodness-of-fit test is the Kolmogorov-Smirnov test [84], the statistic of which is given by

$$KS = sup\left|F(x) - \hat{F}(x)\right| \tag{3.6}$$

where $F(x)$ is the cumulative distribution function under the null hypothesis, and $\hat{F}(x)$ is the empirical cumulative distribution function based on a sample. This test was initially proposed for continuous distributions but was later extended to discrete distributions by Conover [25]. Campbell and Oprian [19] developed an approximate Kolmogorov-Smirnov test specifically for the Poisson distribution, with a series of tables that could be applied to various parameter estimates of the Poisson distribution. Henze [65] overcame the need for tables by applying a bootstrap test. More recently, Frey [53] proposed an exact conditional Poisson Kolmogorov-Smirnov test based on the summation of the observations. The exact conditional Kolmogorov-Smirnov test statistic is given by

$$D = sup\left|\hat{F}(x) - E\left[\hat{F}(x)|T = t\right]\right|. \tag{3.7}$$

Klar [80] proposed using the sum of the absolute differences between the theoretical and empirical distribution functions, rather than the supremum as in the Kolmogorov-Smirnov test. Consequently, the following test statistic was proposed:

$$L = \sqrt{n}\sum_{j=1}^{n}\left|F(x) - \hat{F}(x)\right|. \tag{3.8}$$

Cramér and von Mises (CVM) [34, 140] proposed a non-parametric test that could detect deviation from a specific distribution. In their original papers, the statistic was proposed only for continuous distributions. Choulakain et al. [21] extended upon the definition to allow the test to apply to discrete distributions. Spinelli and Stephens[131] further improved on the research by looking at specific discrete distributions; additionally they also investigated modified versions of the statistic. Henze [65] also discusses a modified CVM type statistic. In all, there are four CVM statistics commonly discussed in the literature:

$$CVM_1 = \frac{1}{n}\sum_{j=0}^{M} Z_j^2 p_j \tag{3.9}$$

$$CVM_2 = \frac{1}{n}\sum_{j=0}^{M} \frac{Z_j^2 p_j}{F(j)(1-F(j))} \tag{3.10}$$

$$CVM_3 = \frac{1}{n}\sum_{j=0}^{M} Z_j^2 \tag{3.11}$$

$$CVM_4 = \frac{1}{n^2}\sum_{j=0}^{M} Z_j^2 O_j \tag{3.12}$$

where $Z_j = \sum_{i=0}^{j}(O_i - E_i)$ and p_j is the theoretical probability of an observation having value j. Statistic 3.9 is considered to be the "standard" CVM statistic. Statistic 3.10 is also known as the Anderson-Darling test. Statistic 3.11 gives more weight to deviations in the tails of the distributions relative to Statistic 3.9, and Statistic 3.12, the one proposed by Henze [65], uses observed relative frequencies instead of the theoretical probabilities in Statistic 3.9. The performance of these four statistics was investigated in-depth by Karlis and Xekalaki [76]. It should be noted that in all four CVM statistics, the theoretical upper summation limit is infinity, but for practical reasons, Karlis and Xekalaki [76] replaced this with a value M such that $p_M < 0.0001$.

3.1.3 Integrated distribution function tests

In addition to the test based on the cumulative distribution function Klar [80] also proposed a test that is based on the integrated distribution function. Define $\Psi(t) = \int_t^\infty (1 - F(x))\,dx$ as the integrated distribution function and $\hat{\Psi}_n(t) = \frac{1}{n}\sum_{j=1}^{n}(X_j - t)\, I_{(X_j > t)}$ as the empirical integrated distribution function, (where $I_{()}$ is the indicator function in Definition 10.14) then the proposed statistic is given by

$$IDF =_{sup_{t\geq 0}} \sqrt{n}\left|\Psi(t) - \hat{\Psi}_n(t)\right|. \tag{3.13}$$

Practically speaking, since the Poisson distribution is discrete, the integrated distribution function would be given by $\Psi(t) = \sum_{j=\lfloor t \rfloor+1} (j-t)\, p_j$ where $\lfloor t \rfloor$ is the integer part of t. In Gürtler and Henze [62] it was shown that this statistic performs very well in detecting deviations from the Poisson distribution, usually outperforming the Kolmogorov-Smirnov and CVM type tests.

3.1.4 Poisson moments tests

One of the most common ways in which Poisson goodness-of-fit tests are constructed is by using the fact that Poisson random variables are equidispersed. A vast number of tests have been proposed based on this fact.

One of the first papers where this equidispersion was noted is Fisher [51]. As a result, the Fisher index as defined in Equation 2.1, is a common statistic used to detect deviations from Poissonity. Note that some authors also define the Fisher index to be $(n-1)\frac{S^2}{\overline{X}}$. In this thesis, this will be referred to as the "variance test" instead. The difference between the two statistics is mostly inconsequential though since n is deterministic. However, Anderson and Siddiqui [3] did note that the Fisher index could be better modelled by the chi-square distribution than the variance test. Another reason the Fisher index is usually preferred over the variance test is because it gives a clear link between the statistic value and the over, under and equidispersion of data.

The variance test

$$VT = (n-1)\frac{S^2}{\overline{X}} \tag{3.14}$$

and Fisher index are typically two-sided tests, where either very large or small values would lead to the null hypothesis of Poissonity being rejected. Due to its widespread use as a goodness-of-fit statistic, the variance test has been the focus of considerable research. Selby [122], Anderson and Siddiqui [3], Bartko et al. [10], Dahiya and Gurland [35], Potthoff and Whittinghill [111], Collings and Margolin [24], Kim and Park [17], Perry and Mead [108] and Kharshikar [**?**] all studied the properties of either the variance test or the Fisher index. Bateman [11] and Darwin [36] studied the power of the variance test against a range of alternative distributions. A crucial caveat must be given regarding the use of both the Fisher index and variance tests; this being that the value of the test statistic has often been erroneously used as a diagnostic to determine if a distribution is over or underdispersed, where large statistic values are associated with overdispersion and small values with underdispersion. Henze and Klar [66], however, showed that this could lead to erroneous conclusions. As a result, they proposed a rescaled version of the variance test, which would only be rejected for large values of the statistic:

$$S^* = \frac{\overline{X}\,(VT-n)^2}{\sum_{j=1}^{n}\left(\left(X_j-\overline{X}\right)^2 - X_j\right)}. \tag{3.15}$$

Another rescaled version of the variance test was proposed by Rayner and McIntyre [115]

$$U^2 = \left(\frac{1}{\sqrt{2n}} \left(VT - n \right) \right)^2 \tag{3.16}$$

Böhning [14] and Potthoff and Whittinghill [111] proposed yet another statistic based on the Fisher index

$$O_2 = \sqrt{\frac{n-1}{2}} \left(FI - 1 \right). \tag{3.17}$$

This test was shown to be the locally most powerful against the negative binomial distribution.

Zelterman [145] proposed a slightly modified version of Statistic 3.17:

$$Z = \sqrt{\frac{n}{2}} \left(FI - 1 \right). \tag{3.18}$$

de Oliveira [37] attempted to model the difference between the sample variance and sample mean $\left(S^2 - \bar{X}\right)$. Consequently, they derived the statistic $O_T = \frac{\sqrt{n}\left(S^2 - \bar{X}\right)}{\sqrt{\left(1 - 2\sqrt{\bar{X}} + 3\bar{X}\right)}}$, and claimed that its limiting distribution under the null hypothesis is standard normal. Prompted by a simulation study, Böhning [71] demonstrated that they had made errors in their derivation of the statistic's variance, and that in fact the statistic should have been $O_T^{new} = \sqrt{\frac{n-1}{2}} \left(\frac{S^2}{\bar{X}} - 1 \right)$, which was noted to be very similar to the variance test, and is in fact exactly the same statistic that was proposed by Böhning [14] and Potthoff and Whittinghill [111].

Kyriakoussis et al. proposed another goodness-of-fit test that is based on the second product moment:

$$c = \frac{\frac{1}{n}\sum_{j=1}^{n} \left(X_j \left(X_j - 1 \right) \right)}{\left(\frac{1}{n}\sum_{j=1}^{n} \left(X_j \right) \right)^2}. \tag{3.19}$$

Rayner and McIntyre [115] proposed a test statistic which was designed to function well against the generalised Poisson distribution (see Consul [27])

$$W = \frac{n}{2} \left(FI - 1 \right)^2. \tag{3.20}$$

For the above tests it has been found that the variance test (Equation 3.14) is the locally most powerful unbiased test against a negative binomial alternative hypothesis, the Z-test (Equation 3.18) against a general central mixture alternative hypothesis and the W-test (Equation 3.20) against the generalised Poisson distribution alternative hypothesis.

Pettigrew and Mohler [109] proposed a test based on the moments of the Poisson distribution. Unlike the tests mentioned earlier in this subsection, this test could be implemented using higher order moments of the Poisson distribution. While it was suggested that lower order moments be used to reduce the variability of the statistic, it is possible to use higher order cumulants. Their proposed statistic is given by

$$Z_p = \frac{k_p - \overline{X}}{\sqrt{var(k_j|X)}} \quad , p = 2, 3, 4 \tag{3.21}$$

where k_p is the p^{th} sample cumulant and $var\,(k_j|X)$ is the variance of the p^{th} cumulant given the sum of the observations. For $p = 2$ this statistic reduces to

$$Z_2 = \frac{S^2 - \overline{X}}{\sqrt{2n\overline{X}\left(n\overline{X}-1\right)}} n\sqrt{n-1} \ . \tag{3.22}$$

Similarly for $p = 3$ and $p = 4$ the statistics become

$$Z_3 = \frac{m_3 - \overline{X}}{\sqrt{6n\overline{X}\left(n\overline{X}-1\right)\left(3+\frac{n\overline{X}-2}{n-2}\right)}} n\sqrt{n-1} \tag{3.23}$$

$$Z_4 = \frac{m_4 - 3S^2 - \overline{X}}{\sqrt{2n\overline{X}\left(n\overline{X}-1\right)\left(49+\frac{108\left(n\overline{X}-2\right)}{n-2}+\frac{12(n+1)\left(n\overline{X}-2\right)\left(n\overline{X}-3\right)}{n(n-2)(n-2)}\right)}} n\sqrt{n-1} \tag{3.24}$$

with $m_p = \sum_{j=1}^{n} \frac{(X_j - \overline{X})^p}{n}$.

Another test which depends on higher order moments was proposed by Gupta et al. [61]

$$M = \frac{1}{2}\sqrt{\frac{n}{1+24\overline{X}+6\overline{X}^2}} \frac{m_2\left(m_4 - 3m_2^2\right) - m_3}{\overline{X}^2}. \tag{3.25}$$

Cox [33] proposed a test based on the log of the probability mass function of the Poisson distribution. The proposed test statistic is given by

$$C = \sum_{j=1}^{n} \left(\left(\frac{\partial ln\left(f\left(x_j;\lambda\right)\right)}{\partial\lambda} \right)^2 + \frac{\partial^2 ln\left(f\left(x_j;\lambda\right)\right)}{\partial\lambda^2} \right). \tag{3.26}$$

3.1.5 Probability generating function tests

Kocherlakota and Kocherlakota [81] proposed a goodness-of-fit test for discrete distributions based on the probability generating function. The probability generating function is defined as $g\left(t\right) = E\left(t^X\right) = \sum_{j=0}^{\infty} P\left(X = j\right) t^j$, and it's empirical equivalent as $g_n\left(t\right) = \frac{1}{n}\sum_{j=1}^{n} t^{X_j}$. The statistic proposed by Kocherlakota and Kocherlakota is $T = \sqrt{n}\left(g_n\left(t\right) - g\left(t\right)\right)$. Which, after taking into account the Poissonity of the null distribution. equates to

$$K = \sqrt{n}\frac{\phi_n\left(t\right) - e^{\bar{x}(t-1)}}{e^{\bar{x}(t^2-1)} - e^{2\bar{x}(t-1)}\left(1+\bar{x}\left(t-1\right)^2\right)}. \tag{3.27}$$

This test does have the drawback that it depends on a specific choice of t, and although the statistic is not very sensitive with respect to the parameter, it does require that some value for t be chosen.

To overcome this problem Rueda et al. [121] and Rueda and O'Reilly [120] proposed using the square of Statistic 3.27, integrated over 0 to 1 with respect to t, and thus the newly proposed statistic becomes $R_n = \int_0^1 \left(\sqrt{n}\left(g_n\left(t\right) - g\left(t\right)\right)\right)^2 dt$. After simplification this statistic is

$$R = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{X_i + X_j + 1} - 2e^{-\lambda}\sum_{i=1}^{n}T\left(X_i, \lambda\right) + n\frac{1 - e^{-2\lambda}}{2\lambda} \tag{3.28}$$

where $T\left(x, \lambda\right) = \int_0^1 t^x e^{\lambda t} dt$.

Baringhaus et al. [7] further generalised on the statistic by adding a weight function, and their resulting statistic is given by

$$R_a = \int_0^1 \left(\sqrt{n}\left(g_n\left(t\right) - g\left(t\right)\right)\right)^2 t^a dt \tag{3.29}$$

which is equivalent to

$$R_a \approx n\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\left(\frac{\left(f_n\left(i\right) - f\left(i\right)\right)\left(f_n\left(j\right) - f\left(j\right)\right)}{i + j + a + 1}\right). \tag{3.30}$$

Additionally, Baringhaus and Henze [8] proposed a goodness-of-fit test based on a unique property of the Poisson probability generating function, this being that $\frac{\partial}{\partial t}g\left(t\right) = \lambda g\left(t\right)$. Consequently their proposed statistic is $T = n\int_0^1 \left(\overline{X}g_n\left(t\right) - g_n'\left(t\right)\right)^2 dt$ which again simplifies to

$$B = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\overline{X}^2}{X_i + X_j + 1} + \frac{X_i X_j}{X_i + X_j - 1}\right) - \bar{X}\left(n - f_0\right). \tag{3.31}$$

Similar to Statistic 3.29, Statistic 3.31 was also generalised with the addition of a weight function by Treutler [138] with the result being the proposed statistic $T = n\int_0^1 \left(\overline{X}g_n\left(t\right) - g_n'\left(t\right)\right)^2 t^a dt$ which simplifies to

$$T = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\overline{X}^2}{X_i + X_j + a + 1} - \frac{\overline{X}\left(X_i + X_j\right)}{X_i + X_j + a} + \frac{X_i X_j}{X_i + X_j + a - 1}\right). \tag{3.32}$$

Nakamura and Perez-Abreu [98] also utilised the unique derivative properties of the Poisson probability generating function by noting that $\frac{\partial^2}{\partial t^2}ln\left(g\left(t\right)\right) = 0$. This led them to propose the following test statistic

$$V = \frac{1}{n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\left(X_i\left(X_i - X_j - 1\right)X_k\left(X_k - X_l - 1\right)I_{\left(X_i + X_j = X_k + X_l\right)}\right), \tag{3.33}$$

which they then modified to

$$V^* = \left(\frac{V_n}{\overline{X}^{1.45}}\right). \tag{3.34}$$

One of the particularly interesting and useful aspects of the Nakamura and Perez-Abreu statistic is that it is asymptotically independent of λ. It should be noted, however, that the four-fold summation is extremely computationally intensive, and for practical purposes, an alternate expression, found in both Nakamura and Perez-Abreu [98] and Gürtler and Henze [62] is implemented in Section 3.2.

Meintanis and Nikitin [94] constructed a test similar to the one proposed by Baringhaus and Henze [8]. Their statistic is given by

$$T_a^* = \frac{1}{n}\sum_{j=1}^{n}\left(\frac{X_j}{X_j + a} + \frac{\overline{X}}{X_j + a - 1}\right). \tag{3.35}$$

3.1.6 Other tests

Rayner and Best [13, 114] proposed the use of Poisson-Charlier polynomials to test the goodness-of-fit for a Poisson distribution

$$S_k = \sum_{i=2}^{k}\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(\sqrt{\frac{\hat{\lambda}^i}{i!}}\sum_{k=0}^{i}(-1)^{i-k}\begin{pmatrix} i \\ k \end{pmatrix}\frac{k!}{\hat{\lambda}^k}\begin{pmatrix} X_j \\ k \end{pmatrix}\right)\right)^2. \tag{3.36}$$

See Ledwina and Wyłupek [87] for a followup paper on Charlier polynomials.

Szekely and Rizzo [135] proposed a test based on the mean distance between a Poisson random variable X and some integer value k. Their proposed test statistic is the same as CVM1 (equation 3.9), however, instead of using $Z_j = \sum_{i=0}^{j}(O_i - E_i)$ they proposed constructing the empirical cumulative mass function out of a recursive formula given by

$$\hat{f}(k) = \frac{\hat{m}_{k+1} - \left(k + 1 - \hat{\lambda}\right)\left(2\hat{F}_X(k-1) - 1\right)}{2(k+1)}. \tag{3.37}$$

The statistic is given by

$$SR = n\sum_{j-0}^{\infty}\left(\hat{F}(j) - F_{\hat{\lambda}}(j)\right)^2 f_{\hat{\lambda}}(j). \tag{3.38}$$

Another test that has been widely implemented (see [137]), not only in the Poisson case, is the likelihood ratio test. The statistic for this test is given by

$$L^* = 2(L_1 - L_0) \tag{3.39}$$

where L_0 and L_1 are the maximised log-likelihoods under the respective hypotheses.

Karlis and Xekalaki [75] proposed a test similar to the likelihood ratio test, but which is based on the minimum Hellinger distance rather than the maximised log-likelihoods. Their proposed statistic is given by

$$HDT = 4n\left(HD_0 - HD_1\right) \tag{3.40}$$

where HD_0 and HD_1 are the minimised Hellinger distances under the two respective hypotheses.

Many graphical methods have also been developed to detect deviations from the Poisson distribution. See Lindsay [91] and Lindsay and Roeder [92] for two graphical examples, and Karlis and Xekalaki [76] for an extensive list of other authors who used or proposed graphical tests.

Cameron and Trivedi [18] in addition to discussing a compound Poisson alternative that can be used to model count data (where the parameter λ was assumed to follow some gamma distribution (see Section 2.1), also proposed tests for Poissonity. Their family of tests can be summarised as follows:

$$\begin{aligned} H_0: &\quad X_i \sim Poisson\left(\lambda\right) \\ H_A: &\quad E\left(X_i\right) = \lambda \\ &\quad Var\left(X_i\right) = \lambda + a\lambda^b. \end{aligned}$$

3.2 Test comparison

As has already been mentioned, not all of the tests discussed in Section 3.1 will be implemented in the computational comparison between the statistics. One reason for this is the sheer amount of tests that were mentioned. More importantly, however, is the fact that many of the tests have been shown not to be extremely powerful when tested against a general range of alternative distributions. Some tests, like a few of those based on the Poisson moments, may be powerful, even the most powerful, against a specific distribution, but perform poorly against more general alternatives. Taking this into account, the tests that will be considered are:

- W: Rayner and McIntyre's test based on the variance test (Equation 3.16).
- KS: The Kolmogorov-Smirnov test (Equation 3.6).
- CVM_1: The Cramér von Mises test (Equation 3.9).
- L: Klar's test based on the differences between the theoretical and empirical distribution functions (Equation 3.8).
- IDF: Klar's test based on the integrated distribution function (Equation 3.13).

- R: The Rueda et al. test based on the probability generating function (Equation 3.28).
- T: Treutler's test based on the probability generating function (Equation 3.32).
- V^*: Nakamura and Perez-Abreu's test based on the probability generating function (Equation 3.34).
- SR : Szekely and Rizzo's test based on mean distances (Equation 3.38).
- T_a^* : Meintanis and Nikitin's test (Equation 3.35).

The performance of the above statistics will be tested against a range of alternative distributions, with various parameter combinations. The distributions are: The Poisson distribution $Poisson\left(\lambda\right)$, the discrete uniform distribution $Unif\left(a,b\right)$, the negative binomial distribution $NegBin\left(r,p\right)$, the mixed Poisson distribution $MixPoi\left(p,\lambda_1,\lambda_2\right) = pPoisson\left(\lambda_1\right) + \left(1-p\right)Poisson\left(\lambda_2\right)$, the generalised Poisson distribution $GenPoi\left(\lambda_1,\lambda_2\right)$, the zero-inflated Poisson distribution $ZiPoi\left(\lambda,\varepsilon\right)$ (see Theorem 4.17) and the weighted Poisson distribution $WeiPoi\left(\lambda,a,b\right)$ (with weight function $w\left(n;\phi\right) = an^2+bn+1$). While some of these distributions have been used before in similar comparison tests, weighted Poisson distributions have never been selected as the alternative distribution.

For each test-distribution combination the sample size will also be varied. The samples sizes that are considered are $n = 30, 50, 100, 200$.

Traditionally, a parametric bootstrap procedure is used in order to evaluate the performance of these tests, see Gürtler and Henze [62]. The process can be described as follows:

- Define W_n to be the statistic that will be tested (which depends on a sample of n observations), $W_n\left(X_1,X_2,...,X_n\right)$
- Let $H_{n,\lambda}\left(t\right) = P\left(W_n \le t\right)$ be the distribution function of the null distribution of W_n. (In other words $H_{n,\lambda}\left(t\right)$ is the distribution of the statistic if the underlying distribution is indeed $Poisson\left(\lambda\right)$ distributed.)
- The critical value(s), c, for this test will then be the $\left(1-\alpha\right)^{th}$ quantile of $H_{n,\lambda}\left(t\right)$ for a one sided test or the $\frac{\alpha}{2}^{th}$ and $\left(1-\frac{\alpha}{2}\right)^{th}$ quantiles of $H_{n,\lambda}\left(t\right)$ for a two sided test.

This percentile will be estimated using the following Monte Carlo procedure:

- Generate a sample of size n from a $Poisson\left(\lambda\right)$ distribution, $\left(X_1,X_2,...,X_n\right)$.
- Using this sample, calculate the maximum likelihood estimate of λ, $\hat{\lambda} = \frac{\sum_{i=1}^n X_i}{n}$.
- Generate M random samples of size n from a $Poisson\left(\hat{\lambda}\right)$ distribution, $\left(X_{j1}^*,X_{j2}^*,...,X_{jn}^*\right), j = 1,2,...M$.
- Calculate the statistic value for each sample, $W_{j,n}^* = W_n\left(X_{j1}^*,X_{j2}^*,...,X_{jn}^*\right), j = 1,2,...M$.

- Based on the set of M statistic values, calculate the empirical distribution function, $H^*_{n,M}(t)$.
- Calculate the critical value(s), $c^*_{n,M}$, of $H^*_{n,M}(t)$. (Baringhaus and Henze [8] found this critical value to be sufficiently accurate only in cases where $M \geq 5n$.)

In general, the above procedure would have to be repeated many (say B) times, and the final critical value would be set as the average of the B $c^*_{n,M}$ values. This can be extremely computationally intensive. However, a much less intensive method was developed by Giacomini et al. [55], which they named the "warp speed" bootstrap method, which significantly reduces the computational time needed to calculate the critical values. In essence their method is the same as the one described above, with the main difference being that $M = 1$. It was found that this method is dramatically faster to calculate, and gives near identical results to the traditional parametric bootstrap approach. For more detail on the implementation of this warp speed approach see Allison et al. [2], and Henze and Klar [67].

The results from the above procedure are given in tables 3.1 to 3.4, where the test values reported are the various percentages that the different tests will result in rejection of the null hypothesis.

While it is impossible to give a set of definitive recommendations on which test will always perform the best, the tables below do give some insights into general situations when the tests might perform well. Some interesting points to note:

- As expected, the tests perform better the more the dispersion of the alternative distributions deviate from equidispersion.
- Treutler's (T) and Meintanis and Nikitin's (T^*_a) tests (both based on the probability generating functions) appear to perform particularly well when data is underdispersed and sample sizes are relatively small, however, both tests are much less competitive options when data is overdispersed (Treutler's test more so than Meintanis and Nikitin's)
- The "historic" tests (W, KS, CVM_1) perform relatively well at detecting deviations from Poissonity when the alternative data is overdispersed (irrespective of sample size), with the only "more modern" tests giving consistent good performance being Nakamura and Perez-Abreu's (V^*) and Meintanis and Nikitin's (T^*_a) tests. It should be noted, however, that the power of the Nakamura and Perez-Abreu test performed relatively poorly for larger sample sizes.

Alternative Distribution	Dispersion of Alternative Distribution	W	KS	CVM_1	L	IDF	R	T	V^*	SR	T_a^*
$Poisson\,(0.5)$	1.00	5	4	4	5	5	4	4	5	5	5
$Poisson\,(1)$	1.00	5	5	5	5	5	4	5	5	5	5
$Poisson\,(5)$	1.00	5	5	5	5	5	5	5	5	5	5
$Poisson\,(10)$	1.00	5	5	5	5	5	5	5	5	5	5
$Uni\,(0,4)$	1.00	2	6	6	32	19	34	10	43	32	4
$WeiPoi\,(1,1,-1)$	1.00	4	6	6	13	11	10	6	10	13	5
$Uni\,(0,2)$	0.67	18	18	18	28	29	33	38	29	28	19
$WeiPoi\,(1,2,1)$	0.68	22	27	27	20	8	20	26	6	19	30
$WeiPoi\,(1,1,1)$	0.75	12	16	17	13	5	12	16	4	13	18
$Bin\,(4,0.25)$	0.75	11	15	15	13	4	11	14	2	13	15
$Bin\,(20,0.25)$	0.75	12	12	11	12	6	11	15	4	11	14
$WeiPoi\,(1,1,0)$	0.83	6	8	8	10	4	8	10	3	10	9
$Bin\,(10,0.1)$	0.90	4	5	5	6	3	5	6	2	6	6
$Bin\,(50,0.1)$	0.90	5	5	5	6	4	5	6	3	6	6
$MixPoi\,(0.01,1,5)$	1.03	6	7	7	5	5	6	5	6	5	6
$NegBin\,(9,0.9)$	1.11	10	8	8	6	9	7	7	10	6	8
$NegBin\,(45,0.9)$	1.11	9	9	8	6	7	8	7	9	6	7
$MixPoi\,(0.05,1,5)$	1.16	12	18	18	7	7	10	10	11	7	15
$Uni\,(0,5)$	1.17	8	26	27	41	31	47	22	56	40	21
$GenPoi\,(4,0.1)$	1.24	17	16	15	8	11	13	13	16	8	13
$MixPoi\,(0.5,3,5)$	1.25	18	17	17	9	13	15	14	17	10	16
$ZiPoi\,(3,0.1)$	1.30	23	33	33	14	26	21	19	24	14	29
$NegBin\,(15,0.75)$	1.33	26	24	23	12	16	20	20	23	12	21
$NegBin\,(3,0.75)$	1.33	25	20	21	13	21	18	18	25	14	21
$Uni\,(0,6)$	1.33	24	54	54	54	44	62	43	68	53	47
$NegBin\,(2,0.667)$	1.50	39	34	34	23	34	28	31	39	23	36
$NegBin\,(10,0.667)$	1.50	42	39	38	20	25	34	34	36	19	35
$ZiPoi\,(3,0.2)$	1.60	55	71	71	44	70	56	55	59	44	67
$MixPoi\,(0.5,2,5)$	1.64	57	57	57	35	41	49	52	52	35	55
$NegBin\,(1,0.5)$	2.00	69	62	63	53	64	56	58	67	53	65

Table 3.1: Test comparison, $n = 30$

Alternative Distribution	Dispersion of Alternative Distribution	W	KS	CVM_1	L	IDF	R	T	V^*	SR	T_a^*
$Poisson\,(0.5)$	1.00	5	5	5	5	5	5	5	5	5	5
$Poisson\,(1)$	1.00	5	5	5	5	5	5	5	5	5	5
$Poisson\,(5)$	1.00	5	5	5	5	5	5	5	5	5	5
$Poisson\,(10)$	1.00	5	5	5	5	5	5	5	5	5	5
$Uni\,(0,4)$	1.00	1	9	9	54	33	60	18	74	53	6
$WeiPoi\,(1,1,-1)$	1.00	4	7	7	21	18	16	8	15	21	6
$Uni\,(0,2)$	0.67	37	29	31	52	53	59	67	75	52	33
$WeiPoi\,(1,2,1)$	0.68	40	45	46	30	13	34	41	15	30	47
$WeiPoi\,(1,1,1)$	0.75	23	26	27	19	7	21	25	8	19	28
$Bin\,(4,0.25)$	0.75	18	22	23	18	5	18	22	5	18	25
$Bin\,(20,0.25)$	0.75	21	20	19	17	7	18	23	8	17	24
$WeiPoi\,(1,1,0)$	0.83	10	11	11	13	5	12	14	6	13	11
$Bin\,(10,0.1)$	0.90	5	7	7	7	3	6	7	2	7	7
$Bin\,(50,0.1)$	0.90	6	6	6	6	4	6	7	4	7	7
$MixPoi\,(0.01,1,5)$	1.03	6	7	7	5	5	6	5	6	5	6
$NegBin\,(9,0.9)$	1.11	11	9	9	7	10	8	8	10	7	9
$NegBin\,(45,0.9)$	1.11	10	10	9	7	8	9	9	10	6	9
$MixPoi\,(0.05,1,5)$	1.16	15	23	23	8	8	13	12	12	8	20
$Uni\,(0,5)$	1.17	11	42	41	66	48	74	40	83	66	34
$GenPoi\,(4,0.1)$	1.24	23	21	20	11	14	18	19	19	11	19
$MixPoi\,(0.5,3,5)$	1.25	24	23	22	13	16	20	21	21	13	22
$ZiPoi\,(3,0.1)$	1.30	31	47	47	22	38	31	30	32	24	42
$NegBin\,(15,0.75)$	1.33	36	32	32	17	21	28	30	29	17	30
$NegBin\,(3,0.75)$	1.33	35	30	30	19	30	25	27	32	19	31
$Uni\,(0,6)$	1.33	37	75	74	79	64	87	66	91	79	68
$NegBin\,(2,0.667)$	1.50	54	48	49	35	47	42	44	50	34	48
$NegBin\,(10,0.667)$	1.50	57	54	54	30	35	47	50	48	30	52
$ZiPoi\,(3,0.2)$	1.60	75	89	89	68	90	79	78	79	67	86
$MixPoi\,(0.5,2,5)$	1.64	77	77	76	55	61	71	74	70	55	75
$NegBin\,(1,0.5)$	2.00	86	81	82	74	82	77	78	82	73	83

Table 3.2: Test comparison, $n = 50$

Alternative Distribution	Dispersion of Alternative Distribution	W	KS	CVM_1	L	IDF	R	T	V^*	SR	T_a^*
$Poisson\,(0.5)$	1.00	5	5	5	5	5	5	5	5	5	5
$Poisson\,(1)$	1.00	5	5	5	5	5	5	5	5	5	5
$Poisson\,(5)$	1.00	5	5	5	5	5	5	5	5	5	5
$Poisson\,(10)$	1.00	5	5	5	5	5	5	5	5	5	5
$Uni\,(0,4)$	1.00	1	18	16	88	68	94	45	99	88	10
$WeiPoi\,(1,1,-1)$	1.00	4	10	9	40	32	35	13	29	40	7
$Uni\,(0,2)$	0.67	82	60	62	92	97	97	98	100	92	69
$WeiPoi\,(1,2,1)$	0.68	73	78	78	57	27	67	74	43	57	78
$WeiPoi\,(1,1,1)$	0.75	47	51	51	34	14	41	46	21	34	51
$Bin\,(4,0.25)$	0.75	45	46	46	35	11	36	43	17	35	49
$Bin\,(20,0.25)$	0.75	45	41	40	30	13	36	44	22	30	45
$WeiPoi\,(1,1,0)$	0.83	20	18	18	22	6	21	24	12	22	20
$Bin\,(10,0.1)$	0.90	8	10	10	8	3	8	9	3	8	10
$Bin\,(50,0.1)$	0.90	9	9	9	8	4	8	10	5	8	10
$MixPoi\,(0.01,1,5)$	1.03	6	8	8	6	5	6	6	6	5	7
$NegBin\,(9,0.9)$	1.11	15	12	12	9	14	10	11	13	8	12
$NegBin\,(45,0.9)$	1.11	14	13	13	8	10	11	12	12	8	12
$MixPoi\,(0.05,1,5)$	1.16	22	35	36	11	11	20	19	18	11	31
$Uni\,(0,5)$	1.17	19	72	70	95	81	98	74	99	95	61
$GenPoi\,(4,0.1)$	1.24	36	32	32	17	22	29	31	27	18	31
$MixPoi\,(0.5,3,5)$	1.25	40	37	37	22	26	33	36	31	22	36
$ZiPoi\,(3,0.1)$	1.30	51	73	72	38	68	55	54	52	39	68
$NegBin\,(15,0.75)$	1.33	57	53	52	30	35	47	50	45	30	52
$NegBin\,(3,0.75)$	1.33	54	49	50	35	47	41	46	47	34	50
$Uni\,(0,6)$	1.33	63	96	95	98	93	100	94	100	98	93
$NegBin\,(2,0.667)$	1.50	79	74	75	60	72	68	71	71	61	75
$NegBin\,(10,0.667)$	1.50	83	80	80	53	58	73	77	71	54	78
$ZiPoi\,(3,0.2)$	1.60	95	99	99	94	100	98	98	97	94	99
$MixPoi\,(0.5,2,5)$	1.64	96	96	96	86	88	94	95	92	86	96
$NegBin\,(1,0.5)$	2.00	99	98	98	96	98	96	97	97	96	98

Table 3.3: Test comparison, $n = 100$

Alternative Distribution	Dispersion of Alternative Distribution	W	KS	CVM_1	L	IDF	R	T	V^*	SR	T_a^*
$Poisson\,(0.5)$	1.00	5	5	5	5	5	5	5	5	5	5
$Poisson\,(1)$	1.00	5	5	5	5	5	5	5	5	5	5
$Poisson\,(5)$	1.00	5	5	5	5	5	5	5	5	5	5
$Poisson\,(10)$	1.00	5	5	5	5	5	5	5	5	5	5
$Uni\,(0,4)$	1.00	1	37	33	100	99	100	93	100	100	20
$WeiPoi\,(1,1,-1)$	1.00	4	15	14	74	59	70	28	57	74	10
$Uni\,(0,2)$	0.67	100	93	94	100	100	100	100	100	100	96
$WeiPoi\,(1,2,1)$	0.68	96	97	98	88	57	94	96	84	88	98
$WeiPoi\,(1,1,1)$	0.75	79	81	82	63	29	72	78	54	63	83
$Bin\,(4,0.25)$	0.75	81	79	79	64	26	69	75	47	64	80
$Bin\,(20,0.25)$	0.75	80	73	73	55	27	68	76	53	55	76
$WeiPoi\,(1,1,0)$	0.83	40	32	33	40	12	40	44	30	40	36
$Bin\,(10,0.1)$	0.90	15	16	16	12	3	13	15	6	13	17
$Bin\,(50,0.1)$	0.90	15	15	15	11	5	13	16	8	11	16
$MixPoi\,(0.01,1,5)$	1.03	7	9	9	6	5	6	6	6	5	8
$NegBin\,(9,0.9)$	1.11	21	19	19	12	19	16	17	17	12	19
$NegBin\,(45,0.9)$	1.11	21	19	19	11	14	16	18	15	11	17
$MixPoi\,(0.05,1,5)$	1.16	35	56	56	18	19	33	33	29	18	53
$Uni\,(0,5)$	1.17	36	95	94	100	99	100	98	100	100	89
$GenPoi\,(4,0.1)$	1.24	58	53	53	32	36	48	52	43	31	53
$MixPoi\,(0.5,3,5)$	1.25	65	61	61	40	44	56	61	51	40	61
$ZiPoi\,(3,0.1)$	1.30	79	95	94	70	94	85	84	81	71	93
$NegBin\,(15,0.75)$	1.33	83	79	79	54	57	74	77	69	53	79
$NegBin\,(3,0.75)$	1.33	80	77	78	61	73	69	73	69	61	78
$Uni\,(0,6)$	1.33	91	100	100	100	100	100	100	100	100	100
$NegBin\,(2,0.667)$	1.50	96	95	95	89	93	92	94	92	88	96
$NegBin\,(10,0.667)$	1.50	98	97	97	84	85	95	96	93	84	97
$ZiPoi\,(3,0.2)$	1.60	100	100	100	100	100	100	100	100	100	100
$MixPoi\,(0.5,2,5)$	1.64	100	100	100	99	99	100	100	100	99	100
$NegBin\,(1,0.5)$	2.00	100	100	100	100	100	100	100	100	100	100

Table 3.4: Test comparison, $n = 200$

Chapter 4

Weighted Poisson Distribution: Theory

In this chapter, various weight functions will be discussed. While other authors have already investigated some of these weight functions, the vast majority are novel. (It should be noted that an infinite number of possible weight functions can be constructed. Consequently, this chapter will only focus on weights that result in closed-form expressions for the weighted Poisson distribution.)

The weights that are included in this thesis are a small subsection of weights that were considered during the research process. They were chosen for a few reasons:

- They could be classified into various overarching families of weights where setting specific parameters equal to 0 would often result in the weighted Poisson distribution reducing to another family member.
- The weights resulted in weighted Poisson distributions that could potentially have niche applications.
- The weights had a "good" trade-off between predictive ability while remaining relatively parsimonious.

As was mentioned previously, one of the main advantages of using the weighted Poisson distribution instead of the original Poisson distribution is that it overcomes the inherent restriction of equidispersion. As will become apparent in this chapter, the weighted Poisson distribution can offer some other interesting properties that could potentially be useful in applications.

The first of these useful properties, that has received considerable attention in the literature, is the concept of the zero-inflated (also known as the zero-modified) Poisson distribution. Real datasets often exhibit an excess of zero counts relative to what a fitted Poisson distribution predicts, and the zero-inflated Poisson distribution is one of the methods that can be employed to overcome this problem. While the zero-inflated Poisson distribution has been described in the framework of weighted Poisson distributions before (see Castillo and Perez [39]), it is usually presented in a way which either results in needlessly complicated expressions, or otherwise in distributions that are piece-wise defined.

Another useful property demonstrated by the weighted Poisson distribution is its ability to model truncated data. This property has only been explored to a minimal extent in previous studies. The most notable distribution that falls into this category is the zero-truncated Poisson distribution (see Cohen [23], and Dietz and Böhning [40] for details and applications). Some of the weighted Poisson distributions in this thesis will explore the modelling of truncated data in much greater detail. The proposed weight functions will allow for arbitrary points of lower and/or upper truncation, which have not been explored to data. These distributions have a wide range of potential applications, in multiple fields of insurance and actuarial modelling.

Furthermore, the weighted Poisson distribution is also capable of modelling multimodal data. These situations are most commonly modelled using mixture distributions; however, the weighted Poisson distribution gives a more unified and coherent way to model these situations.

The remainder of this chapter will be subdivided into sections, each representing a specific family of weight functions. For each weight function, specific properties will be derived. These include the normalising constant that appears in the probability mass function, the probability mass function, the probability generating function, the expected value, as well as the variance. Additionally, parameter and domain restrictions will be provided in each case. Each distribution will also be accompanied by a series of plots of the probability mass function with a fixed expected value of 10 and varying parameter choices to demonstrate potential shapes. These graphs are included in the hope of facilitating a better understanding of the various shapes that the mass functions may take on, as well as to give insight into some of the novel applications mentioned previously. All plots will be superimposed over a $Poisson\,(10)$ probability mass function (in black) to provide a reference distribution. Overdispersed weighted Poisson plots will be shown in red and underdispersed in blue. The discrete points of the various probability mass functions will be joined by lines to avoid confusion, which may otherwise occur due to the number of overlain graphs. This is done purely for aesthetic reasons. These distributions should not be confused with continuous distributions.

For each weight function, the format is similar, with each weight function being stated, followed by the resulting properties of the weighted Poisson distribution, and lastly by graphs of the various probability mass functions. Proofs for some of the weighted Poisson properties will be presented in this chapter, but the vast majority of them can be found in Chapter 10.

While the importance of stating and deriving expressions for the various weighted Poisson distributions may initially seem questionable, there are very real theoretical and practical reasons for doing so. From a theoretical perspective it not only enables insight into the various different data structures that each distribution can accommodate, but also leads to some additional restrictions on parameters which would otherwise not be obvious. From a practical perspective, having closed form representations for the various expressions result in significant time and computational savings when the distributions are fit to observed data.

4.1 Polynomial weight functions

As has already been mentioned in this thesis (Section 2.2), one commonly used weight function is $w(n) = n$. Here this weight function is expanded into a family of a several different functions based on polynomials.

4.1.1 $w(n) = n$

Theorem 4.1. *If the weight function in Equation 2.7 is* $w(n) = n$ *then*

$$\begin{aligned} E(w(N;\phi)) &= \lambda. \\ f_w(n) &= \tfrac{e^{-\lambda}\lambda^{n-1}}{(n-1)!}. \\ g(z) &= e^{\lambda(z-1)}z. \\ E(N^w) &= \lambda+1. \\ Var(N^w) &= \lambda. \end{aligned}$$

Restrictions:

- Domain: $n \in \mathbb{N}_1$.
- Parameters: $\lambda > 0$.

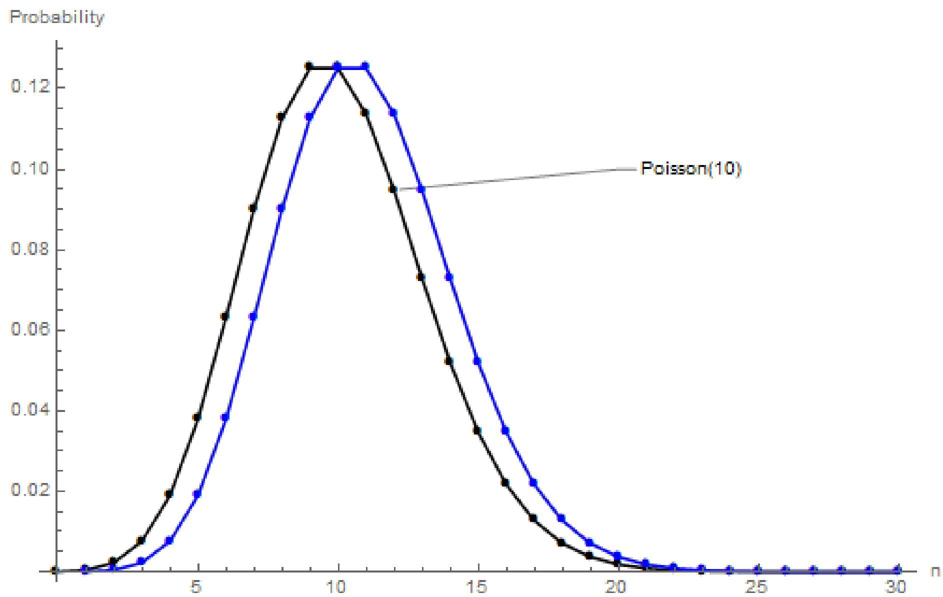


Figure 4.1: Probability mass function - $w(n) = n$

The variance associated with the weighted Poisson probability mass function in the above plot is 9. The domain of this weighted Poisson distribution does not include $n = 0$. This is called a "zero truncated", "shifted" or "one translated" Poisson distribution, but should not be confused with the "zero truncated Poisson distribution" (See Singh [129]). Specifically, a "zero truncated" Poisson distribution is one that cannot have observations of zero. A few weight functions that will be discussed in this chapter result in truncated distributions.

4.1.2 $w(n;\phi) = n^{-a}$

Theorem 4.2. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = n^{-a}$ *then*

$$
\begin{aligned}
E(w(N;\phi)) &= \lambda e^{-\lambda}{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda).\\
f_w(n) &= \tfrac{\lambda^{n-1}}{n^a n!\,{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)}.\\
g(z) &= \tfrac{z\,{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda z)}{{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)}.\\
E(N^w) &= \tfrac{{}_aF_a(1,...,1;2,...,2;\lambda)}{{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)}.\\
Var(N^w) &= \tfrac{{}_{a-1}F_{a-1}(1,...,1;2,...,2;\lambda)-{}_aF_a(1,...,1;2,...,2;\lambda)}{{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)}\\
&+ E(N^w) - (E(N^w))^2,
\end{aligned}
$$

where ${}_pF_q(\cdot;\cdot;\cdot)$ *is the generalised hypergeometric function (Definition 10.17.)*

Restrictions:

- Domain: $n \in \mathbb{N}_1$.
- Parameters: $\lambda > 0,\ a \in \mathbb{N}_1$.

Note: Since the generalised hypergeometric function's subscripts contain a's the number of 1's and 2's will vary depending on the value of a.

Proof. From the definition of the normalising constant, it follows that

$$\begin{aligned} E\left(w\left(N\right)\right) &= \textstyle\sum_{k=0}^{\infty} w\left(k\right) f\left(k\right) \\ &= \textstyle\sum_{k=0}^{\infty} k^{-a} \frac{e^{-\lambda}\lambda^k}{k!}. \end{aligned}$$

Since $0^a = 0$ for all $a > 0$. This fraction is 0 when $k = 0$. Thus it follows that

$$\begin{aligned} E\left(w\left(N\right)\right) &= e^{-\lambda} \textstyle\sum_{k=1}^{\infty} \frac{\lambda^k}{k^a k!} \\ &= \lambda e^{-\lambda} \textstyle\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{k^a k!}. \end{aligned}$$

After reparameterising, it follows that

$$\begin{aligned} E\left(w\left(N\right)\right) &= \lambda e^{-\lambda} \textstyle\sum_{m=0}^{\infty} \frac{\lambda^m}{(m+1)^a m!} \\ &= \lambda e^{-\lambda} \textstyle\sum_{m=0}^{\infty} \frac{\Gamma(m+1)^a}{\Gamma(m+2)^a} \frac{\lambda^m}{m!} \\ &= \lambda e^{-\lambda} \textstyle\sum_{m=0}^{\infty} \frac{\Gamma(m+1)^a \Gamma(2)^a}{\Gamma(1)^a \Gamma(m+2)^a} \frac{\lambda^m}{m!} \\ &= \lambda e^{-\lambda}{}_{a+1}F_{a+1}\left(1, ..., 1; 2, ..., 2; \lambda\right). \end{aligned}$$

This proof, and many of the proofs in this thesis, rely on repeated use of the following result:

$$\begin{aligned} \Gamma\left(1\right) &= 1. \\ \Gamma\left(2\right) &= 1. \\ \tfrac{\Gamma(m+2)}{\Gamma(m+1)} &= m+1. \end{aligned}$$

They are stated here to facilitate understanding of the derivations, and they will not explicitly be stated again.
From the definition of the weighted Poisson probability mass function, it follows that

$$\begin{aligned} f_w\left(n\right) &= \frac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\ &= \frac{n^{-a}\frac{e^{-\lambda}\lambda^n}{n!}}{\lambda e^{-\lambda}{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)} \\ &= \frac{\lambda^{n-1}}{n^a n!{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)}. \end{aligned}$$

From the definition of the probability generating function, it follows that

$$\begin{aligned} g(z) &= \sum_{k=0}^{\infty} f_w(k) z^k \\ &= \sum_{k=0}^{\infty} \frac{\lambda^{k-1}}{k^a k!\,{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)} z^k. \end{aligned}$$

Since $0^a = 0$ for all $a > 0$. This fraction is 0 when $k = 0$. Thus it follows that

$$\begin{aligned} g(z) &= \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{k^a k!\,{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)} z^k \\ &= \frac{z}{{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{k^a k!} z^{k-1} \\ &= \frac{z}{{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)} \sum_{k=1}^{\infty} \frac{(\lambda z)^{k-1}}{k^a k!}. \end{aligned}$$

After reparameterising, it follows that

$$\begin{aligned} g(z) &= \frac{z}{{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)} \sum_{m=0}^{\infty} \frac{(\lambda z)^m}{(m+1)^a m!} \\ &= \frac{z}{{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)} \sum_{m=0}^{\infty} \frac{\Gamma(m+1)^a}{\Gamma(m+2)^a} \frac{(\lambda z)^m}{m!} \\ &= \frac{z}{{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)} \sum_{m=0}^{\infty} \frac{\Gamma(m+1)^a \Gamma(2)^a}{\Gamma(1)^a \Gamma(m+2)^a} \frac{(\lambda z)^m}{m!} \\ &= \frac{z\,{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda z)}{{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)}. \end{aligned}$$

From the definition of the expected value, it follows that

$$E(N^w) = \lim_{z \to -1} \frac{\partial}{\partial z} g(z).$$

Thus it follows that

$$\frac{\partial}{\partial z} g(z) = \frac{\partial}{\partial z} \left(\frac{z\,{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda z)}{{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)} \right).$$

By using Theorem 10.7, it follows that

$$\begin{aligned} \frac{\partial}{\partial z} g(z) &= \frac{\partial}{\partial z} \left(\frac{z\,{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda z)}{{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)} \right) \\ &= \frac{{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda z) + {}_aF_a(1,...,1;2,...,2;\lambda z) - {}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda z)}{{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)} \\ &= \frac{{}_aF_a(1,...,1;2,...,2;\lambda z)}{{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)}. \end{aligned}$$

Consequently, it follows that

$$\begin{aligned} E(N^w) &= \lim_{z \to -1} \frac{{}_aF_a(1,...,1;2,...,2;\lambda z)}{{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)} \\ &= \frac{{}_aF_a(1,...,1;2,...,2;\lambda)}{{}_{a+1}F_{a+1}(1,...,1;2,...,2;\lambda)}. \end{aligned}$$

From the definition of the variance, it follows that

$$Var(N^w) = \lim_{z \to -1} \frac{\partial^2}{\partial z^2} g(z) + E(N^w) - (E(N^w))^2.$$

However,

$$\begin{aligned}
\frac{\partial^2}{\partial z^2} g(z) &= \frac{\partial}{\partial z}\left(\frac{{}_aF_a(1,\ldots,1;2,\ldots,2;\lambda z)}{{}_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)}\right) \\
&= \frac{\frac{1}{z}({}_{a-1}F_{a-1}(1,\ldots,1;2,\ldots,2;\lambda z) - {}_aF_a(1,\ldots,1;2,\ldots,2;\lambda z))}{{}_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)} \\
&= \frac{{}_{a-1}F_{a-1}(1,\ldots,1;2,\ldots,2;\lambda z) - {}_aF_a(1,\ldots,1;2,\ldots,2;\lambda z)}{z_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)}.
\end{aligned}$$

Consequently, it follows that

$$\begin{aligned}
Var(N^w) &= \lim_{z\to -1} \frac{\partial^2}{\partial z^2} g(z) + E(N^w) - (E(N^w))^2 \\
&= \lim_{z\to -1} \frac{{}_{a-1}F_{a-1}(1,\ldots,1;2,\ldots,2;\lambda z) - {}_aF_a(1,\ldots,1;2,\ldots,2;\lambda z)}{z_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)} \\
&+ E(N^w) - (E(N^w))^2 \\
&= \frac{{}_{a-1}F_{a-1}(1,\ldots,1;2,\ldots,2;\lambda) - {}_aF_a(1,\ldots,1;2,\ldots,2;\lambda)}{z_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)} \\
&+ E(N^w) - (E(N^w))^2.
\end{aligned}$$

□

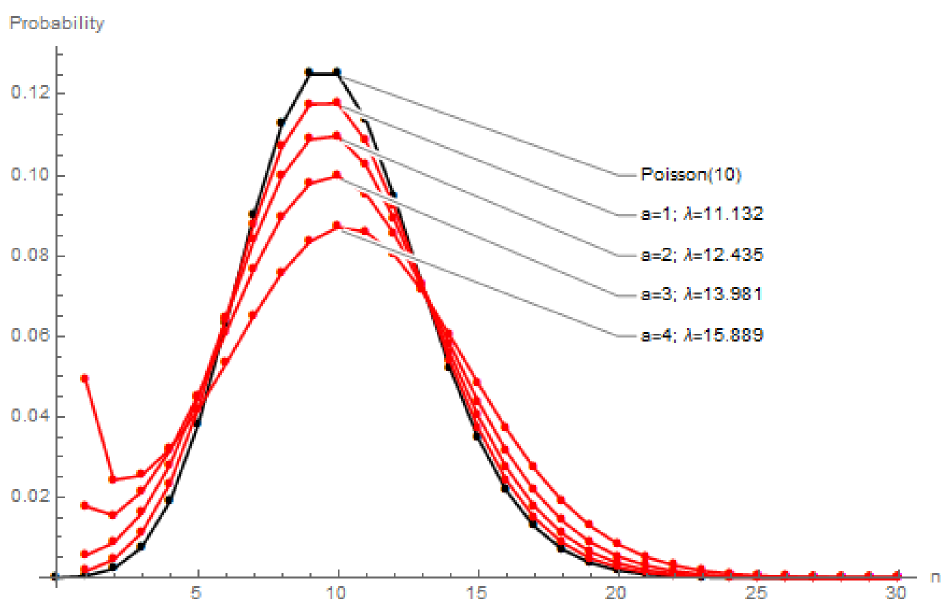


Figure 4.2: Probability mass functions - $w(n;\phi) = n^{-a}$

The variances associated with the weighted Poisson probability mass functions in the above plots are 11.3168, 13.2258, 16.2999 and 21.63 for a equal to 1,2,3 and 4 respectively. From the plots, it may appear as if this weight function always results in an overdispersed distribution. This is not the case. From a numerical investigation, it can be seen that for each a value, there is a corresponding λ for which the distribution is equidispersed. As the a value increases, the corresponding λ value required for equidispersion also increases. The expected values of the respective equidispersed distributions decrease as a function of a. These findings are briefly summarised in the table below.

Weight function	Poisson parameter	$E\left(N^w\right)$
$w(n) = n^{-1}$	$\lambda = 3.75$	2.84
$w(n) = n^{-2}$	$\lambda = 4.189$	2.185
$w(n) = n^{-3}$	$\lambda = 5.292$	1.855
$w(n) = n^{-4}$	$\lambda = 6.967$	1.640
$w(n) = n^{-5}$	$\lambda = 9.253$	1.483
$w(n) = n^{-6}$	$\lambda = 12.173$	1.359
$w(n) = n^{-7}$	$\lambda = 15.689$	1.162

Table 4.1: Equidispersed parameterisations - $w(n;\phi) = n^{-a}$

From the above table, it is clear that this specific weighted Poisson distribution can only be underdispersed for relatively small expected values. The plots below demonstrate some of the varying shapes of the distribution when the expected value is fixed at 1.5.

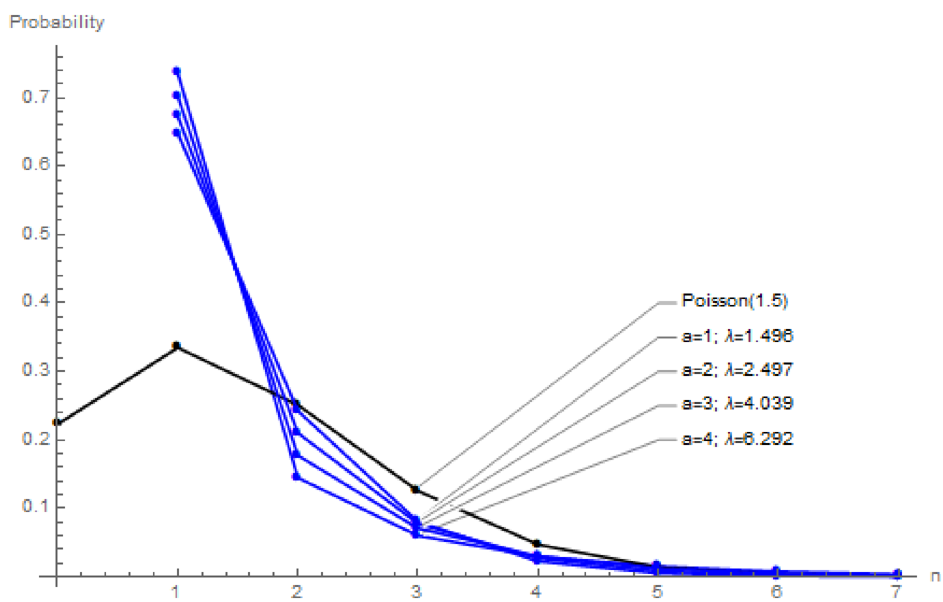


Figure 4.3: Probability mass functions - $w(n;\phi) = n^{-a}$

The variances associated with the weighted Poisson probability mass functions in the above plots are 0.642, 0.753, 0.913 and 1.1594 for a equal to 1,2,3 and 4 respectively.

4.1.3 $w(n;\phi) = n + \varepsilon$

Theorem 4.3. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = n + \varepsilon$ *then*

$$
\begin{aligned}
E(w(N;\phi)) &= \lambda + \varepsilon. \\
f_w(n) &= \tfrac{(n+\varepsilon)e^{-\lambda}\lambda^n}{(\lambda+\varepsilon)n!}. \\
g(z) &= \tfrac{e^{(z-1)\lambda}(\varepsilon+\lambda z)}{\lambda+\varepsilon}. \\
E(N^w) &= \tfrac{\lambda(1+\lambda+\varepsilon)}{\lambda+\varepsilon}. \\
Var(N^w) &= \tfrac{\lambda\left((\lambda+\varepsilon)^2+\varepsilon\right)}{(\lambda+\varepsilon)^2}.
\end{aligned}
$$

Restrictions:

- Domain: $n \in \mathbb{N}_0$.
- Parameters: $\lambda, \varepsilon > 0$.

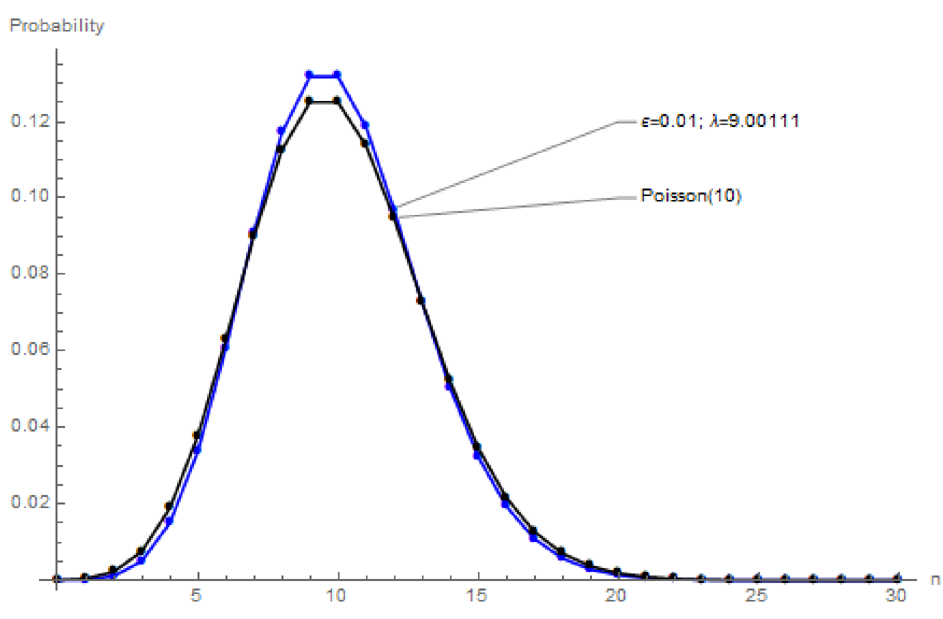


Figure 4.4: Probability mass functions - $w(n;\phi) = n + \varepsilon$

The variance associated with the weighted Poisson probability mass function in the above plot is 9.0022.

When $w(n;\phi) = n$ in Theorem 4.1, the resulting weighted Poisson distribution cannot accommodate zero counts. By adding a constant shift parameter to the weight function, this restriction is overcome while maintaining a similar Fisher index.

4.1.4 $w(n;\phi) = an^3 + bn^2 + cn$

Theorem 4.4. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = an^3 + bn^2 + cn$ *then*

$$\begin{aligned}
E(w(N;\phi)) &= \lambda(a+b+c) + \lambda^2(3a+b+\lambda a). \\
f_w(n) &= \frac{(an^3+bn^2+cn)}{(\lambda(3a+b+\lambda a)+(a+b+c))}\frac{e^{-\lambda}\lambda^n}{n!}. \\
g(z) &= \frac{e^{\lambda(z-1)}z\left(a\left((\lambda z)^2+3\lambda z+1\right)+b(\lambda z+1)+c\right)}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c}. \\
E(N^w) &= \frac{a\left(\lambda^3+6\lambda^2+7\lambda+1\right)+b\left(\lambda^2+3\lambda+1\right)+c(\lambda+1)}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c}. \\
Var(N^w) &= \frac{a\left(\lambda^4+9\lambda^3+19\lambda^2+8\lambda\right)+b\left(\lambda^3+5\lambda^2+4\lambda\right)+c\left(\lambda^2+2\lambda\right)}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c} + E(N^w) - (E(N^w))^2.
\end{aligned}$$

Restrictions:

- Domain: $n \in \mathbb{N}_1$.
- Parameters: $a, b, c \geq 0, \lambda > 0$.

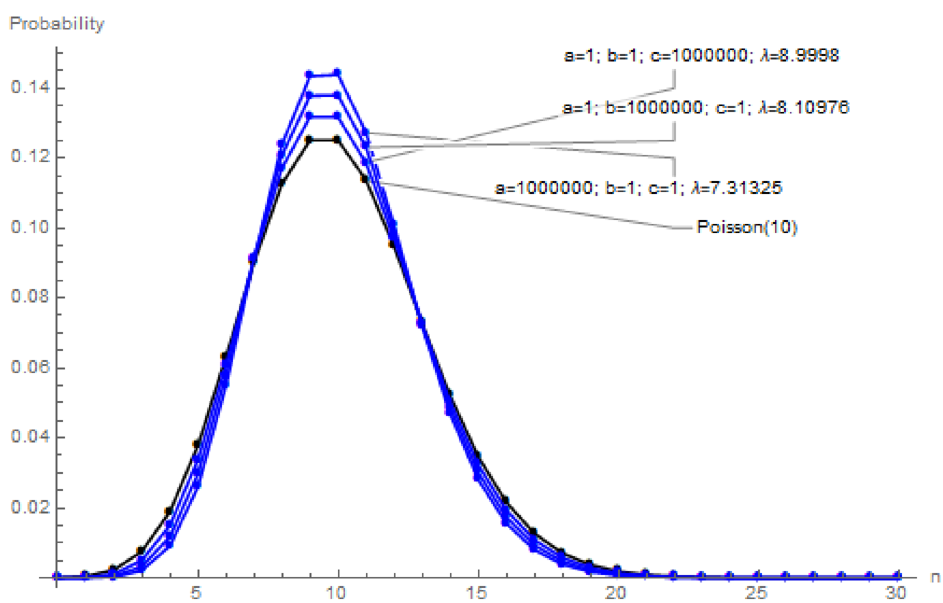


Figure 4.5: Probability mass functions - $w(n) = an^3 + bn^2 + cn$

The variances associated with the weighted Poisson probability mass functions in the above plots are 7.5546, 8.0208 and 9.0003 as a, b and c are increased respectively. This figure demonstrates how slight the change in the probability mass function is for large changes in the parameters of the weight function. The three variables, a, b and c, were each changed from their baseline value of one to one million to demonstrate that even changing these variables by a considerable amount results in very small changes in the variance of the distribution (while keeping the expected value fixed). It was also observed that while a, b and c could vary substantially with minimal effect on the distribution, the expected value and variance are much more sensitive to the value of λ.

4.1.5 $w(n;\phi) = (n+a)(n-b)^2$

As was seen above in Theorem 4.4, the weighted Poisson distribution with weight function $w(n) = an^3 + bn^2 + cn$; $a, b, c \geq 0$ varies minimally as the polynomial parameters change. There is one other potential problem with that specific weight function that must be highlighted: It is well known that a third-degree polynomial (as seen above) has at most two turning points. As a result, it is possible to construct a cubic weight function with turning points that assumes negative values on at least some of the negative integers but which is *a non-negative function on the set of non-negative integers*, and thus a valid weight function. Reparameterising the cubic weight function as $w(n;\phi) = (n+a)(n-b)^2, a, b \geq 0$ gives one such situation which is a special case of Theorem 4.4. In this case, there are zeros at $-a$ and b. Note, however, that the zero at b is also the turning point of the weight function, and as a result, the weight function is never smaller than zero on the non-negative integers. See Figure 4.6 below for four examples of possible parameterisations: $a = 1, b = 3$, $a = 2, b = 5$, $a = 3, b = 8$ and $a = 1000, b = 7$.

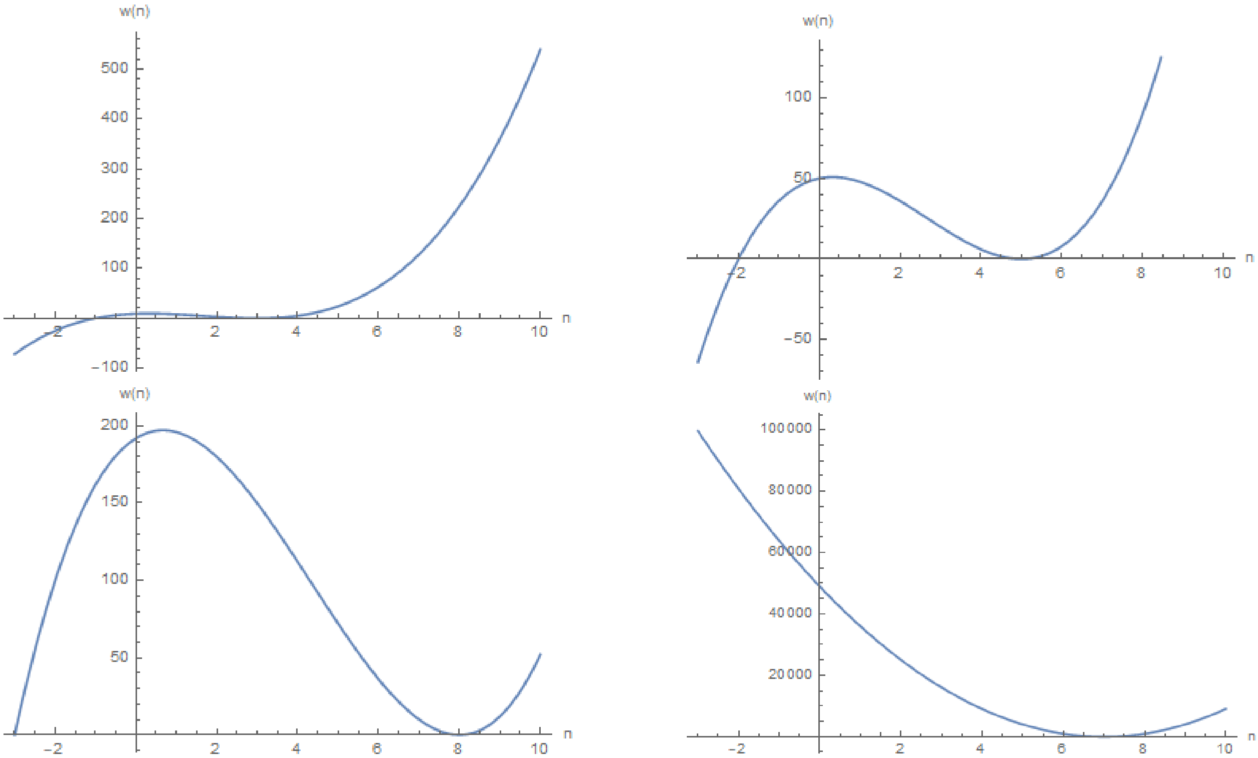


Figure 4.6: $w(n;\phi) = (n+a)(n-b)^2$

Theorem 4.5. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = (n+a)(n-b)^2$ *then*

$$E(w(N;\phi)) = \lambda(b^2 - 2ab - 2b + a + 1) + \lambda^2(\lambda + a - 2b + 3) + ab^2.$$

$$f_w(n) = \frac{(n+a)(n-b)^2}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2}\frac{e^{-\lambda}\lambda^n}{n!}.$$

$$g(z) = \frac{e^{\lambda(z-1)}\left((\lambda z)(b^2-2ab-2b+a+1)+(\lambda z)^2(\lambda z+a-2b+3)+ab^2\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2}.$$

$$E(N^w) = \frac{\lambda\left(1+7\lambda+6\lambda^2+\lambda^3+b^2(1+\lambda)-2b(1+3\lambda+\lambda^2)+a(1+b^2+3\lambda+\lambda^2-2b(1+\lambda))\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2}.$$

$$Var(N^w) = \frac{\lambda^2\left(8+19\lambda+9\lambda^2+\lambda^3+b^2(2+\lambda)-2b(4+5\lambda+\lambda^2)+a(4+b^2+5\lambda+\lambda^2-2b(2+\lambda))\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} + E(N^w) - (E(N^w))^2.$$

Restrictions:

- Domain: $n \in \mathbb{N}_0$.
- Parameters: $a, b \geq 0, \lambda > 0$.

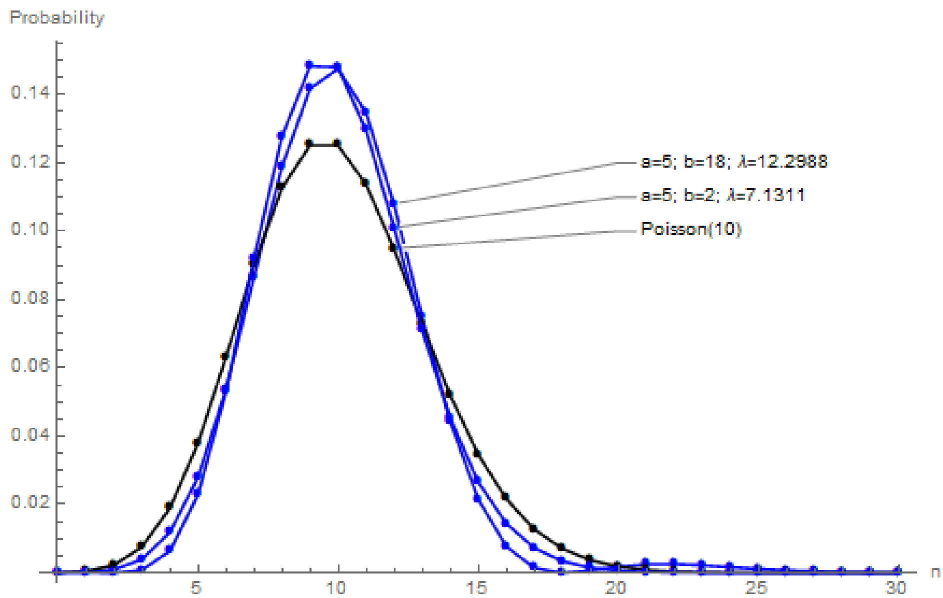


Figure 4.7: Probability mass function - $w(n;\phi) = (n+a)(n-b)^2$; $|b - E(N^w)|$ large

The variances associated with the weighted Poisson probability mass functions in the above plot are 7.0811 and 8.8749 for b values of 2 and 18 respectively.

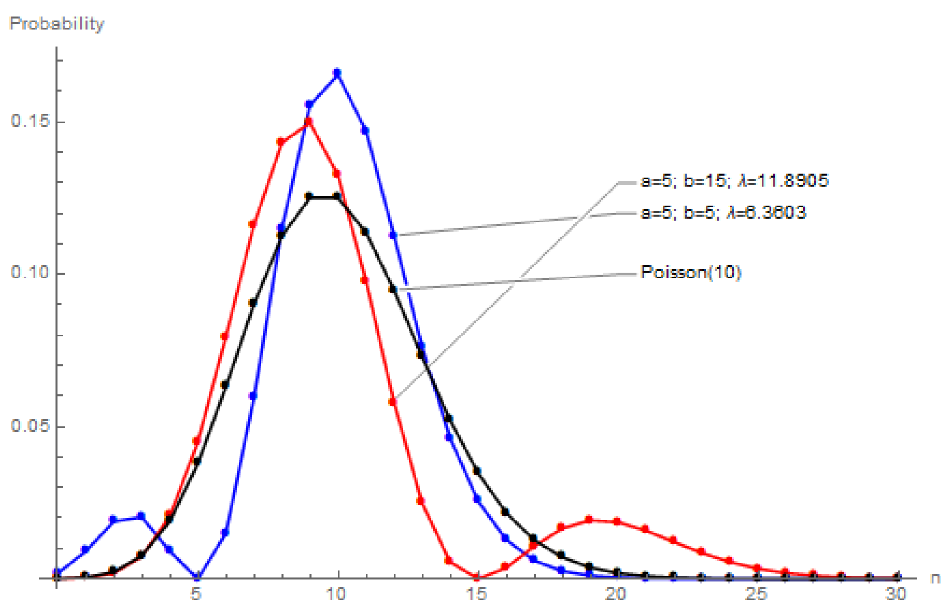


Figure 4.8: Probability mass function - $w(n;\phi) = (n+a)(n-b)^2$; $|b - E(N^w)|$ medium

The variances associated with the weighted Poisson probability mass functions in the above plots are 8.5704 and 19.2228 for b equal to 5 and 15 respectively.

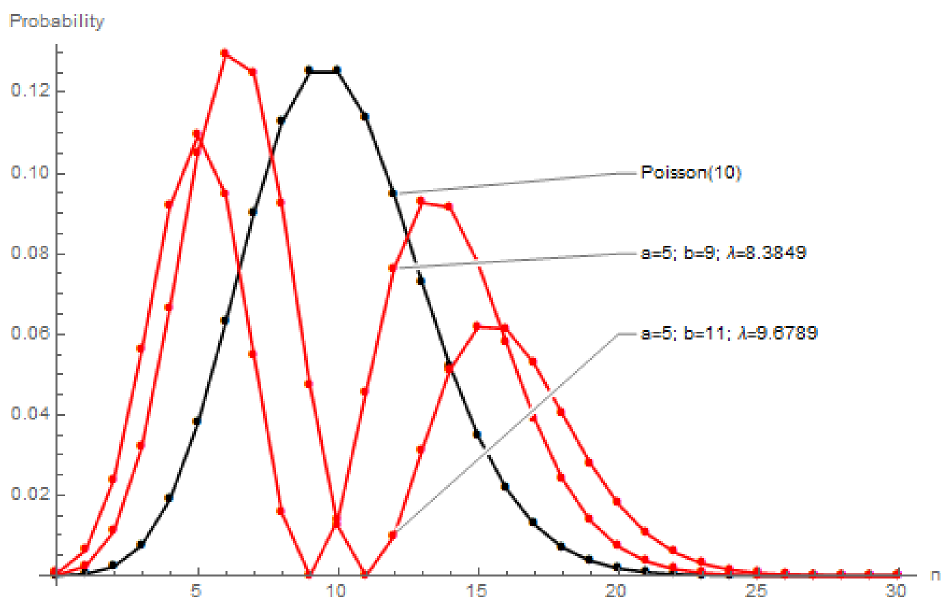


Figure 4.9: Probability mass function - $w(n;\phi) = (n+a)(n-b)^2$; $|b - E(N^w)|$ small

The variances associated with the weighted Poisson probability mass functions in the above plots are 25.8628 and 28.4565 for b equal to 9 and 11 respectively.

From the figures above, it is apparent that using this form of cubic weight function can result in a wide range of shapes for the probability mass function. Some general statements can be made about the parameters of the weight function: The parameter a has a less pronounced effect on the shape of the probability mass function than is the case for b. The parameter b corresponds to the turning point of the probability mass function where there is a zero probability of observing data points, and consequently, the parameter b can be interpreted as a clear divide between mixtures of data. The relationship between b and $E(N^w)$ not only determines the overall shape of the distribution but also whether the resulting distribution will be under or overdispersed. If $|b - E(N^w)|$ is "large", in other words, if the value of b (the turning point) is in the tails of the distribution, it appears as if the probability mass function does not contain a second local mode. Strictly speaking, this is not true; the second local mode is merely being masked by already small probabilities in the tails of the distribution. If $|b - E(N^w)|$ is "small" the distribution will be overdispersed. What constitutes "large" and "small" is not investigated, but the effect of these quantities has been demonstrated graphically. In the above graphs it was assumed that b is an integer. This need not be the case.

4.1.6 $w(n;\phi) = (n+a)\left(n^2 - bn + c\right)$

It is possible to further reparameterise the weight function given in Theorem 4.5. In Theorem 4.5 it is implicitly assumed that the factor $(n-b)^2$ can be factorised from $n^2 -$

$2bn + b^2$. If, however, it cannot, a more general solution is obtained where $w(n;\phi) = (n+a)(n^2 - bn + c)$. This new weight function comes with benefits as well as costs relative to $w(n;\phi) = (n+a)(n-b)^2$. The fact that the weight function in Theorem 4.5 has a turning point and obtains a zero value when $n = b$ potentially allows for a novel modelling opportunity where bi-modal data is completely separable into two distinct groups. However, this narrow application is also its greatest weakness since such data would likely not be observed frequently. By having a factor which does not attain the value of zero for any n, there would be no non-negative integer which will be observed with probability zero, and consequently, bi-modal data could potentially be modelled more realistically. The problem with $w(n;\phi) = (n+a)(n^2 - bn + c)$, however, is that it introduces additional restrictions on the parameters contained in the weight fynction. Due to the fact that $n^2 - bn + c$ must have no real roots, an added restriction is that $b^2 - 4c < 0$. Additional parameter restrictions are also introduced by the expression for probability mass function.

Below in Figure 4.10 some graphs are provided demonstrating the wide range of shapes that this weight function can assume. The parameterisations are $a = 1, b = 3, c = 5$; $a = 1, b = 4, c = 5$; $a = 1, b = 6, c = 15$ and $a = 2, b = 4, c = 5$.

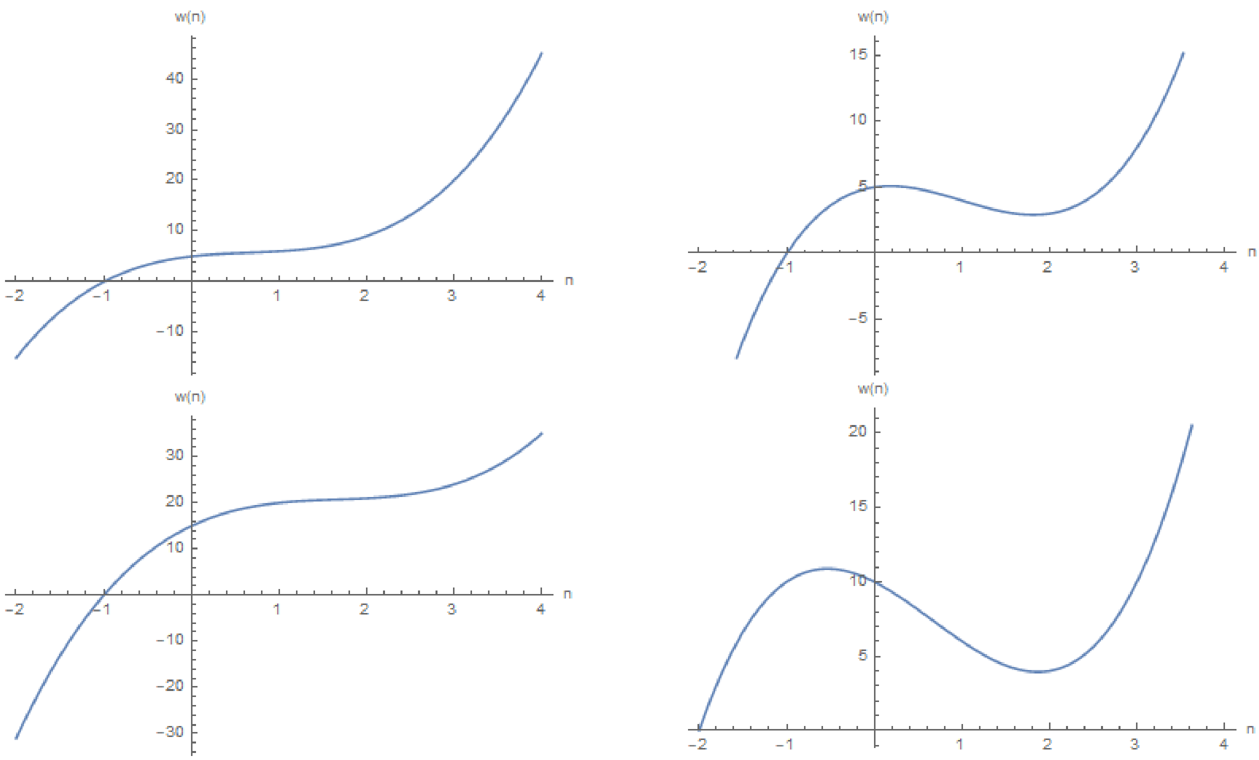


Figure 4.10: $w(n;\phi) = (n+a)(n^2 - bn + c)$

Theorem 4.6. *If the weight function used in the weighted Poisson probability mass function* *is chosen as* $w(n;\phi) = (n+a)(n^2 - bn + c)$ *then*

$$E(w(N;\phi)) = \lambda^3 + \lambda^2(a-b+3) + \lambda(a+c-b-ab+1) + ac.$$

$$f_w(n) = \frac{(n+a)(n^2-bn+c)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac}\frac{e^{-\lambda}\lambda^n}{n!}.$$

$$g(z) = \frac{e^{\lambda(z-1)}(\lambda^3z^3+\lambda^2z^2(a-3+b)+\lambda z(a+c-b-ab+1)+ac)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac}.$$

$$E(N^w) = \frac{\lambda(1+c+7\lambda+c\lambda+6\lambda^2+\lambda^3-b(1+3\lambda+\lambda^2)+a(1+c+3\lambda+\lambda^2-b(1+\lambda)))}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac}.$$

$$\begin{aligned} Var(N^w) &= \frac{\lambda^2(8+2c+19\lambda+c\lambda+9\lambda^2+\lambda^3-b(4+5\lambda+\lambda^2)+a(4+c+5\lambda+\lambda^2-b(2+\lambda)))}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac} \\ &+ E(N^w) - (E(N^w))^2. \end{aligned}$$

Restrictions:

- Domain: $n \in \mathbb{N}_0$.
- Parameters: $a, b, c \geq 0, \lambda > 0, \lambda^3 + \lambda^2(a-b+3) + \lambda(a+c-b-ab+1) + ac > 0$.

Proof. From the definition of the normalising constant, it follows that

$$\begin{aligned} E(w(N;\phi)) &= \sum_{k=0}^{\infty} w(k) f(k) \\ &= \sum_{k=0}^{\infty} (k+a)(k^2-bk+c)\frac{e^{-\lambda}\lambda^k}{k!} \\ &= \sum_{k=0}^{\infty} (k^3+k^2(a-b)+k(c-ab)+ac)\frac{e^{-\lambda}\lambda^k}{k!} \\ &= \sum_{k=0}^{\infty} k^3\frac{e^{-\lambda}\lambda^k}{k!} + (a-b)\sum_{k=0}^{\infty} k^2\frac{e^{-\lambda}\lambda^k}{k!} \\ &+ (c-ab)\sum_{k=0}^{\infty} k\frac{e^{-\lambda}\lambda^k}{k!} + ac\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!}. \end{aligned}$$

Since the first three sums in the above equations are the third, second and first moments respectively of a Poisson distribution with parameter λ, and the fourth sum is the Poisson probability mass function, it follows that

$$\begin{aligned} E(w(N;\phi)) &= (\lambda^3+3\lambda^2+\lambda) + (a-b)(\lambda^2+\lambda) + (c-ab)\lambda + ac \\ &= \lambda^3 + \lambda^2(a-b+3) + \lambda(a+c-b-ab+1) + ac. \end{aligned}$$

From the definition of the weighted Poisson probability mass function, it follows that

$$\begin{aligned} f_w(n) &= \frac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\ &= \frac{(n+a)(n^2-bn+c)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac}\frac{e^{-\lambda}\lambda^n}{n!}. \end{aligned}$$

From the definition of the probability generating function, it follows that

$$
\begin{aligned}
g(z) &= \sum_{k=0}^{\infty} f_w(k)\, z^k \\
&= \sum_{k=0}^{\infty} \frac{(k+a)\left(k^2-bk+c\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac} \frac{e^{-\lambda}\lambda^k}{k!} z^k \\
&= \frac{\sum_{k=0}^{\infty}\left(k^3+k^2(a-b)+k(c-ab)+ac\right)\frac{e^{-\lambda}\lambda^k}{k!} z^k}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac} \\
&= \frac{e^{-\lambda}\sum_{k=0}^{\infty}\left(k^3\frac{\lambda^k}{k!}z^k+k^2(a-b)\frac{\lambda^k}{k!}z^k+k(c-ab)\frac{\lambda^k}{k!}z^k+ac\frac{\lambda^k}{k!}z^k\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac} \\
&= \frac{e^{-\lambda}\left(\sum_{k=0}^{\infty}k^3\frac{\lambda^k}{k!}z^k+\sum_{k=0}^{\infty}k^2(a-b)\frac{\lambda^k}{k!}z^k+\sum_{k=0}^{\infty}k(c-ab)\frac{\lambda^k}{k!}z^k+\sum_{k=0}^{\infty}ac\frac{\lambda^k}{k!}z^k\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac} \\
&= \frac{\frac{e^{-\lambda}}{e^{-\lambda z}}\left(\sum_{k=0}^{\infty}k^3a_{\lambda,k,z}+(a-b)\sum_{k=0}^{\infty}k^2a_{\lambda,k,z}+(c-ab)\sum_{k=0}^{\infty}ka_{\lambda,k,z}+ac\sum_{k=0}^{\infty}a_{\lambda,k,z}\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac}
\end{aligned}
$$

where $a_{\lambda,k,z} = \frac{(\lambda z)^k e^{-\lambda z}}{k!}$.

Since the first three sums in the above equations are the third, second and first moments respectively of a Poisson distribution with parameter λz, and the fourth sum is equals 1, it follows that

$$
\begin{aligned}
g(z) &= \frac{\frac{e^{-\lambda}}{e^{-\lambda z}}\left((\lambda z)^3+3(\lambda z)^2+\lambda z+(a-b)\left((\lambda z)^2+\lambda\right)+(c-2ab)\lambda z+ac\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac} \\
&= \frac{e^{\lambda(z-1)}\left(\lambda z(c-ab-b+a+1)+(\lambda z)^2(\lambda z+a-b+3)+ac\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac}.
\end{aligned}
$$

From the definition of the expected value, it follows that

$$
E\left(N^w\right) = \lim_{z\to^-1}\frac{\partial}{\partial z}g(z).
$$

However,

$$
\begin{aligned}
\frac{\partial}{\partial z}g(z) &= \frac{\partial}{\partial z}\left(\frac{e^{\lambda(z-1)}\left(\lambda z(c-ab-b+a+1)+(\lambda z)^2(\lambda z+a-b+3)+ac\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac}\right) \\
&= \left(\frac{\partial}{\partial z}e^{\lambda(z-1)}\right)\left(\frac{(\lambda z)(c-ab-b+a+1)+(\lambda z)^2(\lambda z+a-b+3)+ac}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac}\right) \\
&+ \left(\frac{e^{\lambda(z-1)}}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac}\right) \\
&\times \left(\frac{\partial}{\partial z}(\lambda z)(c-ab-b+a+1)+(\lambda z)^2(\lambda z+a-b+3)+ac\right) \\
&= \frac{e^{\lambda(z-1)}\lambda\left(\lambda z(c-ab-b+a+1)+(\lambda z)^2(\lambda z+a-b+3)+ac\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac} \\
&+ \left(\frac{e^{\lambda(z-1)}\left(\lambda(c-ab-b+a+1)+\lambda^2 z(3\lambda z+2a-2b+6)\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac}\right) \\
&= \frac{e^{\lambda(z-1)}\lambda\left(1+c+7z\lambda+cz\lambda+6z^2\lambda^2+z^3\lambda^3-b\left(1+3z\lambda+z^2\lambda^2\right)+a\left(1+c+3z\lambda+z^2\lambda^2-b(1+z\lambda)\right)\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac}.
\end{aligned}
$$

Consequently, it follows that

$$
\begin{aligned}
E\left(N^{w}\right) &= \lim_{z\to -1}\frac{e^{\lambda(z-1)}\lambda\left(1+c+7z\lambda+cz\lambda+6z^2\lambda^2+z^3\lambda^3-b\left(1+3z\lambda+z^2\lambda^2\right)+a\left(1+c+3z\lambda+z^2\lambda^2-b(1+z\lambda)\right)\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac} \\
&= \frac{\lambda\left(1+c+7\lambda+c\lambda+6\lambda^2+\lambda^3-b\left(1+3\lambda+\lambda^2\right)+a\left(1+c+3\lambda+\lambda^2-b(1+\lambda)\right)\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac}.
\end{aligned}
$$

From the definition of the variance, it follows that

$$
Var\left(N^{w}\right) = \lim_{z\to -1}\frac{\partial^2}{\partial z^2}g\left(z\right)+E\left(N^{w}\right)-\left(E\left(N^{w}\right)\right)^2.
$$

However,

$$
\begin{aligned}
\frac{\partial^2}{\partial z^2}g\left(z\right) &= \frac{\partial}{\partial z}\left(\frac{e^{\lambda(z-1)}\lambda\left(1+c+7z\lambda+cz\lambda+6z^2\lambda^2+z^3\lambda^3-b\left(1+3z\lambda+z^2\lambda^2\right)+a\left(1+c+3z\lambda+z^2\lambda^2-b(1+z\lambda)\right)\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac}\right) \\
&= \left(\frac{\partial}{\partial z}e^{\lambda(z-1)}\right)\frac{\lambda\left(1+c+7z\lambda+cz\lambda+6z^2\lambda^2+z^3\lambda^3-b\left(1+3z\lambda+z^2\lambda^2\right)+a\left(1+c+3z\lambda+z^2\lambda^2-b(1+z\lambda)\right)\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac} \\
&+ \left(\frac{e^{\lambda(z-1)}}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac}\right) \\
&\times \frac{\partial}{\partial z}\left(\lambda\left(1+c+7z\lambda+cz\lambda+6z^2\lambda^2+z^3\lambda^3-b\left(1+3z\lambda+z^2\lambda^2\right)\right)\right. \\
&\quad \left.+\lambda a\left(1+c+3z\lambda+z^2\lambda^2-b\left(1+z\lambda\right)\right)\right) \\
&= \frac{e^{\lambda(z-1)}\lambda^2\left(z\lambda(1+a-b-ab+c)+z^2\lambda^2(3+a-b+z\lambda)+ac\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac} \\
&+ \frac{2e^{\lambda(z-1)}\lambda^2\left(1+a-b-ab+c+6z\lambda+2az\lambda-2bz\lambda+3z^2\lambda^2\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac} \\
&+ \frac{e^{\lambda(z-1)}\lambda^2(6+2a-2b+6z\lambda)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac} \\
&= \frac{e^{\lambda(z-1)}\lambda^2\left(8+2c+19z\lambda+cz\lambda+9z^2\lambda^2+z^3\lambda^3-b\left(4+5z\lambda+z^2\lambda^2\right)+a\left(4+c+5z\lambda+z^2\lambda^2-b(2+z\lambda)\right)\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac}.
\end{aligned}
$$

Consequently, it follows that

$$
\begin{aligned}
Var\left(N^{w}\right) &= \lim_{z\to -1}\frac{\partial^2}{\partial z^2}g\left(z\right)+E\left(N^{w}\right)-\left(E\left(N^{w}\right)\right)^2 \\
&= \lim_{z\to -1}\frac{e^{\lambda(z-1)}\lambda^2\left(8+2c+19z\lambda+cz\lambda+9z^2\lambda^2+z^3\lambda^3-b\left(4+5z\lambda+z^2\lambda^2\right)+a\left(4+c+5z\lambda+z^2\lambda^2-b(2+z\lambda)\right)\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac} \\
&+ E\left(N^{w}\right)-\left(E\left(N^{w}\right)\right)^2 \\
&= \frac{\lambda^2\left(8+2c+19\lambda+c\lambda+9\lambda^2+\lambda^3-b\left(4+5\lambda+\lambda^2\right)+a\left(4+c+5\lambda+\lambda^2-b(2+\lambda)\right)\right)}{\lambda^3+\lambda^2(a-b+3)+\lambda(a+c-b-ab+1)+ac}+E\left(N^{w}\right)-\left(E\left(N^{w}\right)\right)^2.
\end{aligned}
$$

□

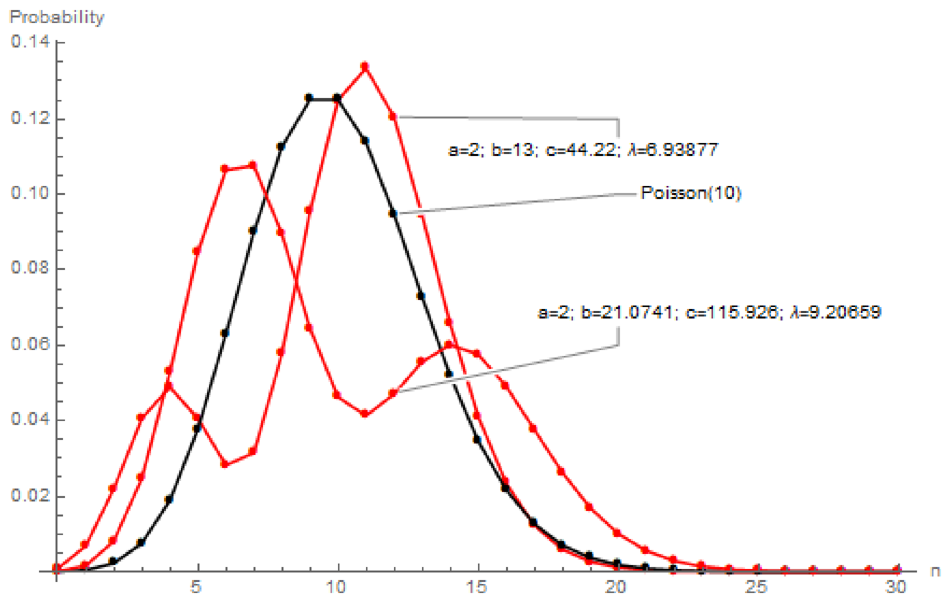


Figure 4.11: Probability mass function - $w(n;\phi) = (n+a)(n^2 - bn + c)$

The variances associated with the weighted Poisson probability mass functions in the above plots are 13.7047 and 21.4081 in the cases where b equals 13 and 21.0741 respectively. From the graph it is apparent that if the added parameter restrictions do not limit the shape of the probability mass function, this weighted Poisson distribution is more flexible, and therefore it may be a more realistic model for observed bi-modal data.

4.1.7 $w(n;\phi) = \frac{a^*n+b^*}{c^*n+d^*} = a + \frac{b-ac}{n+c}$

Another potential weight function is the ratio between two first order polynomials $a^*n + b^*$ and $c^*n + d^*$. The number of parameters in this expression can be reduced as follows:

$$\frac{a^*n+b^*}{c^*n+d^*} = \frac{a^*n+b^*}{\frac{c^*}{a^*}a^*n+d^*} = \frac{a^*n+b^*}{\frac{c^*}{a^*}a^*n+d^*} = \frac{a^*n+b^*}{c^*\left(n+\frac{d^*}{c^*}\right)}$$
$$= \frac{\frac{a^*}{c^*}n+\frac{b^*}{c^*}}{\left(n+\frac{d^*}{c^*}\right)} = \frac{an+b}{n+c} = \frac{a(n+c)+b-ac}{n+c} = a + \frac{b-ac}{n+c}$$

If the weight function used in the weighted Poisson probability mass function is chosen as

$w(n) = a + \frac{b-ac}{n+c}$ then

$$\begin{aligned}
E(w(N;\phi)) &= e^{-\lambda}(-\lambda)^{-c}\left(ae^{\lambda}(-\lambda)^{c} + (b-ac)\gamma(c,-\lambda)\right). \\
f_w(n) &= \frac{(b+an)(-\lambda)^{c}\lambda^{n}}{(c+n)n!\left(ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)\right)}. \\
g(z) &= \frac{(-\lambda)^{c}\left(ae^{z\lambda}(-z\lambda)^{c}+(b-ac)\gamma(c,-\lambda z)\right)}{\left(ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)\right)(-z\lambda)^{c}}. \\
E(N^{w}) &= \frac{\left(e^{\lambda}(-\lambda)^{c}(b-ac+a\lambda)+c(ac-b)\gamma(c,-\lambda)\right)}{\left(ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)\right)}. \\
Var(N^{w}) &= \frac{(-\lambda)^{c}\left(e^{\lambda}(-\lambda)^{c}\left((1+c)(ac-b)+(b-ac)\lambda+a\lambda^{2}\right)+c(1+c)(b-ac)\gamma(c,-\lambda)\right)}{(-\lambda)^{c}\left(ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)\right)} \\
&+ E(N^{w}) - \left(E(N^{w})\right)^{2}.
\end{aligned}$$

Restrictions:

- Domain: $n \in \mathbb{N}_0$.
- Parameters: $a, b > 0, \lambda > 0, c \in \mathbb{N}_0$.

The restriction that $c \in \mathbb{N}_0$ comes from the $(-\lambda)^c$ term in the equations above. Since $\lambda > 0$, $-\lambda < 0$. Raising a negative value to some exponent may lead to multiple complex solutions if the exponent is not an integer. To avoid these potential complications it is assumed that $c \in \mathbb{N}_0$.

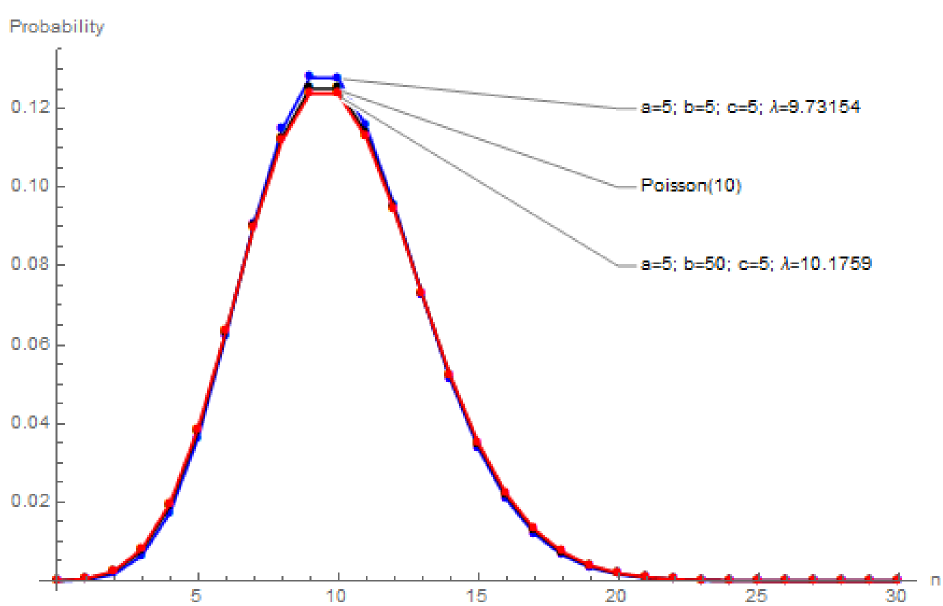


Figure 4.12: Probability mass function - $w(n;\phi) = a + \frac{b-ac}{n+c}$

The variances associated with the weighted Poisson probability mass functions in the above plots are 9.5753 and 10.2040 for b values equal to 5 and 50 respectively. Based on the plots for this distribution, it appears as if it can model relatively small deviations from the Poisson distribution.

4.2 Probability mass/density functions as weight functions

One class of potential weight function that has received no research attention to date are weight functions that are in and of themselves probability mass functions. In this section, some of these weights will be investigated.

4.2.1 $w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$

If $w(n;\phi)$ is assumed to be the probability mass function of the negative binomial distribution the following results are obtained.

Theorem 4.7. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$ *then*

$$
\begin{aligned}
E\left(w\left(N;\phi\right)\right) &= e^{-\lambda}\left(1-p\right)^r {}_1F_1\left(r;1;p\lambda\right). \\
f_w\left(n\right) &= \frac{\binom{n+r-1}{n}}{{}_1F_1(r;1;p\lambda)}\frac{(p\lambda)^n}{n!}. \\
g\left(z\right) &= \frac{{}_1F_1(r;1;pz\lambda)}{{}_1F_1(r;1;p\lambda)}. \\
E\left(N^w\right) &= \frac{\lambda p r {}_1F_1(r+1;2;p\lambda)}{{}_1F_1(r;1;p\lambda)}. \\
Var\left(N^w\right) &= \frac{\lambda^2 p^2 r(r+1) {}_1F_1(r+2;3;p\lambda)}{2 {}_1F_1(r;1;p\lambda)} + E\left(N^w\right) - \left(E\left(N^w\right)\right)^2.
\end{aligned}
$$

where ${}_1F_1(\cdot;\cdot;\cdot)$ *is the confluent hypergeometric function in Definition 10.15.*

Restrictions:

- Domain: $n \in \mathbb{N}_0$.
- Parameters: $r \in \mathbb{N}_1, 0 < p < 1, \lambda > 0$.

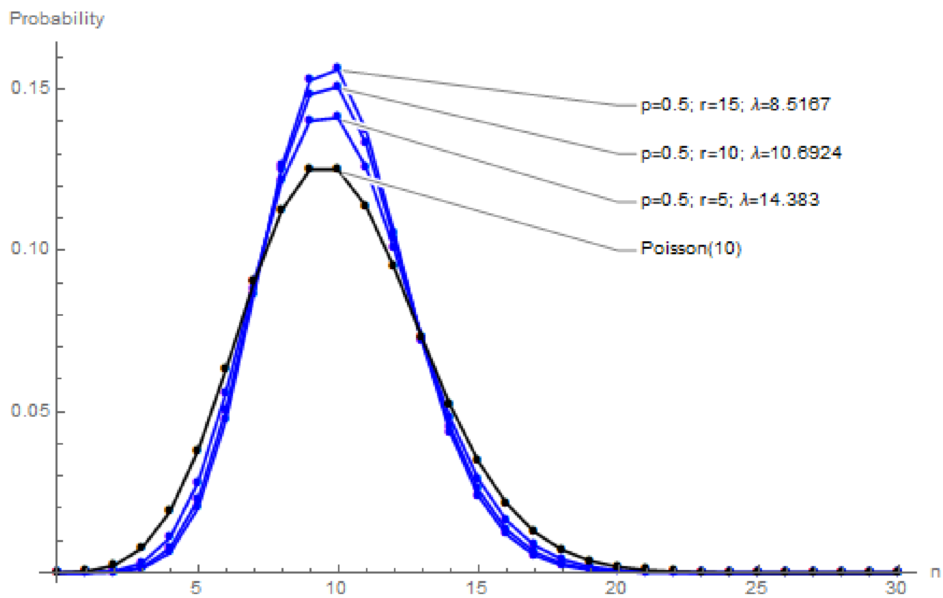


Figure 4.13: Probability mass function - $w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$, varying r

The variances associated with the weighted Poisson probability mass functions in the above plots are 7.8725, 6.9238 and 6.4648 for r equal to 5, 10 and 15 respectively.

and

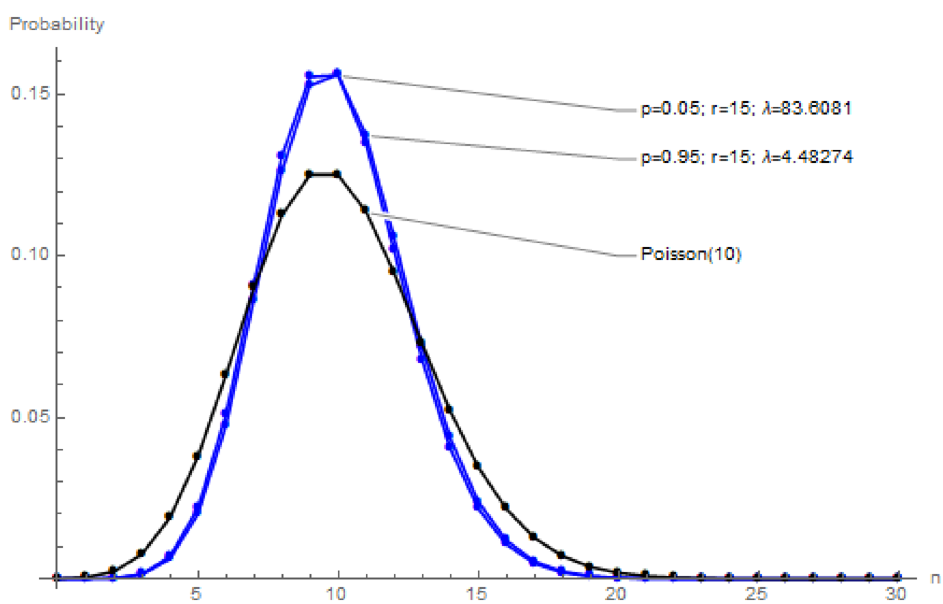


Figure 4.14: Probability mass function - $w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$, varying p

The variances associated with the weighted Poisson probability mass functions in the above plots are 6.3786 and 6.465 for p equal to 0.05 and 0.95 respectively. From the above plots it is apparent that the distribution's shape is more sensitive to changes in r, than in p. It may seem counter intuitive that the negative binomial distribution, which is itself overdispersed, results in an underdispersed weighted Poisson distribution when it is used as a weight function. However, once the logconcavity result from Kokonendi et al. [83] is taken into account, this conclusion becomes apparent.

4.2.2 $w(n;\phi) = \binom{m}{n} p^n (1-p)^{m-n}$

If $w(n;\phi)$ is assumed to be the probability mass function of the binomial distribution the following results are obtained.

Theorem 4.8. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \binom{m}{n} p^n (1-p)^{m-n}$ *then*

$$\begin{aligned}
E(w(N;\phi)) &= e^{-\lambda}(1-p)^m L_m\left(\tfrac{p\lambda}{p-1}\right). \\
f_w(n) &= \binom{m}{n} \frac{(p\lambda)^n}{(1-p)^n n! L_m\left(\frac{p\lambda}{p-1}\right)}. \\
g(z) &= \frac{L_m\left(\frac{p\lambda z}{p-1}\right)}{L_m\left(\frac{p\lambda}{p-1}\right)}. \\
E(N^w) &= \frac{p\lambda L^1_{m-1}\left(\frac{p\lambda}{p-1}\right)}{(1-p)L_m\left(\frac{p\lambda}{p-1}\right)}. \\
Var(N^w) &= +\frac{(p\lambda)^2 L^2_{m-2}\left(\frac{p\lambda}{p-1}\right)}{(p-1)^2 L_m\left(\frac{p\lambda}{p-1}\right)} + E(N^w) - (E(N^w))^2.
\end{aligned}$$

where $L_m(\cdot)$ is the Laguerre polynomial and $L^\alpha_m(\cdot)$ is the generalised Laguerre polynomial (Definition 10.19).

Restrictions:

- Domain: $n \in \{0,1,...,m\}$. This restriction follows from the fact that if $n > m$, $\binom{m}{n}$ is defined to be 0.
- Parameters: $0 < p < 1, \lambda > 0, L_m\left(\frac{p\lambda}{p-1}\right) \geq 0$. The restriction that $L_m\left(\frac{p\lambda}{p-1}\right) \geq 0$ is a consequence of requiring that the probability mass function be non-negative. Since no general expressions exist for when the Laguerre polynomial is positive, this general statement is given.

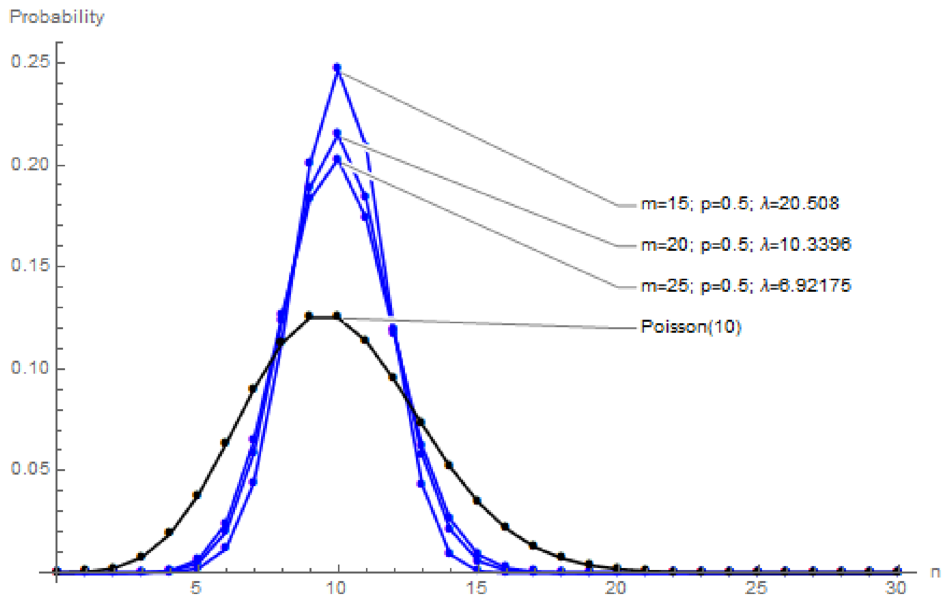


Figure 4.15: Probability mass function - $w(n;\phi) = \begin{pmatrix} m \\ n \end{pmatrix} p^n (1-p)^{m-n}$, varying m

The variances associated with the weighted Poisson probability mass functions in the above plots are 2.5398, 3.3963 and 3.8263 for m equal to 15, 20 and 25 respectively.

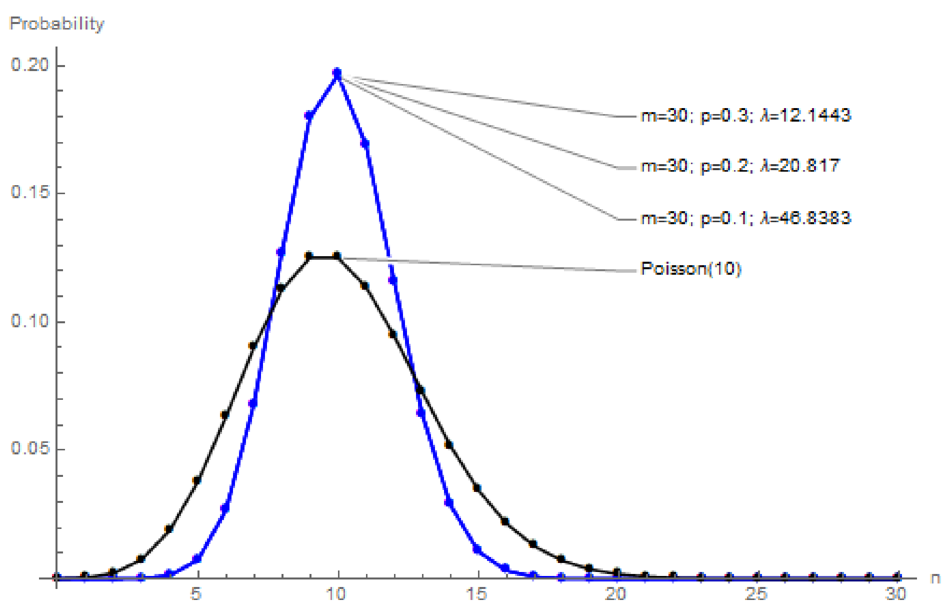


Figure 4.16: Probability mass function - $w(n;\phi) = \begin{pmatrix} m \\ n \end{pmatrix} p^n (1-p)^{m-n}$, varying p

The variance associated with the weighted Poisson probability mass functions in the above plot is 4.0851 for all values of p. This implies that varying the parameter p is completely ineffective as a means of changing the shape of the probability mass function of this specific weighted Poisson distribution (when the expected value is fixed). The reason for this can clearly be seen in the expressions of the equations as well as in the above plot: In all cases, when p decreases, λ is required to increase in order for the mean of the distribution to remain constant. The interaction between these parameters is of such a nature an increase in one cancels out the decrease in the other.

4.2.3 $w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$

If $w(n;\phi)$ is assumed to be the probability mass function of the Yule-Simon distribution the following results are obtained.

Theorem 4.9. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$ *then*

$$\begin{aligned}
E(w(N;\phi)) &= ae^{-\lambda}\lambda\tfrac{\Gamma(1+a)}{\Gamma(2+a)}\,{}_2F_2(1,1;2,2+a;\lambda). \\
f_w(n) &= \frac{\lambda^{n-1}\frac{\Gamma(n)\Gamma(1+a)}{\Gamma(n+a+1)}}{ae^{-\lambda}\lambda\frac{\Gamma(1+a)}{\Gamma(2+a)}\,{}_2F_2(1,1;2,2+a;\lambda)}. \\
g(z) &= z\frac{{}_2F_2(1,1;2,2+a;\lambda z)}{{}_2F_2(1,1;2,2+a;\lambda)}. \\
E(N^w) &= \frac{{}_2F_2(1,1;2,2+a;\lambda)+\frac{\lambda}{2(2+a)}\,{}_2F_2(2,2;3,3+a;\lambda)}{{}_2F_2(1,1;2,2+a;\lambda)}. \\
Var(N^w) &= \frac{\lambda^{-1-a}\left(e^{\lambda}(\lambda-1-a)\right)\Gamma(2+a)+(1+a)\left(\lambda^{1+a}-e^{\lambda}(\lambda-1-a)\Gamma(1+a,\lambda)\right)}{{}_2F_2(1,1;2,2+a;\lambda)} \\
&+ E(N^w)-\left(E(N^w)\right)^2.
\end{aligned}$$

Restrictions:

- Domain: $n \in \mathbb{N}_1$.
- Parameters: $a > 0, \lambda > 0$.

Proof. From the definition of the normalising constant, it follows that

$$\begin{aligned}
E\left(w\left(N;\phi\right)\right) &= \textstyle\sum_{k=0}^{\infty} w\left(k\right) f\left(k\right) \\
&= \textstyle\sum_{k=0}^{\infty} a \frac{\Gamma(k)\Gamma(a+1)}{\Gamma(k+a+1)} \frac{e^{-\lambda}\lambda^{k}}{k!} \\
&= ae^{-\lambda}\Gamma\left(a+1\right) \textstyle\sum_{k=0}^{\infty} \frac{\Gamma(k)}{\Gamma(k+a+1)} \frac{\lambda^{k}}{k!} \\
&= ae^{-\lambda} \frac{\Gamma(a+1)}{\Gamma(a+2)} \textstyle\sum_{k=0}^{\infty} \frac{\Gamma(k)\Gamma(a+2)}{\Gamma(k+a+1)} \frac{\lambda^{k}}{k!} \\
&= ae^{-\lambda}\lambda \frac{\Gamma(a+1)}{\Gamma(a+2)} \textstyle\sum_{k=0}^{\infty} \frac{\Gamma(k)\Gamma(a+2)}{\Gamma(k+a+1)} \frac{\lambda^{k-1}}{k!} \\
&= ae^{-\lambda}\lambda \frac{\Gamma(a+1)}{\Gamma(a+2)} \textstyle\sum_{k=1}^{\infty} \frac{\Gamma(a+2)}{\Gamma(k+a+1)} \frac{\lambda^{k-1}}{k}.
\end{aligned}$$

By reparameterising $m = k - 1$, it follows that

$$\begin{aligned}
E\left(w\left(N;\phi\right)\right) &= ae^{-\lambda}\lambda \frac{\Gamma(a+1)}{\Gamma(a+2)} \textstyle\sum_{m=0}^{\infty} \frac{\Gamma(a+2)}{\Gamma(k+a+1)} \frac{\lambda^{m}}{m+1} \\
&= ae^{-\lambda}\lambda \frac{\Gamma(a+1)}{\Gamma(a+2)} \textstyle\sum_{m=0}^{\infty} \frac{\Gamma(m+2)}{\Gamma(m+2)} \frac{\Gamma(a+2)}{\Gamma(k+a+1)} \frac{\lambda^{m}}{m+1} \\
&= ae^{-\lambda}\lambda \frac{\Gamma(a+1)}{\Gamma(a+2)} \textstyle\sum_{m=0}^{\infty} \frac{\Gamma(m+1)}{\Gamma(m+2)} \frac{\Gamma(a+2)}{\Gamma(k+a+1)} \lambda^{m} \\
&= ae^{-\lambda}\lambda \frac{\Gamma(1+a)}{\Gamma(2+a)} \textstyle\sum_{m=0}^{\infty} \frac{\Gamma(1+m)\Gamma(1+m)}{\Gamma(m+2)} \frac{\Gamma(a+2)}{\Gamma(m+a+2)} \frac{\lambda^{m}}{m!} \\
&= ae^{-\lambda}\lambda \frac{\Gamma(1+a)}{\Gamma(2+a)} \textstyle\sum_{m=0}^{\infty} \frac{\Gamma(1+m)}{\Gamma(1)} \frac{\Gamma(1+m)}{\Gamma(1)} \frac{\Gamma(2)}{\Gamma(2+m)} \frac{\Gamma(a+2)}{\Gamma(2+a+m)} \frac{\lambda^{m}}{m!} \\
&= ae^{-\lambda}\lambda \frac{\Gamma(1+a)}{\Gamma(2+a)} {}_2F_2\left(1,1;2,2+a;\lambda\right).
\end{aligned}$$

From the definition of the weighted Poisson probability mass function, it follows that

$$\begin{aligned}
f_w\left(n\right) &= \frac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\
&= \frac{a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}}{ae^{-\lambda}\lambda\frac{\Gamma(a+1)}{\Gamma(a+2)} {}_2F_2(1,1;2,2+a;\lambda)} \frac{e^{-\lambda}\lambda^{n}}{n!} \\
&= \frac{\Gamma(n)\Gamma(2+a)}{\Gamma(n+a+1) {}_2F_2(1,1;2,2+a;\lambda)} \frac{\lambda^{n-1}}{n!}.
\end{aligned}$$

From the definition of the probability generating function, it follows that

$$\begin{aligned}
g(z) &= \sum_{k=0}^{\infty} f_w(k)\, z^k \\
&= \sum_{k=0}^{\infty} \frac{\Gamma(k)\Gamma(a+2)}{\Gamma(k+a+1)\,{}_2F_2(1,1;2,2+a;\lambda)} \frac{\lambda^{k-1}}{k!} z^k \\
&= \frac{z}{{}_2F_2(1,1;2,2+a;\lambda)} \sum_{k=0}^{\infty} \frac{\Gamma(k)\Gamma(a+2)}{\Gamma(k+a+1)} \frac{(\lambda z)^{k-1}}{k!} \\
&= \frac{z}{{}_2F_2(1,1;2,2+a;\lambda)} \sum_{k=1}^{\infty} \frac{\Gamma(a+2)}{\Gamma(k+a+1)} \frac{(\lambda z)^{k-1}}{k}.
\end{aligned}$$

By reparameterising $m = k - 1$, it follows that

$$\begin{aligned}
g(z) &= \frac{z}{{}_2F_2(1,1;2,2+a;\lambda)} \sum_{m=0}^{\infty} \frac{\Gamma(a+2)}{\Gamma(k+a+1)} \frac{(\lambda z)^m}{m+1} \\
&= \frac{z}{{}_2F_2(1,1;2,2+a;\lambda)} \sum_{m=0}^{\infty} \frac{\Gamma(m+2)}{\Gamma(m+2)} \frac{\Gamma(a+2)}{\Gamma(k+a+1)} \frac{(\lambda z)^m}{m+1} \\
&= \frac{z}{{}_2F_2(1,1;2,2+a;\lambda)} \sum_{m=0}^{\infty} \frac{\Gamma(m+1)}{\Gamma(m+2)} \frac{\Gamma(a+2)}{\Gamma(k+a+1)} (\lambda z)^m \\
&= \frac{z}{{}_2F_2(1,1;2,2+a;\lambda)} \sum_{m=0}^{\infty} \frac{\Gamma(1+m)\Gamma(1+m)}{\Gamma(m+2)} \frac{\Gamma(a+2)}{\Gamma(m+a+2)} \frac{(\lambda z)^m}{m!} \\
&= \frac{z}{{}_2F_2(1,1;2,2+a;\lambda)} \sum_{m=0}^{\infty} \frac{\Gamma(1+m)}{\Gamma(1)} \frac{\Gamma(1+m)}{\Gamma(1)} \frac{\Gamma(2)}{\Gamma(2+m)} \frac{\Gamma(a+2)}{\Gamma(2+a+m)} \frac{(\lambda z)^m}{m!}.
\end{aligned}$$

The sum in the above equation is ${}_2F_2(1,1;2,2+a;\lambda z)$, and thus it follows that

$$g(z) = \frac{z\,{}_2F_2(1,1;2,2+a;\lambda z)}{{}_2F_2(1,1;2,2+a;\lambda)}.$$

From the definition of the expected value, it follows that

$$E(N^w) = \lim_{z\to -1} \frac{\partial}{\partial z} g(z).$$

However,

$$\frac{\partial}{\partial z} g(z) = \frac{\partial}{\partial z}\left(\frac{z\,{}_2F_2(1,1;2,2+a;\lambda z)}{{}_2F_2(1,1;2,2+a;\lambda)}\right).$$

The derivative of generalised hypergeometric functions with respect to z is well known. This result is given in Theorem 10.6. Consequently,

$$\begin{aligned}
\frac{\partial}{\partial z} g(z) &= \frac{\frac{\partial}{\partial z} z\,{}_2F_2(1,1;2,2+a;\lambda z)}{{}_2F_2(1,1;2,2+a;\lambda)} \\
&= \frac{{}_2F_2(1,1;2,2+a;\lambda z) + \frac{z\lambda}{2(2+a)}\,{}_2F_2(2,2;3,3+a;\lambda z)}{{}_2F_2(1,1;2,2+a;\lambda)}.
\end{aligned}$$

It then follows that

$$\begin{aligned}
E(N^w) &= \lim_{z\to -1} \frac{{}_2F_2(1,1;2,2+a;\lambda z) + \frac{z\lambda}{2(2+a)}\,{}_2F_2(2,2;3,3+a;\lambda z)}{{}_2F_2(1,1;2,2+a;\lambda)} \\
&= \frac{{}_2F_2(1,1;2,2+a;\lambda) + \frac{\lambda}{2(2+a)}\,{}_2F_2(2,2;3,3+a;\lambda)}{{}_2F_2(1,1;2,2+a;\lambda)}.
\end{aligned}$$

From the definition of the variance, it follows that

$$Var\left(N^w\right) = \lim_{z\to -1}\frac{\partial^2}{\partial z^2}g\left(z\right) + E\left(N^w\right) - \left(E\left(N^w\right)\right)^2.$$

However,

$$\frac{\partial^2}{\partial z^2}g\left(z\right) = \frac{\partial}{\partial z}\left(\frac{{}_2F_2(1,1;2,2+a;\lambda z)+\frac{z\lambda}{2(2+a)}{}_2F_2(2,2;3,3+a;\lambda z)}{{}_2F_2(1,1;2,2+a;\lambda)}\right)$$

$$= \frac{\frac{\lambda}{2(2+a)}{}_2F_2(2,2;3,3+a;\lambda z)+\frac{\lambda}{2(2+a)}{}_2F_2(2,2;3,3+a;\lambda z)+\frac{z2^2\lambda^2}{2(2+a)3(3+a)}{}_2F_2(3,3;4,4+a;\lambda z)}{{}_2F_2(1,1;2,2+a;\lambda)}$$

$$= \frac{\frac{\lambda}{(2+a)}{}_2F_2(2,2;3,3+a;\lambda z)+\frac{z2\lambda^2}{(2+a)3(3+a)}{}_2F_2(3,3;4,4+a;\lambda z)}{{}_2F_2(1,1;2,2+a;\lambda)}.$$

Consequently, the variance can be expressed as

$$Var\left(N^w\right) = \lim_{z\to -1}\frac{\partial^2}{\partial z^2}g\left(z\right) + E\left(N^w\right) - \left(E\left(N^w\right)\right)^2$$

$$= \lim_{z\to -1}\frac{\frac{\lambda}{(2+a)}{}_2F_2(2,2;3,3+a;\lambda z)+\frac{z2\lambda^2}{(2+a)3(3+a)}{}_2F_2(3,3;4,4+a;\lambda z)}{{}_2F_2(1,1;2,2+a;\lambda)} + E\left(N^w\right) - \left(E\left(N^w\right)\right)^2$$

$$= \frac{\frac{\lambda}{(2+a)}{}_2F_2(2,2;3,3+a;\lambda)+\frac{2\lambda^2}{(2+a)3(3+a)}{}_2F_2(3,3;4,4+a;\lambda)}{{}_2F_2(1,1;2,2+a;\lambda)} + E\left(N^w\right) - \left(E\left(N^w\right)\right)^2.$$

□

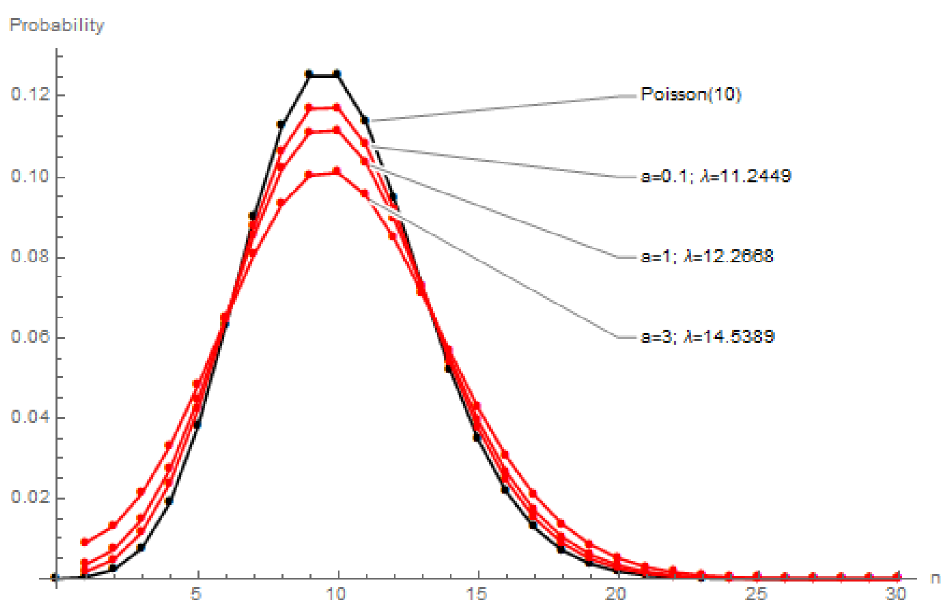


Figure 4.17: Probability mass function - $w\left(n;\phi\right) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$, small a

The variances associated with the weighted Poisson probability mass functions in the above plots are 11.4508, 12.6751 and 15.4346 for a equal to 0.1, 1 and 3 respectively.

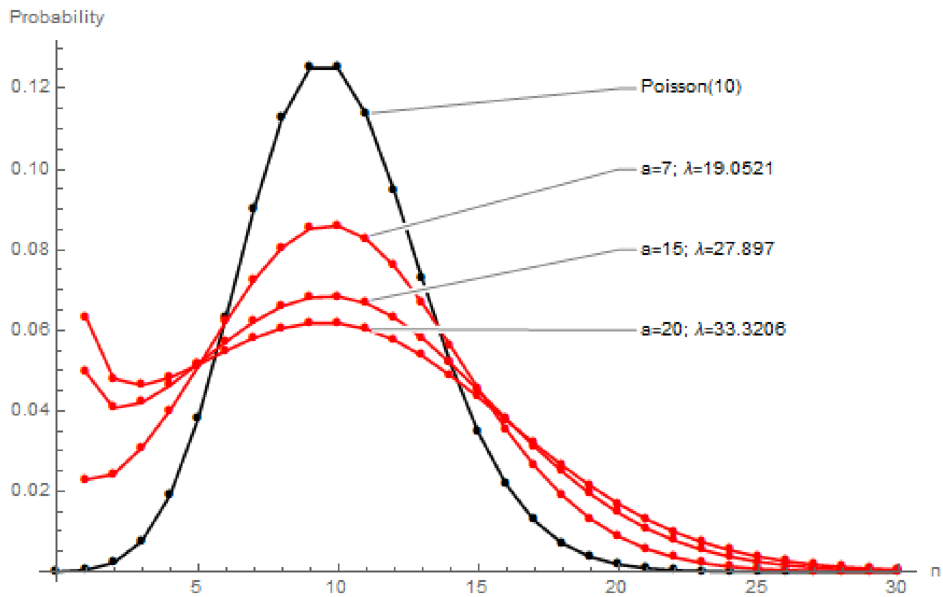


Figure 4.18: Probability mass function - $w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$, large a

The variances associated with the weighted Poisson probability mass functions in the above plots are 20.7045, 29.897 and 34.5357 for a equal to 7, 15 and 20 respectively. This distribution can assume a very wide range of shapes. When a is very large (larger than the values in the above graphs), the distribution assumes an almost "exponential decay" shape. Additionally, if λ is large the tails of this distribution can be quite thick. (As will be seen in Figure 10.27.)

4.2.4 $w(n;\phi) = \frac{ab^a}{n^{a+1}}$

If $w(n;\phi)$ is assumed to be the probability density function of the Pareto distribution the following results are obtained.

Theorem 4.10. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \frac{ab^a}{n^{a+1}}$ *then the following results hold:*

$$E(w(N;\phi)) = \frac{ae^{-\lambda}\lambda^b\left({}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)\right)}{b!b}.$$

$$f_w(n) = \frac{\lambda^{n-b}b^{a+1}b!}{n^{a+1}n!\,{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)}.$$

$$g(z) = \frac{z^b\,{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda z)}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)}.$$

$$E(N^w) = \frac{b\,{}_{a+1}F_{a+1}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)}.$$

$$Var(N^w) = \frac{b^2\,{}_aF_a(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda) - b\,{}_{a+1}F_{a+1}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)} + E(N^w) - (E(N^w))^2.$$

Note: The classical Pareto distribution has continuous support, while the weighted Poisson has discrete support. While it might initially seem that the Pareto probability density function and Poisson probability mass function are incompatible, this is not the case. The only requirement for a weight function is that it must be non-negative on $\mathbb{N}_0$. This restriction is clearly met by any probability mass/density function.

Restrictions:

- Domain: $n \in \{b, b+1, b+2, ...\}$. From the Pareto distribution, the parameter b is defined to be the minimum possible value that the distribution can assume. Since the Poisson distribution is discrete, the domain of the weighted Poisson distribution becomes $\{b, b+1, b+2, ...\}$.

- Parameters: $a \in \mathbb{N}_1, b \in \mathbb{N}_1, \lambda > 0$. These restrictions originate from the expressions given above. If $b = 0$, then $f_w(n) = 0$. The above equations all include a ${}_{a+2}F_{a+2}(;;)$ term. The number of parameters of a generalised hypergeometric function must be integers, and if $a = 0$ then $w(n;\phi) = \frac{0}{n^1} = 0$. It should also be noted that the number of finite moments of the distribution is directly linked to the value a. (Only the first a moments exist.)

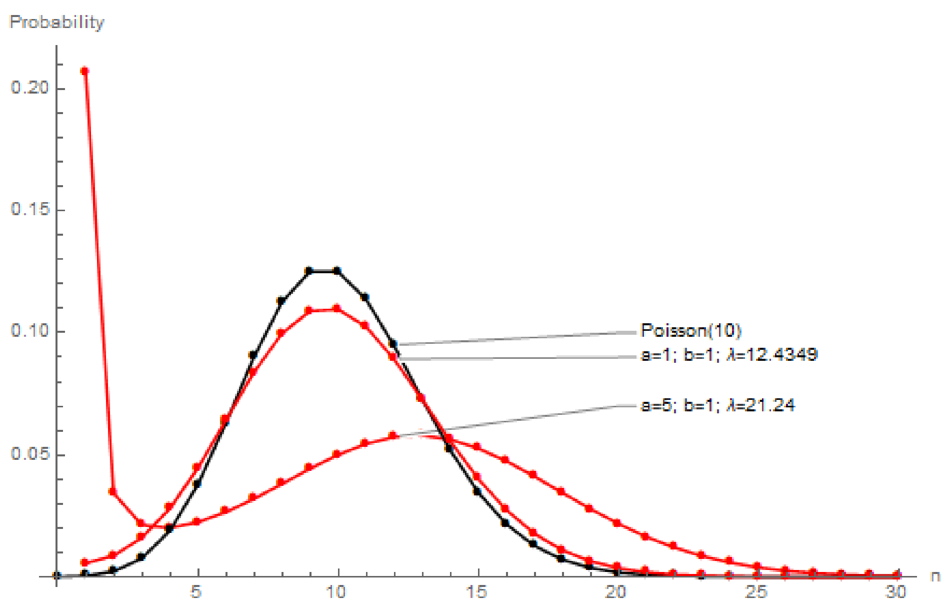


Figure 4.19: Probability mass function - $w(n;\phi) = \frac{ab^a}{n^{a+1}}$, varying a

The variances associated with the weighted Poisson probability mass functions in the above plots are 13.2257 and 44.6502 for a equal to 1 and 5 respectively.

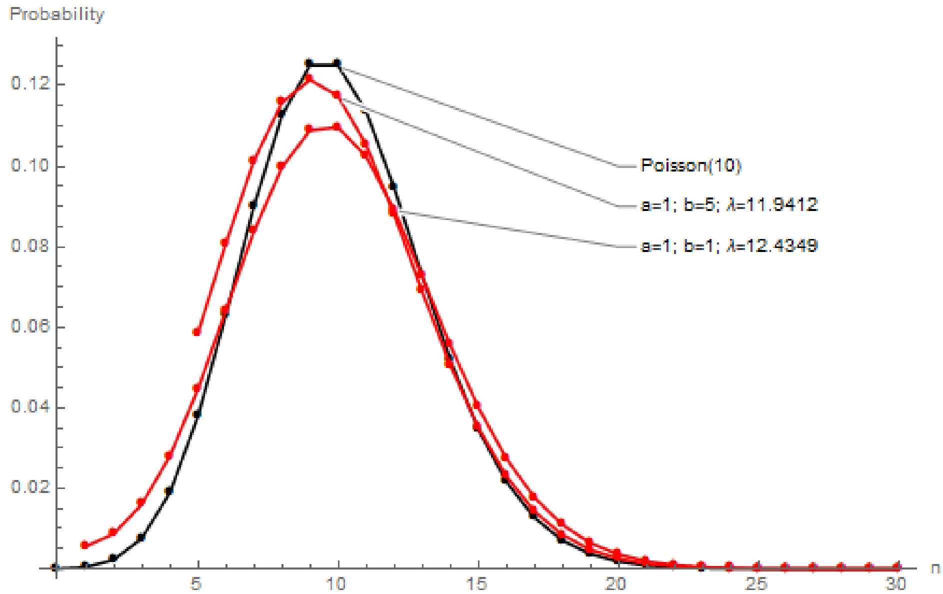


Figure 4.20: Probability mass function - $w\left(n;\phi\right)=\frac{ab^a}{n^{a+1}}$, varying b (small)

The variances associated with the weighted Poisson probability mass functions in the above plots are 13.2257 and 10.1139 for b equal to 1 and 5 respectively.

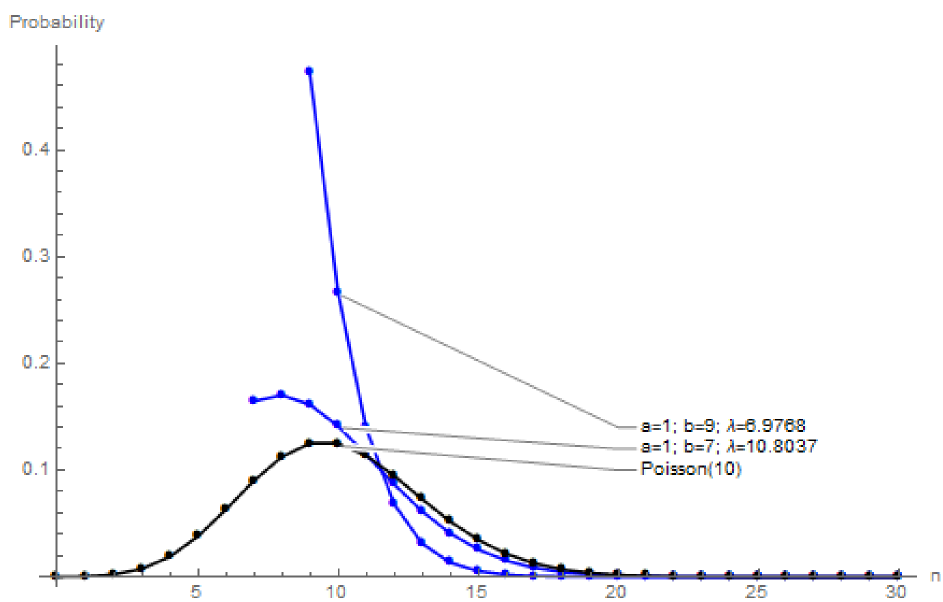


Figure 4.21: Probability mass function - $w\left(n;\phi\right)=\frac{ab^a}{n^{a+1}}$, varying b (large)

The variances associated with the weighted Poisson probability mass functions in the above plots are 6.3313 and 1.6567 for b equal to 7 and 9 respectively. For the majority of parameter combinations considered this distribution is overdispersed; however, if b is close to the mean of the data, it will be underdispersed.

4.2.5 $w(n;\phi) = \frac{-1}{\ln(1-p)}\frac{p^n}{n}$

If $w(n;\phi)$ is assumed to be the probability mass function of the logarithmic distribution the following results are obtained.

Theorem 4.11. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \frac{-1}{\ln(1-p)}\frac{p^n}{n}$ *then the following results hold:*

$$
\begin{aligned}
E(w(N;\phi)) &= \tfrac{-p}{\ln(1-p)}e^{-\lambda}\lambda {}_2F_2(1,1;2,2;\lambda p). \\
f_w(n) &= \tfrac{(p\lambda)^{n-1}}{n.n!{}_2F_2(1,1;2,2;\lambda p)}. \\
g(z) &= \tfrac{z {}_2F_2(1,1;2,2;\lambda pz)}{{}_2F_2(1,1;2,2;\lambda p)}. \\
E(N^w) &= \tfrac{{}_1F_1(1;2;\lambda p)}{{}_2F_2(1,1;2,2;\lambda p)}. \\
Var(N^w) &= \tfrac{\lambda p {}_1F_1(2;3;\lambda p)}{2 {}_2F_2(1,1;2,2;\lambda p)} + E(N^w) - (E(N^w))^2.
\end{aligned}
$$

Restrictions:

- Domain: $n \in \mathbb{N}_1$.
- Parameters: $0 < p < 1, \lambda > 0$.

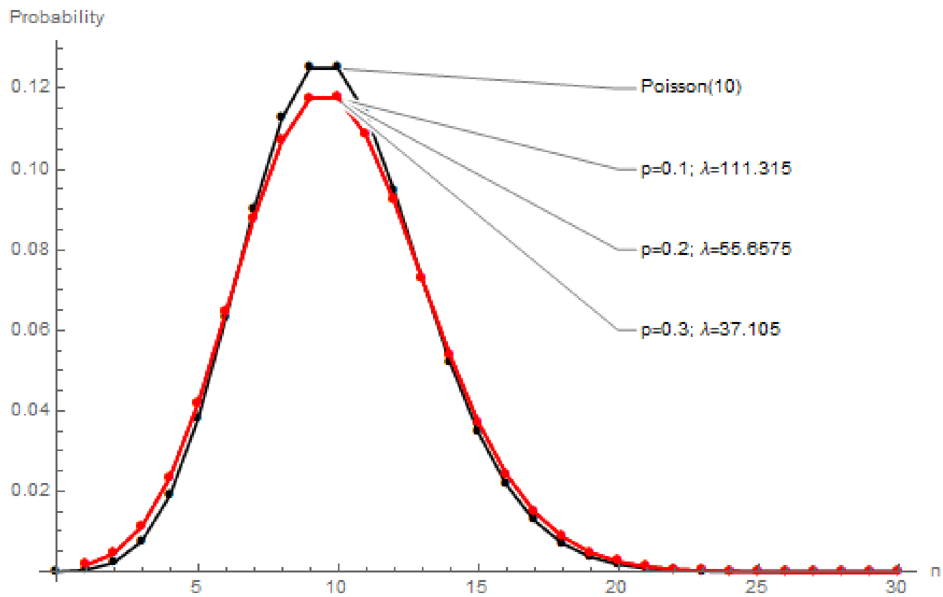


Figure 4.22: Probability mass function - $w(n;\phi) = \frac{-1}{\ln(1-p)} \frac{p^n}{n}$

The variance associated with the weighted Poisson probability mass functions in the above plot is 11.3163 for all values of p. Similar to Theorem 4.8, when p decreases, λ increases. The interaction between these parameters is of such a nature that an increase in one negates the decrease in the other.

4.2.6 $w(n;\phi) = \frac{\Gamma(r+n)}{n!\Gamma(r)} \frac{Beta(a+r,b+n)}{Beta(a,b)}$

If $w(n;\phi)$ is assumed to be the probability mass function of the beta-negative binomial distribution the following results are obtained.

Theorem 4.12. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \frac{\Gamma(r+n)}{n!\Gamma(r)} \frac{Beta(a+r,b+n)}{Beta(a,b)}$ *then the following results hold:*

$$E(w(N;\phi)) = \frac{e^{-\lambda} Beta(a+r,b) {}_2F_2(b,r;1,a+b+r;\lambda)}{Beta(a,b)}.$$

$$f_w(n) = \frac{\lambda^n Beta(a+r,b+n)\Gamma(r+n)}{Beta(a+r,b)(n!)^2\Gamma(r) {}_2F_2(b,r;1,a+b+r;\lambda)}.$$

$$g(z) = \frac{{}_2F_2(b,r;1,a+b+r;\lambda z)}{{}_2F_2(b,r;1,a+b+r;\lambda)}.$$

$$E(N^w) = \frac{br\lambda {}_2F_2(1+b,1+r;2,1+a+b+r;\lambda)}{(a+b+r) {}_2F_2(b,r;1,a+b+r;\lambda)}.$$

$$Var(N^w) = +\frac{b(1+b)r(1+r)\lambda^2 {}_2F_2(2+b,2+r;3,2+a+b+r;\lambda)}{2(a+b+r)(1+a+b+r) {}_2F_2(b,r;1,a+b+r;\lambda)} + E(N^w) - (E(N^w))^2.$$

Restrictions:

- Domain: $n \in \mathbb{N}_0$.
- Parameters: $a, b, r, \lambda > 0$.

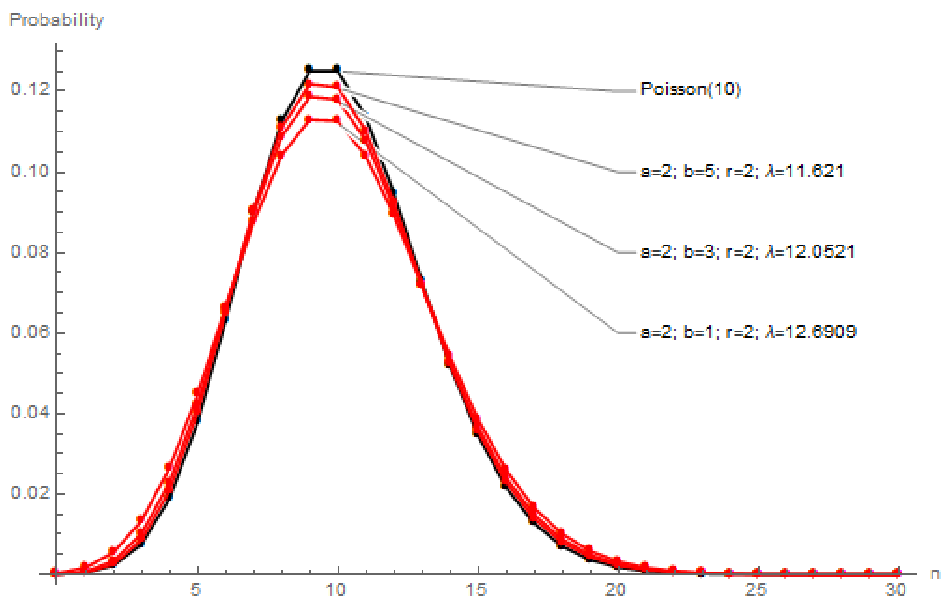


Figure 4.23: Probability mass function - $w\left(n;\phi\right)=\frac{\Gamma(r+n)}{n!\Gamma(r)}\frac{Beta(a+r,b+n)}{Beta(a,b)}$, varying b

The variances associated with the weighted Poisson probability mass functions in the above plots are 12.2903, 11.1712 and 10.6113 for b equal to 1, 3 and 5 respectively.

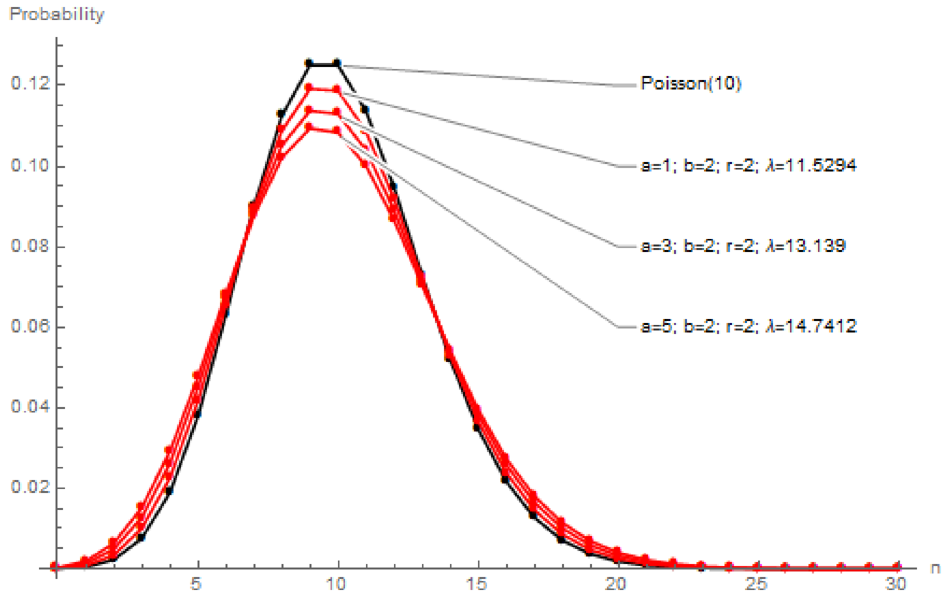


Figure 4.24: Probability mass function - $w(n;\phi) = \frac{\Gamma(r+n)}{n!\Gamma(r)} \frac{Beta(a+r,b+n)}{Beta(a,b)}$, varying a

The variances associated with the weighted Poisson probability mass functions in the above plots are 11.0492, 12.1551 and 13.1668 for a equal to 1, 3 and 5 respectively.

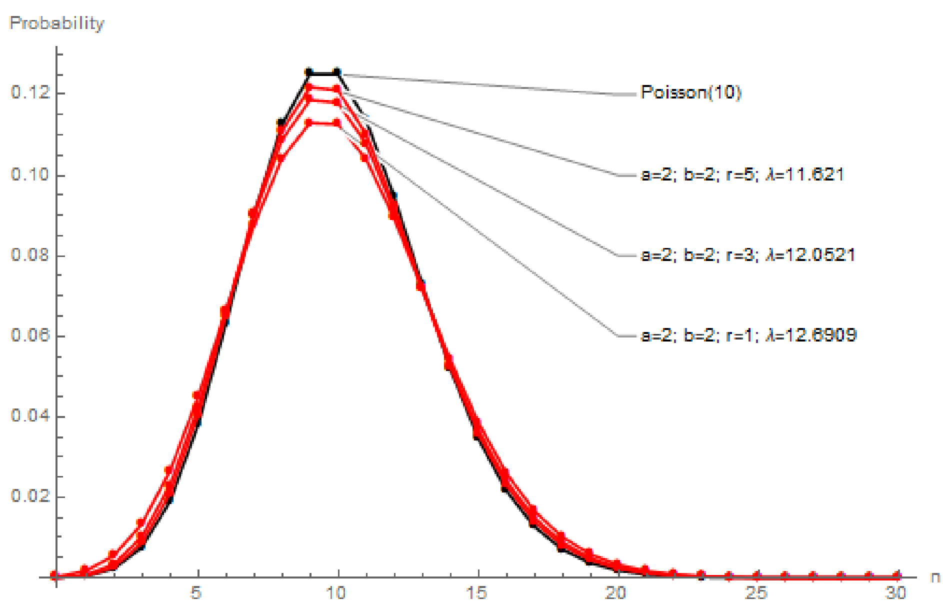


Figure 4.25: Probability mass function - $w(n;\phi) = \frac{\Gamma(r+n)}{n!\Gamma(r)} \frac{Beta(a+r,b+n)}{Beta(a,b)}$, varying r

The variances associated with the weighted Poisson probability mass functions in the above plots are 12.2903, 11.1712 and 10.6113 for r equal to 1, 3 and 5 respectively.

4.3 Truncating weight functions

4.3.1 $w(n;\phi) = I(n \geq a)$

Theorem 4.13. *If the weight function used in the weighted Poisson probability mass function is chosen as $w(n;\phi) = I(n \geq a)$ then*

$$E(w(N;\phi)) = \frac{\gamma(a,\lambda)}{\Gamma(a)}.$$

$$f_w(n) = \frac{I(n\geq a)e^{-\lambda}\lambda^n\Gamma(a)}{n!\gamma(a,\lambda)}.$$

$$g(z) = \frac{e^{\lambda(z-1)}\gamma(a,\lambda z)}{\gamma(a,\lambda)}.$$

$$E(N^w) = \frac{e^{-\lambda}\lambda^a+\lambda\gamma(a,\lambda)}{\gamma(a,\lambda)}.$$

$$Var(N^w) = \frac{e^{-\lambda}(\lambda^{1+a}+\lambda^a(a-1)+e^\lambda\lambda^2\gamma(a,\lambda))}{\gamma(a,\lambda)} + E(N^w) - (E(N^w))^2.$$

Restrictions:

- Domain: $n \in \{a, a+1, ...\}$.
- Parameters: $a \in \mathbb{N}_0, \lambda > 0$. The restriction that $a \in \mathbb{N}_0$ is a practical one since this weighted Poisson distribution is only defined on $n \in \mathbb{N}_0$. If a were assumed to be a real number between two integers, $b_1 < b_2$, the parameter used in the equations below would be b_1, since the decimal component of a does not affect the mapping of the discrete distribution. A similar argument holds for the next few weight functions.

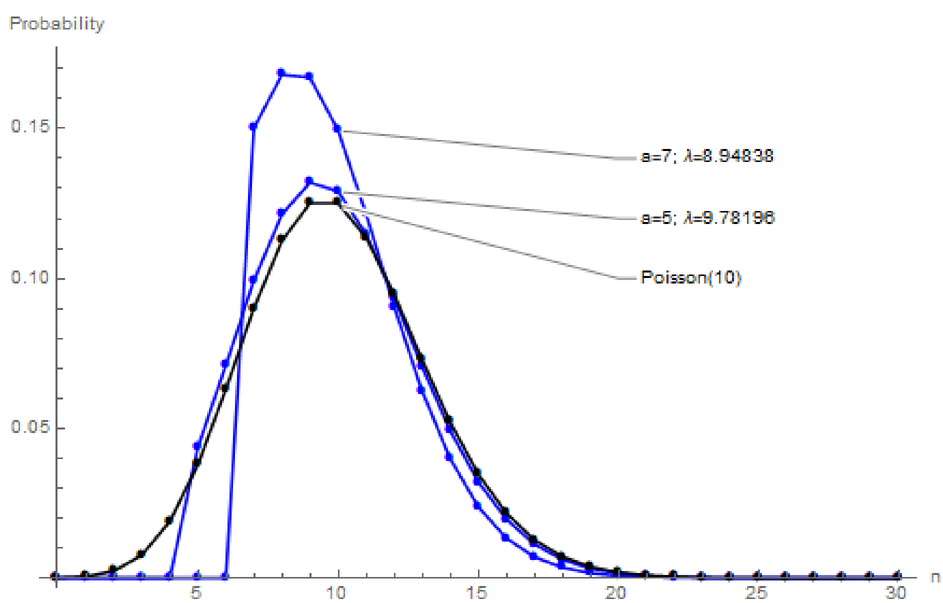


Figure 4.26: Probability mass function - $w(n;\phi) = I(n \geq a)$

The variances associated with the weighted Poisson probability mass functions in the above plots are 8.6918 and 5.7935 for a equal to 5 and 7 respectively.

4.3.2 $w(n;\phi) = I(n \le b)$

Theorem 4.14. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = I(n \le b)$ *then*

$$\begin{aligned} E(w(N;\phi)) &= \tfrac{\Gamma(1+b,\lambda)}{\Gamma(1+b)}. \\ f_w(n) &= \tfrac{I(n\le b)e^{-\lambda}\lambda^n\Gamma(1+b)}{n!\Gamma(1+b,\lambda)}. \\ g(z) &= \tfrac{e^{\lambda(z-1)}\Gamma(1+b,\lambda z)}{\Gamma(1+b,\lambda)}. \\ E(N^w) &= \lambda - \tfrac{e^{-\lambda}\lambda^{1+b}}{\Gamma(1+b,\lambda)}. \\ Var(N^w) &= \lambda\left(\lambda - \tfrac{e^{-\lambda}\lambda^b(\lambda+b)}{\Gamma(1+b,\lambda)}\right) + E(N^w) - (E(N^w))^2. \end{aligned}$$

Restrictions:

- Domain: $n \in \{0, 1, ..., b-1, b\}$.
- Parameters: $b \in \mathbb{N}_0, \lambda > 0$.

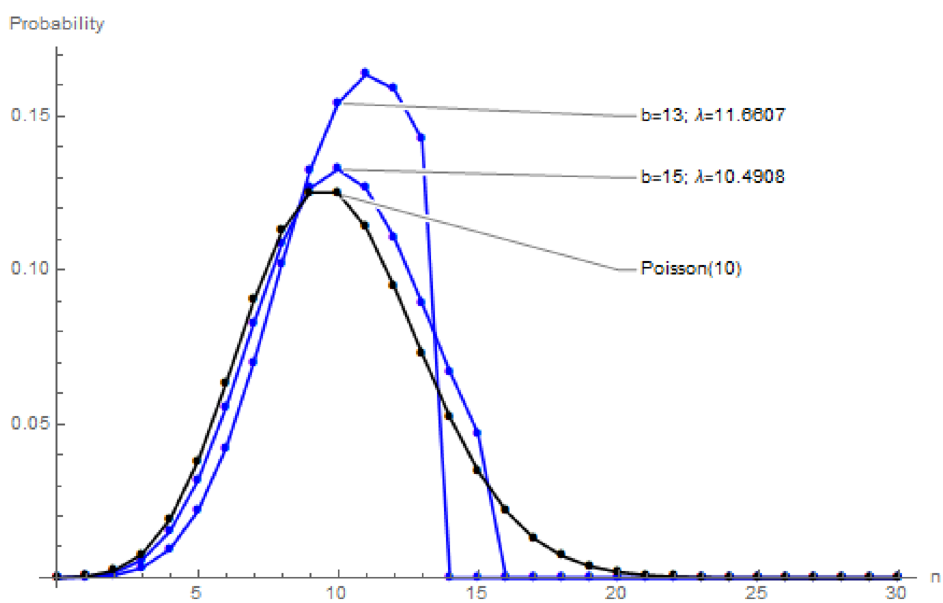


Figure 4.27: Probability mass function - $w(n;\phi) = I(n \le b)$

The variances associated with the weighted Poisson probability mass functions in the above plots are 5.0179 and 7.5461 for b equal to 13 and 15 respectively.

4.3.3 $w(n;\phi) = I(n \geq a)\, I(n \leq b)$

Theorem 4.15. *If the weight function used in the weighted Poisson probability mass function is chosen as $w(n;\phi) = I(n \geq a)\, I(n \leq b)$ then*

$$E(w(N;\phi)) = \frac{\Gamma(a)\Gamma(1+b,\lambda)-\Gamma(1+b)\Gamma(a,\lambda)}{\Gamma(a)\Gamma(1+b)}.$$

$$f_w(n) = \frac{\Gamma(a)\Gamma(1+b)}{\Gamma(a)\Gamma(1+b,\lambda)-\Gamma(1+b)\Gamma(a,\lambda)} \frac{I(n\geq a)I(n\leq b)e^{-\lambda}\lambda^n}{n!}.$$

$$g(z) = \frac{e^{\lambda(z-1)}(\Gamma(1+b)\Gamma(a,\lambda z)-\Gamma(a)\Gamma(1+b,\lambda z))}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)}.$$

$$E(N^w) = \frac{e^{-\lambda}\left(\lambda\Gamma(a)\left(\lambda^b-e^{\lambda}\Gamma(1+b,\lambda)\right)-\Gamma(1+b)\left(\lambda^a-e^{\lambda}\lambda\Gamma(a,\lambda)\right)\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)}.$$

$$\begin{aligned} Var(N^w) &= \frac{e^{-\lambda}\left(\left(\lambda^{2+b}+b\lambda^{1+b}\right)\Gamma(a)+\left(\lambda^a-\lambda^{1+a}-a\lambda^a\right)\Gamma(1+b)+e^{\lambda}\lambda^2(\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda))\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)} \\ &+ E(N^w) - (E(N^w))^2. \end{aligned}$$

Restrictions:

- Domain: $n \in \{a, a+1, ..., b-1, b\}$.
- Parameters: $a, b \in \mathbb{N}_0, a < b, \lambda > 0$.

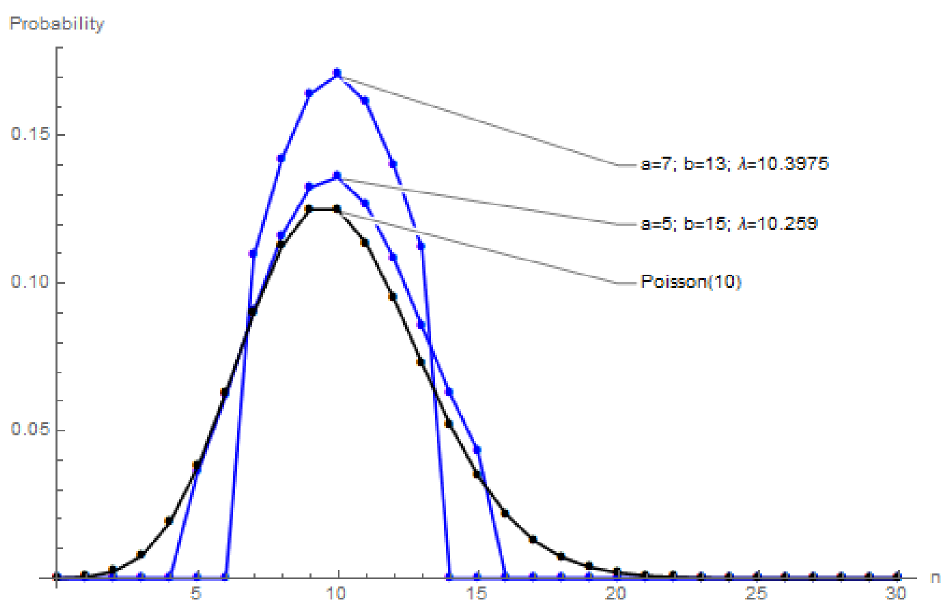


Figure 4.28: Probability mass function - $w(n;\phi) = I(n \geq a)\, I(n \leq b)$

The variances associated with the weighted Poisson probability mass functions in the above plots are 3.4460 and 6.7146 for b equal to 13 and 15 respectively.

4.3.4 $w(n;\phi) = \binom{n}{a}$

Note that if $p = 0.5$ in Theorem 4.8, the resulting weight function will be equal in distribution to the weight function discussed below.

Theorem 4.16. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \binom{n}{a}$ *where* $a \in \{0, 1, 2, ...\}$ *then*

$$
\begin{aligned}
E(w(N;\phi)) &= \tfrac{\lambda^a}{a!}. \\
f_w(n) &= \tfrac{e^{-\lambda}\lambda^{n-a}}{(n-a)!}. \\
g(z) &= e^{\lambda(z-1)}z^a. \\
E(N^w) &= a + \lambda. \\
Var(N^w) &= \lambda.
\end{aligned}
$$

Restrictions:

- Domain: $n \in \{a, a+1, ...\}$.
- Parameters: $a \in \mathbb{N}_0, \lambda > 0$.

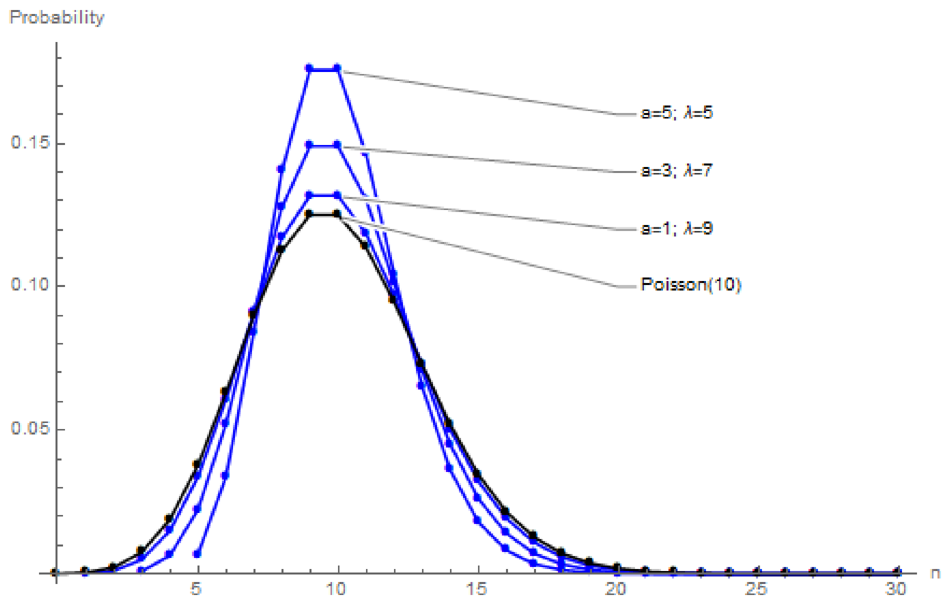


Figure 4.29: Probability mass function - $w(n;\phi) = \begin{pmatrix} n \\ a \end{pmatrix}$

The variances associated with the weighted Poisson probability mass functions in the above plots are 9, 7 and 5 for a equal to 1, 3 and 5 respectively.

4.4 Miscellaneous weight functions

In this section, weight functions that could not be classified into one of the previous categories are discussed.

4.4.1 $w(n;\phi) = \varepsilon \frac{n!}{e^{-\lambda}\lambda^n} I(n=0) + (1-\varepsilon)$

The zero-inflated Poisson distribution is defined as a Poisson distribution that has been modified to have probability ε of realising a value of zero and has the Poisson probability mass function on the rest of the domain, normalised by $1-\varepsilon$.

Theorem 4.17. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \varepsilon \frac{n!}{e^{-\lambda}\lambda^n} I(n=0) + (1-\varepsilon)$ *then*

$$E\left(w\left(N;\phi\right)\right) = 1.$$

$$f_w\left(n\right) = \left(\varepsilon\frac{n!}{e^{-\lambda}\lambda^n}I\left(n=0\right)+\left(1-\varepsilon\right)\right)\frac{e^{-\lambda}\lambda^n}{n!}.$$

$$g\left(z\right) = \varepsilon + e^{\lambda(z-1)}\left(1-\varepsilon\right).$$

$$E\left(N^w\right) = \lambda\left(1-\varepsilon\right).$$

$$Var\left(N^w\right) = \lambda\left(1-\varepsilon\right)\left(1+\varepsilon\lambda\right).$$

Restrictions:

- Domain: $n \in \mathbb{N}_0$ if $0 < \varepsilon < 1$, $n \in \mathbb{N}_1$ if $\varepsilon = 0$ and $n = 0$ if $\varepsilon = 1$.
- Parameters: $0 \leq \varepsilon \leq 1, \lambda > 0$.

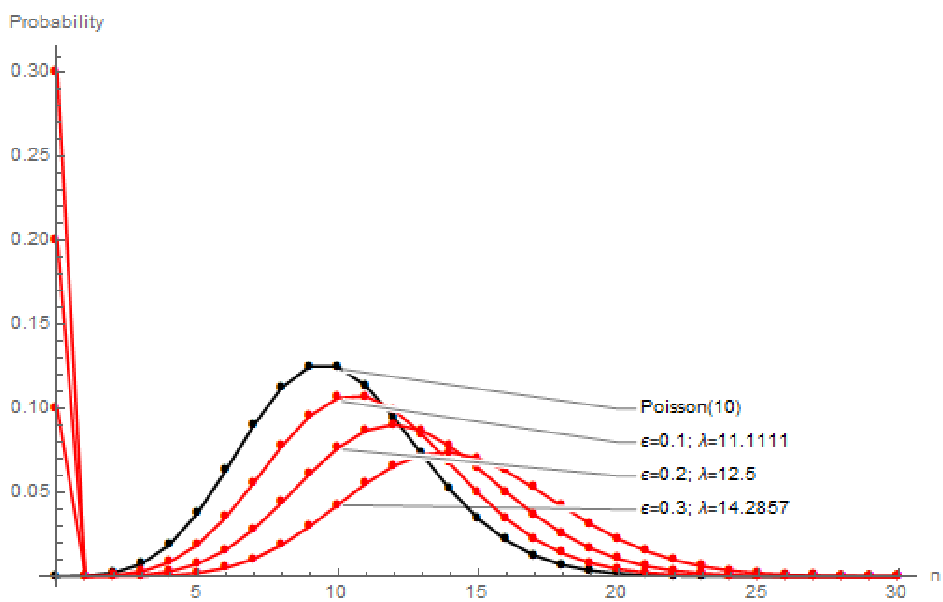


Figure 4.30: Probability mass function - $w\left(n;\phi\right) = \varepsilon\frac{n!}{e^{-\lambda}\lambda^n}I\left(n=0\right)+\left(1-\varepsilon\right)$, small epsilon

The variances associated with the weighted Poisson probability mass functions in the above plots are 21.1111, 35 and 52.857 for ε equal to 0.1, 0.2 and 0.3 respectively.

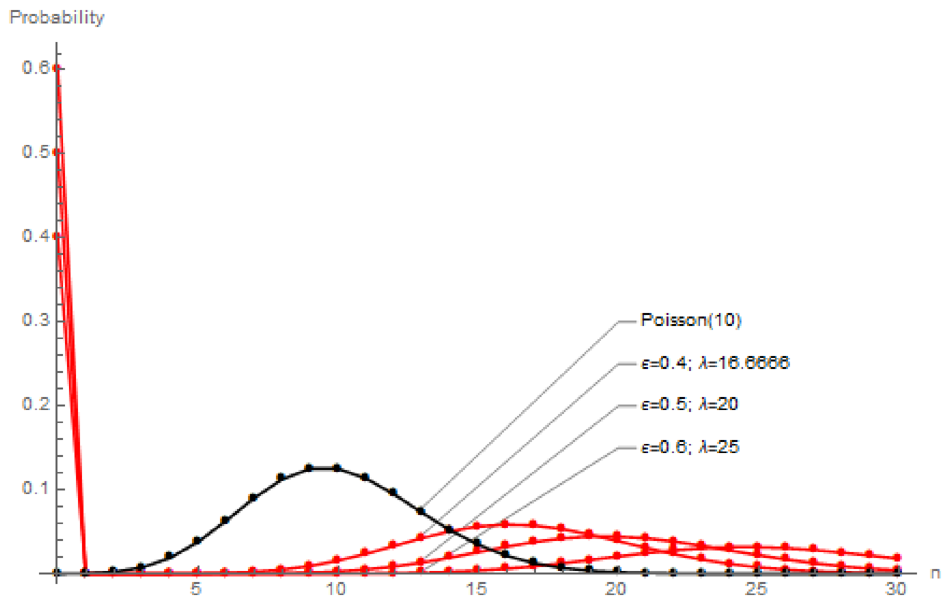


Figure 4.31: Probability mass function - $w(n;\phi) = \varepsilon \frac{n!}{e^{-\lambda}\lambda^n} I(n=0) + (1-\varepsilon)$, medium epsilon

The variances associated with the weighted Poisson probability mass functions in the above plots are 76.6661, 110 and 160 for ε equal to 0.4, 0.5 and 0.6 respectively.

It should be noted that by merely changing the index of the indicator function, it is possible to inflate any part of the probability mass function,
If $w(n;\phi) = \varepsilon \frac{n!}{e^{-\lambda}\lambda^n} I(n=b) + (1-\varepsilon)$ then

$$
\begin{aligned}
E(w(N;\phi)) &= 1. \\
f_w(n) &= \left(\varepsilon \frac{n!}{e^{-\lambda}\lambda^n} I(n=b) + (1-\varepsilon)\right) \frac{e^{-\lambda}\lambda^n}{n!}. \\
g(z) &= \varepsilon z^b + e^{\lambda(z-1)}(1-\varepsilon). \\
E(N^w) &= \lambda + (b-\lambda)\varepsilon. \\
Var(N^w) &= (1-\varepsilon)\left(\varepsilon(\lambda-b)^2 + \lambda\right).
\end{aligned}
$$

4.4.2 $w(n;\phi) = (a)_n = \frac{\Gamma(a+n)}{\Gamma(a)}$

Theorem 4.18. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = (a)_n$ *then*

$$
\begin{aligned}
E(w(N;\phi)) &= e^{-\lambda}(1-\lambda)^{-a}. \\
f_w(n) &= \frac{(1-\lambda)^a \lambda^n (a)_n}{n!}. \\
g(z) &= (1-\lambda)^a (1-\lambda z)^{-a}. \\
E(N^w) &= \frac{a\lambda}{1-\lambda}. \\
Var(N^w) &= \frac{a\lambda}{(\lambda-1)^2}.
\end{aligned}
$$

Restrictions:

- Domain: $n \in \mathbb{N}_0$.
- Parameters: $a > 0, 0 < \lambda < 1$.

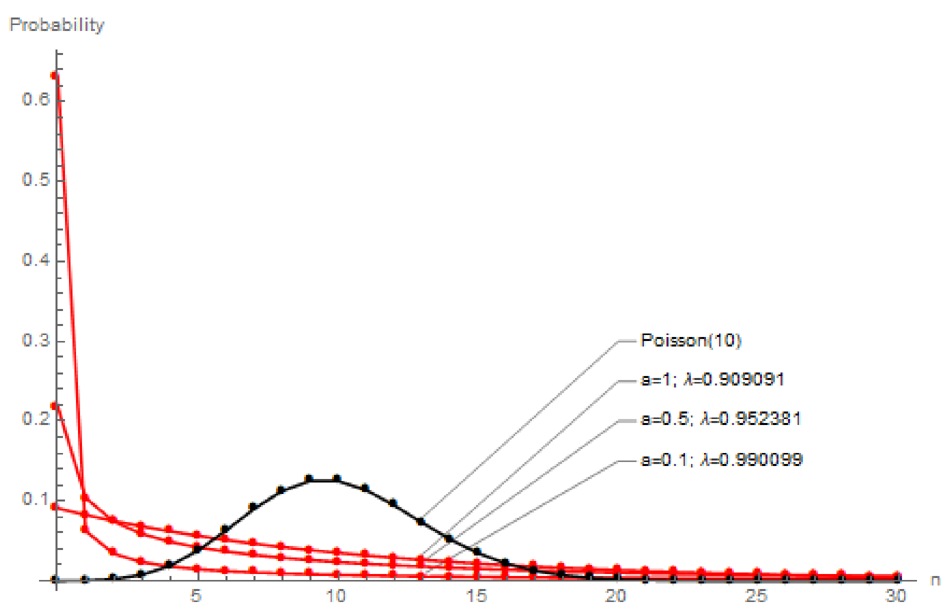


Figure 4.32: Probability mass function - $w(n;\phi) = (a)_n$, small a

The variances associated with the weighted Poisson probability mass functions in the above plots are 1010, 210 and 110 for a equal to 0.1, 0.5 and 1 respectively.

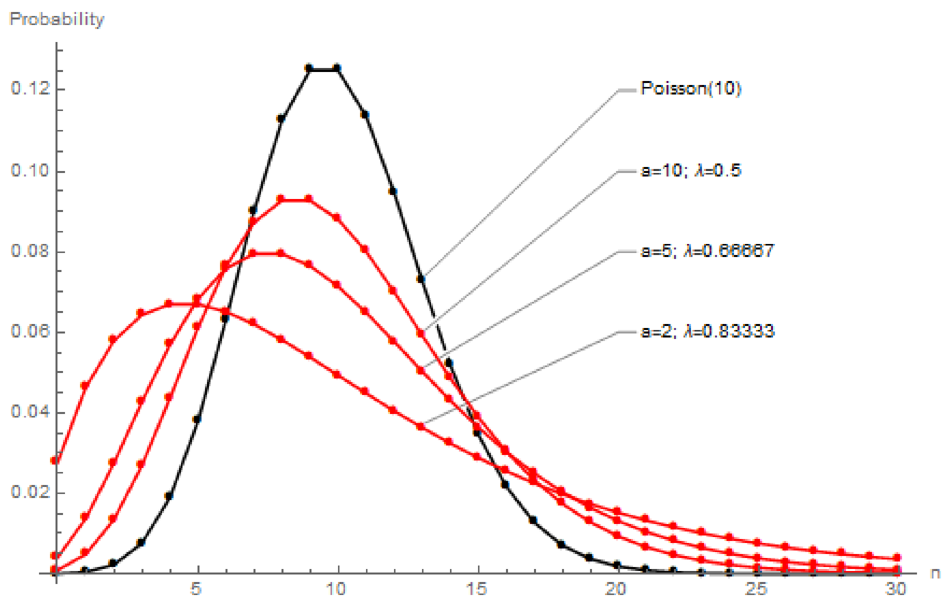


Figure 4.33: Probability mass function - $w(n;\phi) = (a)_n$, large a

The variances associated with the weighted Poisson probability mass functions in the above plots are 54.9974, 30 and 20 for a equal to 2, 5 and 10 respectively.

4.4.3 $w(n;\phi) = (n)_a = \frac{\Gamma(a+n)}{\Gamma(n)}$

Theorem 4.19. *If the weight function used in the weighted Poisson probability mass function is chosen as $w(n;\phi) = (n)_a$ then*

$$\begin{aligned}
E(w(N;\phi)) &= e^{-\lambda}\lambda\Gamma(1+a)\,{}_1F_1(1+a;2;\lambda).\\
f_w(n) &= \frac{\lambda^{n-1}(n)_a}{n!\Gamma(1+a)\,{}_1F_1(1+a;2;\lambda)}.\\
g(z) &= \frac{z\,{}_1F_1(1+a;2;\lambda z)}{{}_1F_1(1+a;2;\lambda)}.\\
E(N^w) &= \frac{2\,{}_1F_1(1+a;2;\lambda)+(1+a)\lambda\,{}_1F_1(2+a;3;\lambda)}{2\,{}_1F_1(1+a;2;\lambda)}.\\
Var(N^w) &= \frac{(1+a)\lambda(6\,{}_1F_1(2+a;3;\lambda)+(2+a)\lambda\,{}_1F_1(3+a;4;\lambda))}{6\,{}_1F_1(1+a;2;\lambda)} + E(N^w) - (E(N^w))^2.
\end{aligned}$$

Restrictions:

- Domain: $n \in \mathbb{N}_1$.
- Parameters: $a > 0, \lambda > 0$.

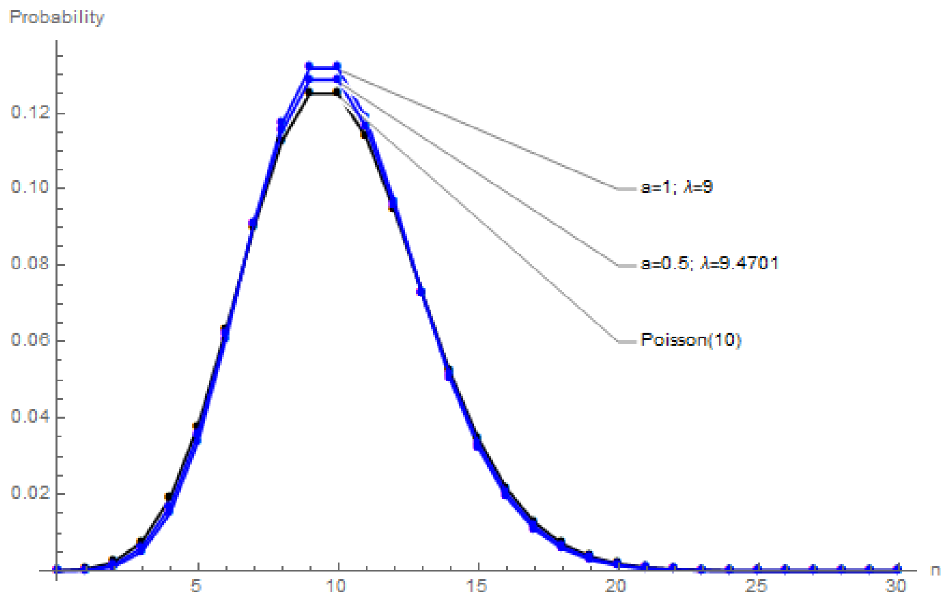


Figure 4.34: Probability mass function - $w(n;\phi) = (n)_a$, small a

The variances associated with the weighted Poisson probability mass functions in the above plots are 9.4353 and 9 for a equal to 0.5 and 1 respectively.

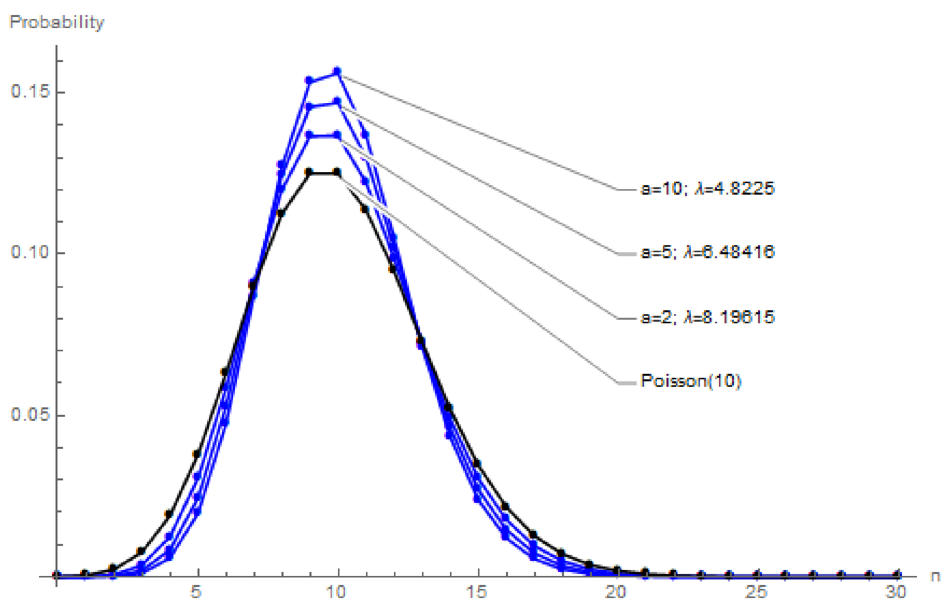


Figure 4.35: Probability mass function - $w(n;\phi) = (n)_a$, large a

The variances associated with the weighted Poisson probability mass functions in the above plots are 8.3538, 7.2624 and 6.4501 for a equal to 2, 5 and 10 respectively.

4.5 Inverted weight functions

The notion of a dual pair of weight functions was discussed in Kokonendji et al. [83] (as well as Section 2.2). In this section, potential dual partners of the previously discussed weight functions be will investigated. As will be seen, while some inverted versions of the weight functions result in valid weighted Poisson distributions, the majority of them are not suitable for the construction of dual partners.

In this section, proofs of the relevant derivations will not be shown. These proofs are omitted due to their similarity to others presented in this chapter and in chapter 10.

In Table 4.2 below the weight functions from the previous part of the chapter that cannot be inverted (or are excluded for other reasons) are listed. Following the table, the various weighted Poisson distributions that have closed-form expressions are discussed.

Table 4.2: Non invertable weight functions

Weight function	Reason for lack of closed-form probability mass function
n^a	While closed-form expressions exist for specific integer values of a, no general formula exists.
$(n+\varepsilon)^{-1}$	Closed-form expressions can be derived for this weight function, however, the expressions include a term $(-\lambda)^{\varepsilon}$. Since $\lambda > 0$ and $\varepsilon \in (0,1)$ the term is complex, and thus the resulting probability mass function is invalid.
$(an^3+bn^2+cn)^{-1}$	Closed-form expressions for this weight function exist. However, the expressions, even after fully simplifying them, are extremely lengthy. For instance, the expressions of the normalising constant and probability mass function are each more than half a page in length, with the other expressions being dramatically longer. For this reason, this weight function has been omitted from the thesis.
$\left((n+a)(n-b)^2\right)^{-1}$	Closed-form expressions can be derived for this weight function, however, the expressions include many terms that lead to contradictions in the parameter restrictions. The most obvious example is that λ is required to be both greater and less than 0. Consequently this weight function does not lead to a valid probability mass function.

$\left(\left(n+a\right)\left(n^2-bn+c\right)\right)^{-1}$	Closed-form expressions can be derived for this weight function, however, the expressions include terms that lead to contradictions in the parameter restrictions. Consequently this weight function does not lead to a valid probability mass function.
$\left(\binom{m}{n}p^n\left(1-p\right)^{m-n}\right)^{-1}$	A general closed-form expression for this weight function does not exist. However, if the parameter m is assumed to be fixed, an expression does exist. Since m is the upper bound of the domain of this specific weighted Poisson distribution, it is not unrealistic to assume that practically the value of m could be determined by real world constrains.
$\left(\frac{ab^a}{n^{a+1}}\right)^{-1}$	While closed-form expressions exist for specific values of a and b, no general formula exists.
$\left(I\left(n\geq a\right)\right)^{-1}$	The non-inverted weight function has 0 probability when $n<a$. By contrast the inverted weighted Poisson distribution is undefined when $n<a$. If the domain of the function is assumed to be the more practical $n\in\{a,a+1,...\}$ the function exists, however, since $\left(I\left(n\geq a\right)\right)^{-1}$ returns a 1 if true, the non-zero values of the probability mass function are the same as the non-inverted version. The same fact applies to all three indicator weight functions.
$\left(I\left(n\leq b\right)\right)^{-1}$	The non-inverted weight function has 0 probability when $n>b$. By contrast the inverted weighted Poisson distribution is undefined when $n>b$.
$\left(I\left(n\geq a\right)I\left(n\leq b\right)\right)^{-1}$	The non-inverted weight function has 0 probability when $n<a$ or $n>b$. By contrast the inverted weighted Poisson distribution is undefined when $n<a$ or $n>b$.
$\left(\varepsilon\frac{n!}{e^{-\lambda}\lambda^n}I\left(n=0\right)+\left(1-\varepsilon\right)\right)^{-1}$	A closed-form expression exists for this specific weight function. However, the shape of the resulting probability mass function is not substantially affected by the value of ε. Since the objective is the construction of flexible distributions, this weight function is omitted below.

4.5.1 $w\left(n;\phi\right)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$

Theorem 4.20. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w\left(n;\phi\right)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$ *then*

$$E\left(w\left(N;\phi\right)\right) = \frac{ce^{-\lambda}{}_2F_2\left(\frac{b}{a},1+c;1+\frac{b}{a},c;\lambda\right)}{b}.$$

$$f_w\left(n\right) = \frac{b\lambda^n}{c\left(a+\frac{b-ac}{n+c}\right)n!\,{}_2F_2\left(\frac{b}{a},1+c;1+\frac{b}{a},c;\lambda\right)}.$$

$$g\left(z\right) = \frac{{}_2F_2\left(\frac{b}{a},1+c;1+\frac{b}{a},c;\lambda z\right)}{{}_2F_2\left(\frac{b}{a},1+c;1+\frac{b}{a},c;\lambda\right)}.$$

$$E\left(N^w\right) = \frac{b\left(ae^{\lambda}(ac+a\lambda-b)+b(b-ac)\left(\Gamma\left(\frac{b}{a}\right)-\Gamma\left(\frac{b}{a},-\lambda\right)\right)(-\lambda)^{-\frac{b}{a}}\right)}{a^3\,{}_2F_2\left(\frac{b}{a},1+c;1+\frac{b}{a},c;\lambda\right)}.$$

$$\begin{aligned} Var\left(N^w\right) &= \frac{ae^{\lambda}(-\lambda)^{\frac{b}{a}}\left(b^2-ab(c+\lambda-1)+a^2\left(\lambda^2+c\lambda-c\right)\right)+b(a+b)(ac-b)\left(\Gamma\left(\frac{b}{a}\right)-\Gamma\left(\frac{b}{a},-\lambda\right)\right)}{a^3e^{\lambda}(-\lambda)^{\frac{b}{a}}+a^2(ac-b)\left(\Gamma\left(\frac{b}{a}\right)-\Gamma\left(\frac{b}{a},-\lambda\right)\right)} \\ &+ E\left(N^w\right)-\left(E\left(N^w\right)\right)^2. \end{aligned}$$

Restrictions:

- Domain: $n \in \mathbb{N}_0$.
- Parameters: $a, b, c > 0, \lambda > 0$.

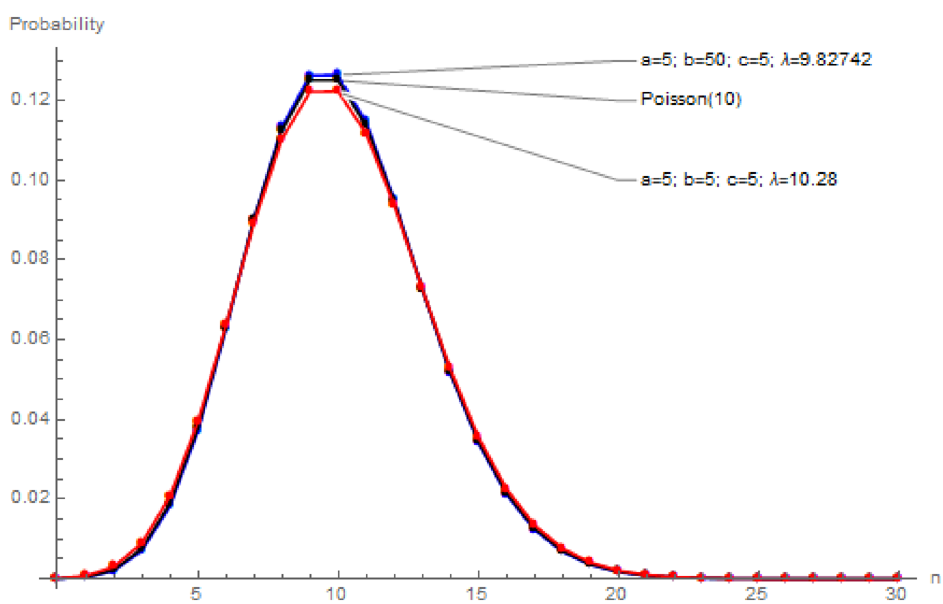


Figure 4.36: Probability mass function - $w\left(n;\phi\right)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$

The variances associated with the weighted Poisson probability mass functions in the above plots are 10.4806 and 9.8041 for b equal to 5 and 50 respectively.

4.5.2 $w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$

Theorem 4.21. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$ *then*

$$
\begin{aligned}
E(w(N;\phi)) &= \frac{e^{-\lambda+\frac{\lambda}{p}}(r-1)\lambda\left(\Gamma(r-1)-\Gamma\left(r-1,\frac{\lambda}{p}\right)\right)}{p\left(\frac{\lambda}{p}(1-p)\right)^r}. \\
f_w(n) &= \frac{\lambda^{n-1}\left(\frac{\lambda}{p}\right)^r}{\binom{n+r-1}{n} n!\left(\Gamma(r-1)-\Gamma\left(r-1,\frac{\lambda}{p}\right)\right)e^{\frac{\lambda}{p}}p^{n-1}(r-1)}. \\
g(z) &= \frac{ze^{\frac{\lambda(z-1)}{p}}\left(\frac{\lambda}{p}\right)^r\left(\Gamma(r-1)-\Gamma\left(r-1,\frac{\lambda z}{p}\right)\right)}{\left(\frac{z\lambda}{p}\right)^r\left(\Gamma(r-1)-\Gamma\left(r-1,\frac{\lambda}{p}\right)\right)}. \\
E(N^w) &= \frac{e^{-\frac{\lambda}{p}}p^2(r-1)\left(\frac{\lambda}{p}\right)^r+\lambda(\lambda+p(1-r))\left(\Gamma(r)-(r-1)\Gamma\left(r-1,\frac{\lambda}{p}\right)\right)}{p(r-1)\lambda\left(\Gamma(r-1)-\Gamma\left(r-1,\frac{\lambda}{p}\right)\right)}. \\
Var(N^w) &= \frac{e^{-\frac{\lambda}{p}}\left(-p^2(r-1)(pr-\lambda)\left(\frac{\lambda}{p}\right)^r+e^{\frac{\lambda}{p}}\lambda(p^2(r-1)r-2p(r-1)\lambda+\lambda^2)\left(\Gamma(r-1)-\Gamma\left(r-1,\frac{\lambda}{p}\right)\right)\right)}{p^2(r-1)\lambda\left(\Gamma(r-1)-\Gamma\left(r-1,\frac{\lambda}{p}\right)\right)} \\
&+ E(N^w)-(E(N^w))^2.
\end{aligned}
$$

Restrictions:

- Domain: $n \in \mathbb{N}_0$.
- Parameters: $r \in \{2,3,...\}, 0 < p < 1, \lambda > 0$.

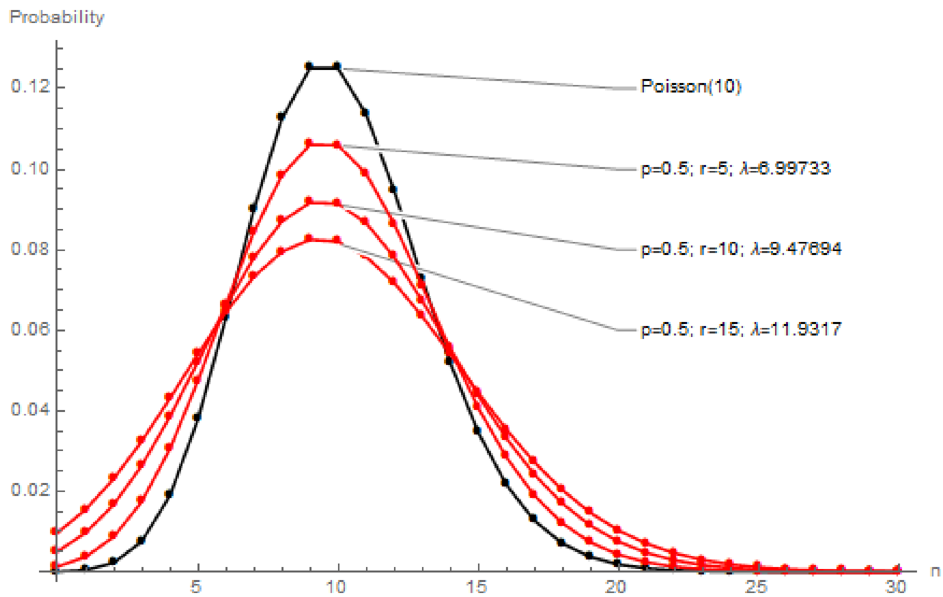


Figure 4.37: Probability mass function - $w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$

The variances associated with the weighted Poisson probability mass functions in the above plots are 13.9412, 18.4926 and 22.4981 for r equal to 5, 10 and 15 respectively.

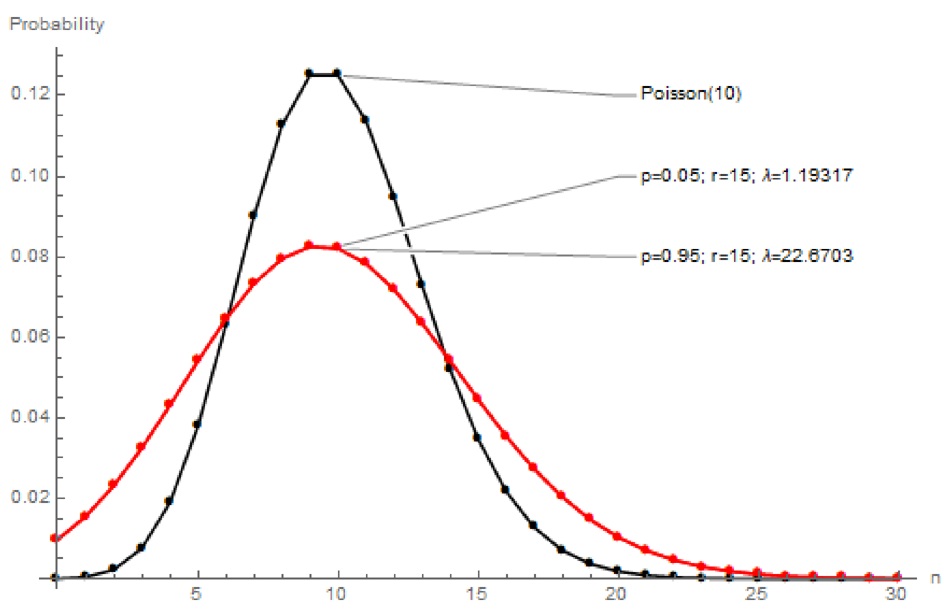


Figure 4.38: Probability mass function - $w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$

The variance associated with the weighted Poisson probability mass functions in the above plot is 22.4982 for both values of p. This, together with Figure 4.38, shows varying the parameter p is completely ineffective in changing the shape of the probability mass function of this specific weighted Poisson distribution (when the expected value is fixed). The reason for this can clearly be seen in the expressions of the equations as well as in the above plot: In all cases, when p decreases, λ is required to increase in order for the mean of the distribution to remain constant. The interaction between these parameters is of such a nature that an increase in one cancels out the decrease in the other.

4.5.3 $w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$

Theorem 4.22. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$ *then*

$$\begin{aligned}
E(w(N;\phi)) &= \frac{e^{-\lambda}\lambda\Gamma(a+2)_1F_1(a+2;2;\lambda)}{a\Gamma(a+1)}. \\
f_w(n) &= \frac{\lambda^{n-1}\Gamma(1+a+n)}{n!\Gamma(a+2)\Gamma(n)_1F_1(a+2;2;\lambda)}. \\
g(z) &= \frac{z_1F_1(a+2;2;\lambda z)}{_1F_1(a+2;2;\lambda)}. \\
E(N^w) &= 1+\frac{(a+2)\lambda_1F_1(a+3;3;\lambda)}{2_1F_1(a+2;2;\lambda)}. \\
Var(N^w) &= \frac{(a+2)\lambda(6_1F_1(a+3;3;\lambda)+(a+3)\lambda_1F_1(a+4;4;\lambda))}{6_1F_1(a+2;2;\lambda)} + E(N^w) - (E(N^w))^2.
\end{aligned}$$

Restrictions:

- Domain: $n \in \mathbb{N}_1$.
- Parameters: $a > 0, \lambda > 0$.

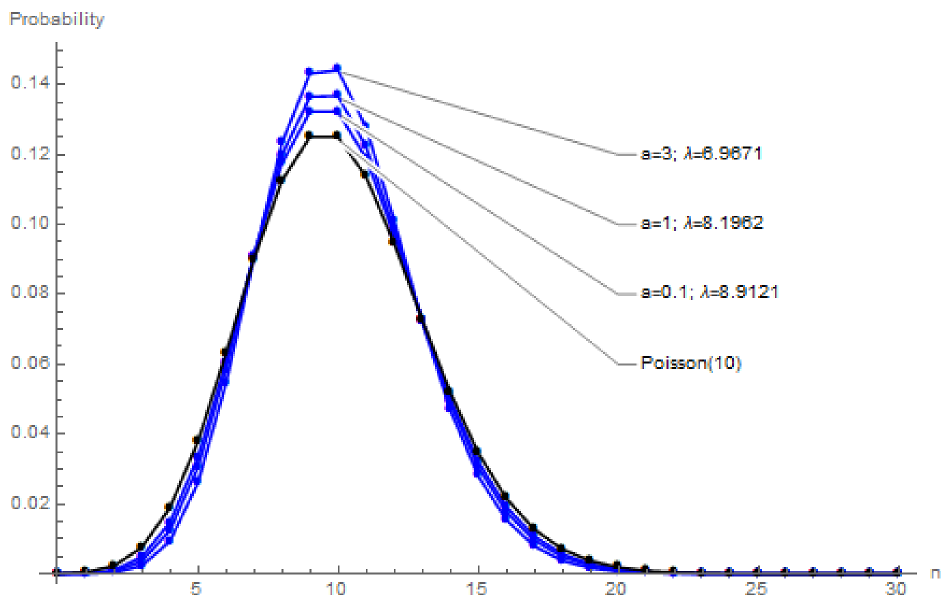


Figure 4.39: Probability mass function - $w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$, small a

The variances associated with the weighted Poisson probability mass functions in the above plots are 8.9238, 8.3539 and 7.539 for a equal to 0.1, 1 and 3 respectively.

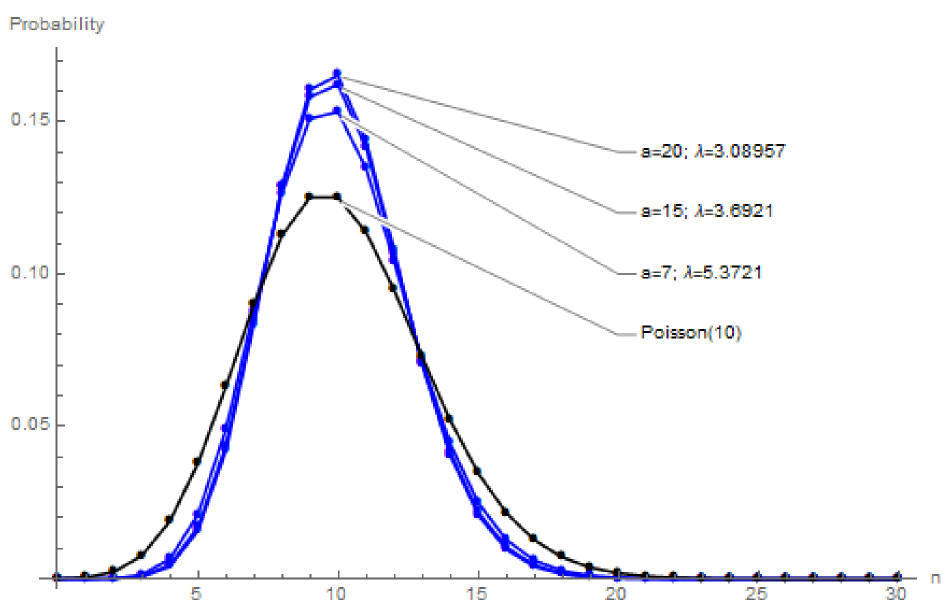


Figure 4.40: Probability mass function - $w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$, large a

The variances associated with the weighted Poisson probability mass functions in the above plots are 6.6978, 5.9939 and 5.7759 for a equal to 7, 15 and 20 respectively.

4.5.4 $w(n;\phi) = \left(\frac{-1}{\ln(1-p)}\frac{p^n}{n}\right)^{-1}$

Theorem 4.23. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \left(\frac{-1}{\ln(1-p)}\frac{p^n}{n}\right)^{-1}$ *then*

$$
\begin{aligned}
E(w(N;\phi)) &= -\frac{e^{\left(\frac{\lambda}{p}-\lambda\right)}\lambda\ln(1-p)}{p}. \\
f_w(n) &= \frac{e^{-\frac{\lambda}{p}}p^{1-n}\lambda^{n-1}}{(n-1)!}. \\
g(z) &= ze^{\frac{\lambda(z+p-1)}{p}-\lambda}. \\
E(N^w) &= \frac{p+\lambda}{p}. \\
Var(N^w) &= \frac{\lambda}{p}.
\end{aligned}
$$

Restrictions:

- Domain: $n \in \mathbb{N}_1$.
- Parameters: $0 < p < 1, \lambda > 0$.

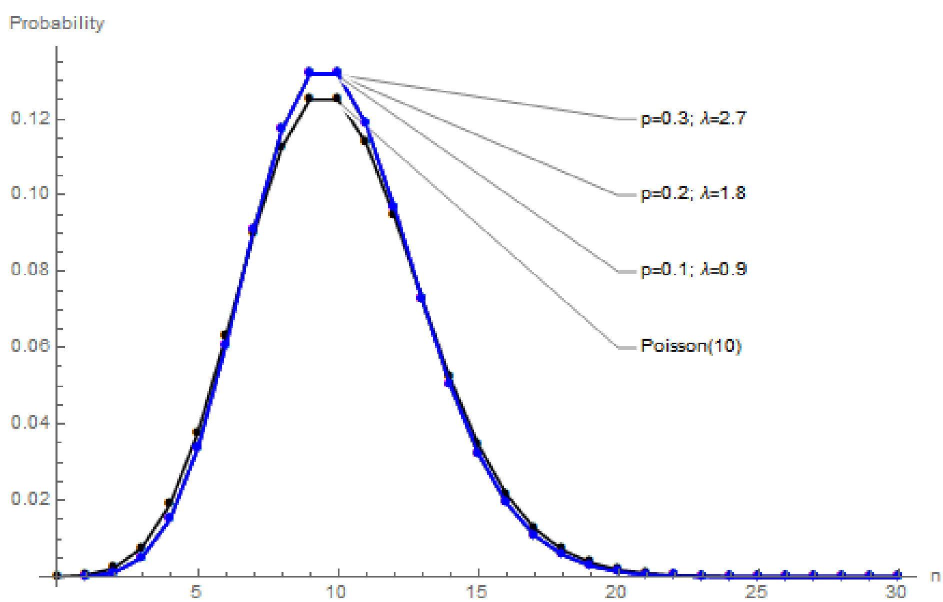


Figure 4.41: Probability mass function - $w(n;\phi) = \left(\frac{-1}{\ln(1-p)}\frac{p^n}{n}\right)^{-1}$

The variance associated with the weighted Poisson probability mass functions in the above plot is 9 for all values of p. This implies that varying the parameter p is entirely ineffective in affecting the shape of the probability mass function of this specific weighted Poisson distribution (when the expected value is fixed). The reason for this can be seen in the expressions of the equations as well as in the above plot: In all cases, when p decreases, λ is required to increase in order for the mean of the distribution to remain constant. The interaction between these parameters is of such a nature that an increase in one nullifies the decrease in the other.

4.5.5 $w(n;\phi) = \left(\frac{\Gamma(r+n)}{n!\Gamma(r)}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$

Theorem 4.24. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \left(\frac{\Gamma(r+n)}{n!\Gamma(r)}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$ *then*

$$
\begin{aligned}
E(w(N;\phi)) &= \frac{e^{-\lambda}Beta(a,b)\,{}_2F_2(1,a+b+r;b,r;\lambda)}{Beta(a+r,b)}. \\
f_w(n) &= \frac{\lambda^n Beta(a+r,b)\Gamma(r)}{Beta(a+r,b+n)\Gamma(r+n)\,{}_2F_2(1,a+b+r;b,r;\lambda)}. \\
g(z) &= \frac{{}_2F_2(1,a+b+r;b,r;\lambda z)}{{}_2F_2(1,a+b+r;b,r;\lambda)}. \\
E(N^w) &= \frac{(a+b+r)\lambda\,{}_2F_2(2,1+a+b+r;1+b,1+r;\lambda)}{br\,{}_2F_2(1,a+b+r;b,r;\lambda)}. \\
Var(N^w) &= \frac{2(a+b+r)(1+a+b+r)\lambda^2\,{}_2F_2(3,2+a+b+r;2+b,2+r;\lambda)}{b(1+b)r(1+r)\,{}_2F_2(1,a+b+r;b,r;\lambda)} + E(N^w) - (E(N^w))^2.
\end{aligned}
$$

Restrictions:

- Domain: $n \in \mathbb{N}_0$.
- Parameters: $a, b, r, \lambda > 0$.

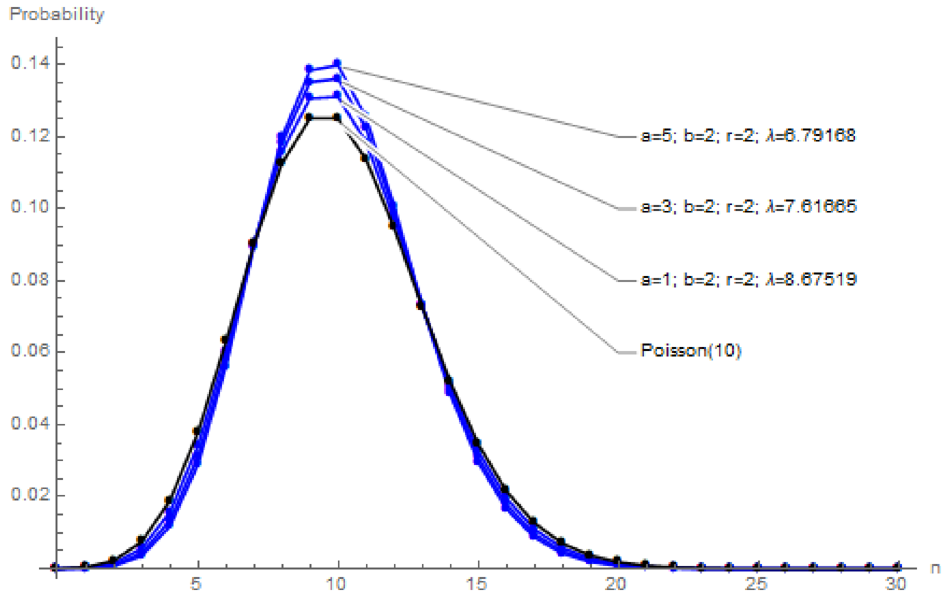


Figure 4.42: Probability mass function - $w(n;\phi) = \left(\frac{\Gamma(r+n)}{n!\Gamma(r)} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$, varying a

The variances associated with the weighted Poisson probability mass functions in the above plots are 9.1278, 8.4831 and 8.0419 for a equal to 1, 3 and 5 respectively.

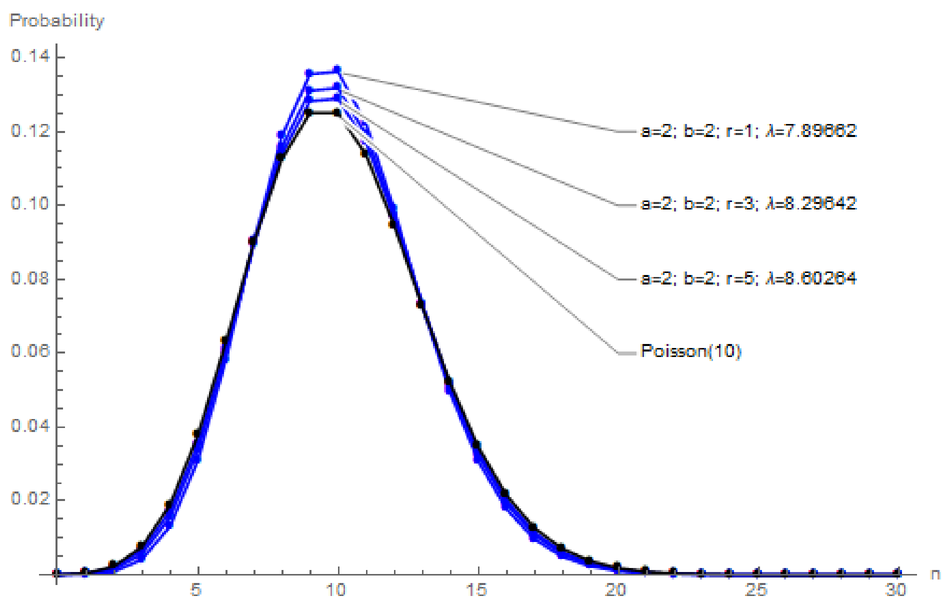


Figure 4.43: Probability mass function - $w(n;\phi) = \left(\frac{\Gamma(r+n)}{n!\Gamma(r)} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$, varying r

The variances associated with the weighted Poisson probability mass functions in the above plots are 8.4493, 9.0393 and 9.4505 for r equal to 1, 3 and 5 respectively.

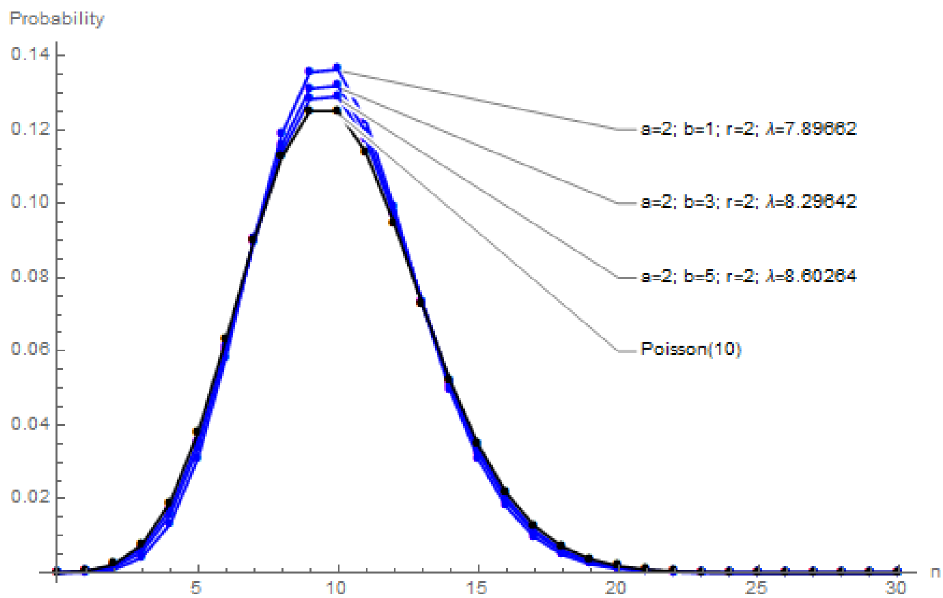


Figure 4.44: Probability mass function - $w(n;\phi) = \left(\frac{\Gamma(r+n)}{n!\Gamma(r)} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$, varying b

The variances associated with the weighted Poisson probability mass functions in the above plots are 8.4493, 9.0393 and 9.4505 for b equal to 1, 3 and 5 respectively. From the above graphs, it is apparent that changing r or b from some baseline results in the same change in the other properties of the weighted Poisson distribution.

4.5.6 $w(n;\phi) = \binom{n}{a}^{-1}$

Theorem 4.25. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \binom{n}{a}^{-1}$ *where* $a \in \mathbb{N}_0$ *then*

$$E\left(w\left(N;\phi\right)\right) = \frac{e^{-\lambda}\lambda^{a}{}_2F_2(1,1;1+a,1+a;\lambda)}{a!}.$$

$$f_w\left(n\right) = \frac{\lambda^{n-a}a!}{\begin{pmatrix} n \\ a \end{pmatrix} n!\,{}_2F_2(1,1;1+a,1+a;\lambda)}.$$

$$g\left(z\right) = \frac{z^{a}{}_2F_2(1,1;1+a,1+a;\lambda z)}{{}_2F_2(1,1;1+a,1+a;\lambda)}.$$

$$E\left(N^w\right) = a + \frac{\lambda{}_2F_2(2,2;2+a,2+a;\lambda)}{(1+a)^2{}_2F_2(1,1;1+a,1+a;\lambda)}.$$

$$\begin{aligned} Var\left(N^w\right) &= \frac{a(a-1)\left(a^2+3a+2\right)^2{}_2F_2(1,1;1+a,1+a;\lambda)}{(1+a)^2(2+a)^2{}_2F_2(1,1;1+a,1+a;\lambda)} \\ &+ \frac{2\lambda\left(a(2+a)^2{}_2F_2(2,2;2+a,2+a;\lambda)+2\lambda{}_2F_2(3,3;3+a,3+a;\lambda)\right)}{(1+a)^2(2+a)^2{}_2F_2(1,1;1+a,1+a;\lambda z)} + E\left(N^w\right) - \left(E\left(N^w\right)\right)^2. \end{aligned}$$

Restrictions:

- Domain:$n \in \{a, a+1, ...\}$.
- Parameters: $a \in \mathbb{N}_0, \lambda > 0$.

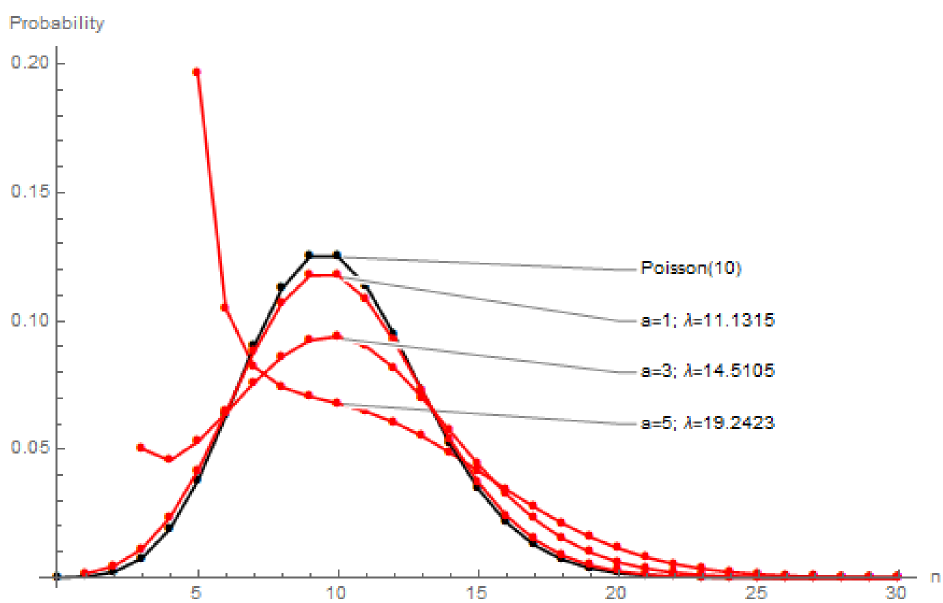


Figure 4.45: Probability mass function - $w\left(n;\phi\right) = \begin{pmatrix} n \\ a \end{pmatrix}^{-1}$

The variances associated with the weighted Poisson probability mass functions in the above plots are 11.3163, 16.5437 and 20.3559 for a equal to 1, 3 and 5 respectively.

4.5.7 $w(n;\phi) = ((a)_n)^{-1} = \frac{\Gamma(a)}{\Gamma(a+n)}$

Theorem 4.26. *If the weight function used in the weighted Poisson probability mass function is chosen as $w(n;\phi) = (a)_n$ then*

$$E(w(N;\phi)) = e^{-\lambda}{}_0F_1(a;\lambda).$$

$$f_w(n) = \frac{\lambda^n}{n!{}_0F_1(a;\lambda)(a)_n}.$$

$$g(z) = \frac{{}_0F_1(a;\lambda z)}{{}_0F_1(a;\lambda)}.$$

$$E(N^w) = \frac{\lambda^{1-\frac{a}{2}} I_a\left(2\sqrt{\lambda}\right)\Gamma(1+a)}{a{}_0F_1(a;\lambda)}.$$

$$Var(N^w) = \frac{\lambda^{\frac{3}{2}-\frac{a}{2}} I_{1+a}\left(2\sqrt{\lambda}\right)\Gamma(2+a)}{a(1+a){}_0F_1(a;\lambda)} + E(N^w) - (E(N^w))^2.$$

where $I_a(\cdot)$ is the modified Bessel function of the first kind (Definition 10.20)

Restrictions:

- Domain: $n \in \mathbb{N}_0$.
- Parameters: $a, \lambda > 0$.

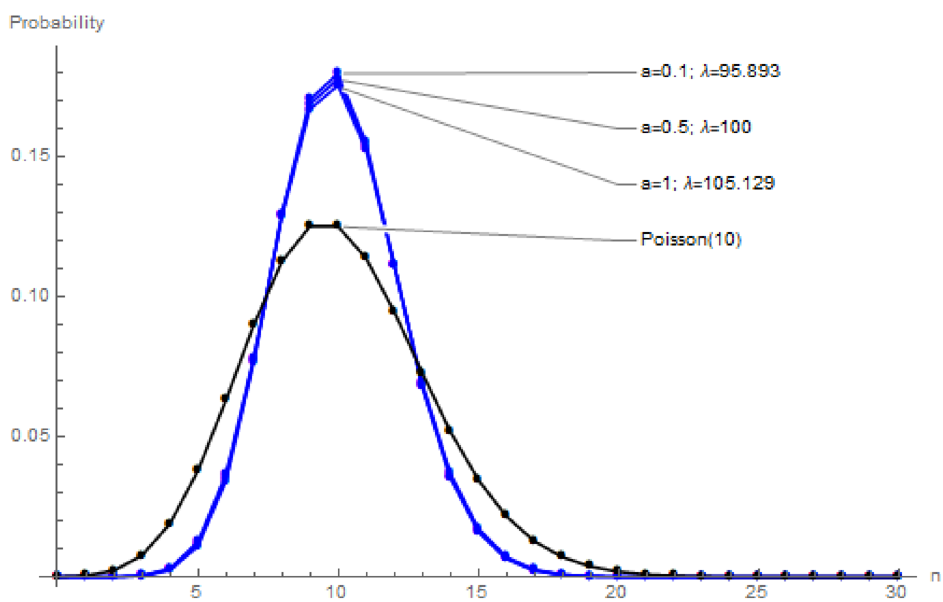


Figure 4.46: Probability mass function - $w(n;\phi) = ((a)_n)^{-1}$, small a

The variances associated with the weighted Poisson probability mass functions in the above plots are 4.8923, 5 and 5.1283 for a equal to 0.1, 0.5 and 1 respectively.

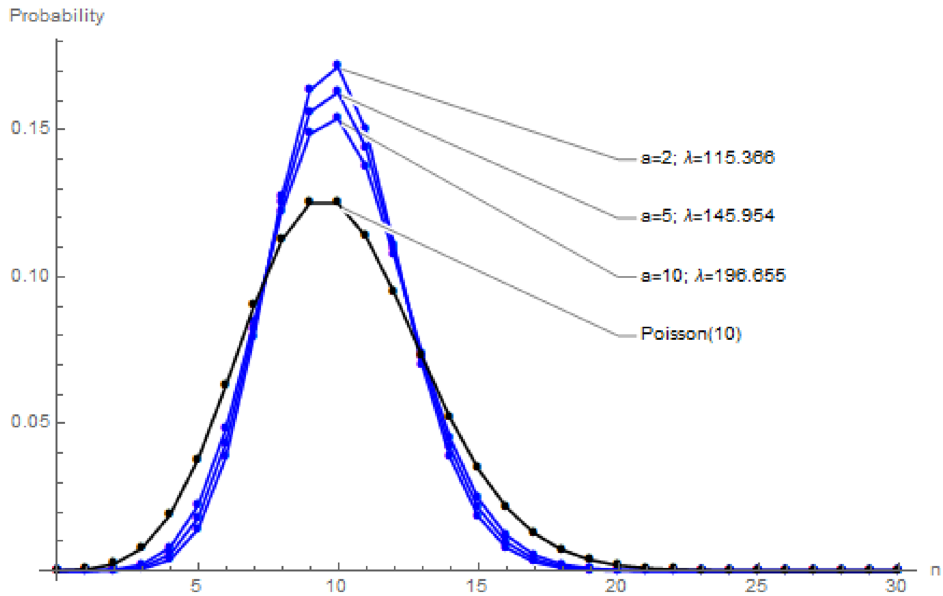


Figure 4.47: Probability mass function - $w(n;\phi) = \left((a)_n\right)^{-1}$, large a

The variances associated with the weighted Poisson probability mass functions in the above plots are 5.3656, 5.9537 and 6.6547 for a equal to 2, 5 and 10 respectively.

4.5.8 $w(n;\phi) = \left((n)_a\right)^{-1} = \frac{\Gamma(n)}{\Gamma(n+a)}$

Theorem 4.27. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = (a)_n$ *then*

$$E(w(N;\phi)) = \frac{e^{-\lambda}\lambda\,{}_2F_2(1,1;2,1+a,\lambda)}{(1)_a}.$$

$$f_w(n) = \frac{\lambda^{n-1}(1)_a}{n!\,{}_2F_2(1,1;2,1+a,\lambda)(n)_a}.$$

$$g(z) = \frac{z\,{}_2F_2(1,1;2,1+a,\lambda z)}{{}_2F_2(1,1;2,1+a,\lambda)}.$$

$$E(N^w) = \frac{e^{\lambda}\lambda^{-a}(\Gamma(1+a)-a\Gamma(a,\lambda))}{{}_2F_2(1,1;2,1+a,\lambda)}.$$

$$Var(N^w) = \frac{\lambda^{-a}(a\lambda^a+e^{\lambda}(a-\lambda))(a\Gamma(a,\lambda)-\Gamma(1+a))}{{}_2F_2(1,1;2,1+a,\lambda)} + E(N^w) - (E(N^w))^2.$$

Restrictions:

- Domain: $n \in \mathbb{N}_1$.
- Parameters: $a, \lambda > 0$.

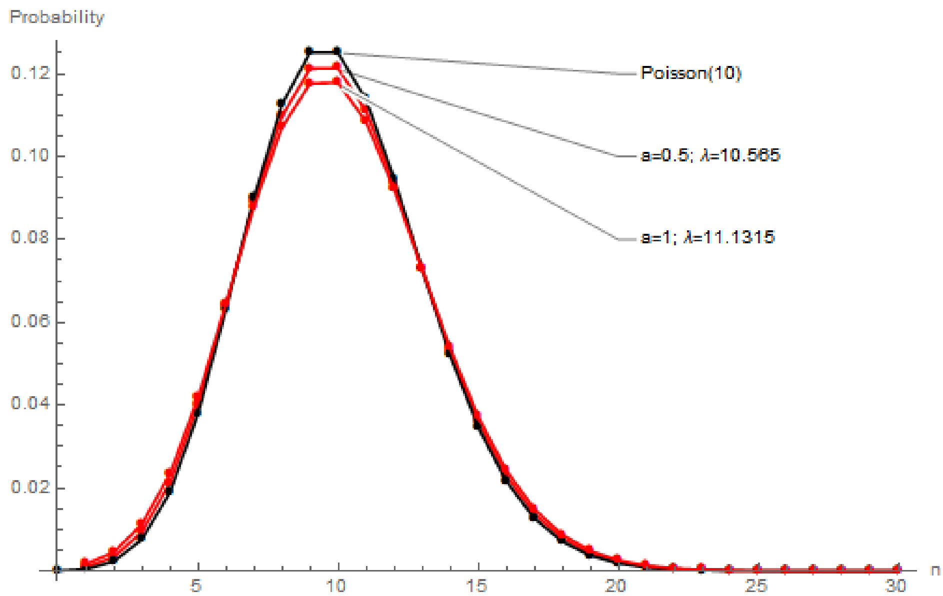


Figure 4.48: Probability mass function - $w(n;\phi) = ((a)_n)^{-1}$, small a

The variances associated with the weighted Poisson probability mass functions in the above plots are 10.6498 and 11.3163 for a equal to 0.5 and 1 respectively.

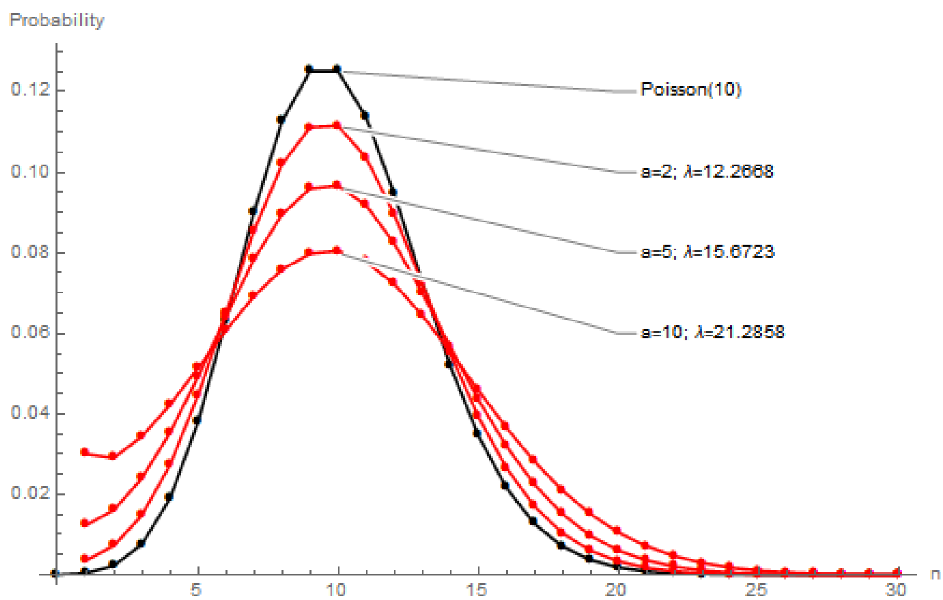


Figure 4.49: Probability mass function - $w(n;\phi) = ((a)_n)^{-1}$, large a

The variances associated with the weighted Poisson probability mass functions in the above plots are 12.6751, 16.784 and 24.1581 for a equal to 2, 5 and 10 respectively.

While some of the weight functions listed above are suitable dual partners for their non-inverted counterparts, some are not. In Table 4.3 below, the duality of the various weight functions is briefly summarised.

Table 4.3: Duality of weight functions

Weight function	If applicable, reason for non-duality
$(an^3 + bn^2 + cn)^{-1}$	While a closed-form expression exists (see Table 4.2), the original weight function results in a probability mass function that is defined on $n \in \mathbb{N}_0$. By contrast the inverse is defined on $n \in \mathbb{N}_1$.
$\left(a + \frac{b-ac}{n+c}\right)^{-1}$	This weight function results in a dual pair. However, the original function had the restriction that $c \in \mathbb{N}_0$ whereas the inverted counterpart only requires that $c > 0$.
$\left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	This weight function results in a dual pair. However, the original function had the restriction that $r \in \mathbb{N}_1$ whereas the inverted counterpart requires that $r \in \{2, 3, ...\}$.
$\left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	This weight function results in a dual pair.
$\left(\frac{-1}{\ln(1-p)} \frac{p^n}{n}\right)^{-1}$	This weight function results in a dual pair.
$\left(\frac{\Gamma(r+n)}{n!\Gamma(r)} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	This weight function results in a dual pair.
$\binom{n}{a}^{-1}$	While this weight function does not result in a dual pair as defined by Kokonendji et al. [83]. $(w_1(n) w_2(n) = 1 \forall n \in \mathbb{N}_0)$, it adheres to this property on the domain $n \in \{a, a+1, ...\}$.
$((a)_n)^{-1}$	This weight function results in a dual pair. However, the original function had the restriction that $0 < \lambda < 1$ whereas the inverted counterpart only requires that $0 < \lambda$.
$((n)_a)^{-1}$	While this weight function does not result in a dual pair as defined by Kokonendji et al. [83], it adheres to this property on the domain $n \in \mathbb{N}_1$.

Chapter 5

Weighted Poisson Distribution: Applications

In this chapter, the weighted Poisson distributions derived earlier in the thesis will be fitted to a variety of discrete datasets to demonstrate the flexibility and wide range of applications of the distributions. In Section 5.1 the numerical methods used to find the parameter estimates and their confidence intervals will be discussed. In Section 5.2 datasets that were obtained mostly from Kaggle.com will be modelled using the various weighted Poisson distributions. This is done to investigate how well the newly derived distributions perform. In Section 5.3, the weighted Poisson distributions will be applied to some datasets that were used in previously published papers that proposed alternate methods for modelling discrete data. This section is included to compare the performance of the newly derived distributions with those that were previously proposed in an attempt to compare the efficacy of the different methodologies.

5.1 Parameter estimation and confidence intervals

In this thesis parameter estimates will be calculated using a maximum likelihood approach. Since many of the weight functions result in complicated probability mass functions for the weighted Poisson distributions, obtaining explicit solutions for the parameter estimates is usually impossible. As a result, numerical methods have to be applied to find the estimates.

Many different methods exist which could be used to find these estimates. In general, the methods can be categorised as either being based on the gradient of the function being optimised, or as "direct search" methods. The gradient based approaches typically rely on the first or second derivatives of the function being optimised, and include the Newton-Raphson method which was first discussed by Newton [101] in 1669, the augmented Lagrangian method (originally known as the method of multipliers) which was discussed by Hestenes [68] and Powell [112], the interior point methods (Karmarkar [78]) and the sequential quadratic programming method (Nocedal and Wright [103]). In contrast, direct search methods do not rely on the derivatives of the function, and include the genetic algorithm (Holland [72] and Goldberg and Holland [57]), differential evolution (Storn [132] and Storn and Price [133]), simulated annealing (Kirkpatrick et al. [79]), and the Nelder–Mead (Nelder and Mead [100]) methods.

While an in-depth discussion about the various methods is beyond the scope of this thesis, in general, the direct search methods tend to converge more slowly, but are more tolerant to the presence of noise in the dataset relative to gradient based methods. Of the direct search methods mentioned, the Nelder–Mead is one of the most commonly used, with relatively fast and stable convergence properties. For this reason it is the algorithm that will be implemented to find the maximum likelihood estimates. The Nelder–Mead algorithm that is implemented can be described as follows:

- Let $l\left(\phi|n\right)$ be the log-likelihood function of a specific weighted Poisson distribution.
- Suppose that $l\left(\phi|n\right)$ consists of m variables that have to be estimated.
- A set of $m+1$ points $(x_1, x_2, ..., x_{m+1})$, will be selected which forms the vertices of a polytope in m dimensional space.
- For each iteration of the algorithm the log-likelihood function will be calculated at the $m+1$ points, and the points will be ordered such that $l\left(x_1|n\right) \geq l\left(x_2|n\right) \geq ... \geq l\left(x_{m+1}|n\right)$.
- Since the log-likelihood functions are ordered, it is known that $(x_1, x_2, ..., x_m)$ is the set of the best m points. The aim is to replace the worst point, x_{m+1}, with an improved one.
- This is achieved by first calculating the centre of the best m points $c = \frac{1}{m}\sum_{i=1}^{m} x_i$.
- A new "test point" x_t is generated by reflecting x_{m+1} though c. $x_t = c + \alpha\left(c - x_{m+1}\right)$, $(\alpha > 0)$.
- This test point can either be the best new point $l\left(x_t|n\right) \geq l\left(x_1|n\right)$, the worst new point $l\left(x_m|n\right) \geq l\left(x_t|n\right)$, or neither the best or the worst $l\left(x_1|n\right) \geq l\left(x_t|n\right) \geq l\left(x_m|n\right)$. The next step of the algorithm depends on which of the three cases is true:
- If x_t is the worst point, a new test point x_t^* is generated $x_t^* = \begin{cases} c + \gamma\left(x_{m+1} - c\right) & \text{,if } l\left(x_t|n\right) \leq l\left(x_{m+1}|n\right) \\ c + \gamma\left(x_t - c\right) & \text{,if } l\left(x_t|n\right) > l\left(x_{m+1}|n\right) \end{cases}$, $(0 < \gamma < 1)$ and x_t^* replaces x_{m+1} in $(x_1, x_2, ..., x_{m+1})$.
- If x_t is the best point, the reflection that resulted in x_t was very successful, and the reflection is further expanded. $x_t^* = c + \beta\left(x_t - r\right)$ where $\beta > 1$. If $l\left(x_t^*|n\right) \geq l\left(x_t|n\right)$ the further expansion is successful and x_t^* replaces x_{m+1} in the list $(x_1, x_2, ..., x_{m+1})$. Conversely, if the further expansion was not successful, $l\left(x_t|n\right) \geq l\left(x_t^*|n\right)$, and x_t replaces x_{m+1}.

- If x_t is neither the best or the worst, x_t replaces x_{m+1}.
- These steps are repeated until the best point, x_1, and best function, $l\left(x_1|n\right)$, do not change significantly between iterations.
- The maximum likelihood estimates are then set as the m parameters that make up x_1.

While many different values could have been chosen for α, β and γ in the Nelder–Mead algorithm, the selection in this chapter is based on values that have previously been used in this model's applications. Consequently $\alpha = 1, \beta = 2, \gamma = 0.5$ have been selected.

Once parameters have been estimated, the next question that has be addressed is how confidence intervals for the parameters will be constructed. Many software packages, specifically those designed with mathematical or statistical implementations in mind, have builtin confidence interval calculation functionalities. While these functions are extremely useful in general, the computational methods that they use to calculate the intervals can either result in errors, or in intervals that are narrower or wider than they should be when dealing with novel distributions. As a result, the intervals in this thesis will be coded manually, first using R and then with Mathematica to validate the results. While the confidence intervals will receive minimal comment, they have been included to give a sense of the accuracy of the point estimates, as well as to indicate whether certain special cases of the weight function could potentially be more suitable.

Three approaches are commonly used to construct confidence intervals for parameters. The first method, which is what most software packages provide, calculates asymptotic confidence intervals based on the Fisher information matrix. The second method, which was proposed by Efron and Tibshirani [44], is called the non-parametric bootstrap method, and the third is called the parametric bootstrap. While many different alterations and modifications have been proposed to the above methods, in general they can be described as follows.

5.1.1 The Fisher information matrix

The first confidence interval method based on the Fisher information matrix can be described as follows:

- Let $l\left(\phi|n\right)$ be the log-likelihood function of a specific distribution.
- Let $\phi = \left[\phi_1, ..., \phi_m\right]^T$ be a vector of the parameters of the distribution, and $\hat{\phi} = \left[\hat{\phi}_1, ..., \hat{\phi}_m\right]^T$ be a vector of maximum likelihood estimates of the parameters of the distribution.
- The Fisher information matrix is defined as the negative matrix consisting of the partial

second derivatives of the log likelihood function:

$$I\left(\phi\right)=-\begin{pmatrix}\frac{\partial^2 l(\phi|n)}{\partial\phi_1^2} & \frac{\partial^2 l(\phi|n)}{\partial\phi_1\partial\phi_2} & \cdots & \frac{\partial^2 l(\phi|n)}{\partial\phi_1\partial\phi_m}\\ \frac{\partial^2 l(\phi|n)}{\partial\phi_2\partial\phi_1} & \frac{\partial^2 l(\phi|n)}{\partial\phi_2^2} & \cdots & \frac{\partial^2 l(\phi|n)}{\partial\phi_2\partial\phi_m}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial^2 l(\phi|n)}{\partial\phi_m\partial\phi_1} & \frac{\partial^2 l(\phi|n)}{\partial\phi_m\partial\phi_2} & \cdots & \frac{\partial^2 l(\phi|n)}{\partial\phi_m^2}\end{pmatrix}.$$

- The observed Fisher information matrix is merely the Fisher information matrix evaluated the maximum likelihood estimates:

$$I\left(\hat{\phi}\right)=-\begin{pmatrix}\frac{\partial^2 l\left(\hat{\phi}|n\right)}{\partial\phi_1^2} & \frac{\partial^2 l\left(\hat{\phi}|n\right)}{\partial\phi_1\partial\phi_2} & \cdots & \frac{\partial^2 l\left(\hat{\phi}|n\right)}{\partial\phi_1\partial\phi_m}\\ \frac{\partial^2 l\left(\hat{\phi}|n\right)}{\partial\phi_2\partial\phi_1} & \frac{\partial^2 l\left(\hat{\phi}|n\right)}{\partial\phi_2^2} & \cdots & \frac{\partial^2 l\left(\hat{\phi}|n\right)}{\partial\phi_2\partial\phi_m}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial^2 l\left(\hat{\phi}|n\right)}{\partial\phi_m\partial\phi_1} & \frac{\partial^2 l\left(\hat{\phi}|n\right)}{\partial\phi_m\partial\phi_2} & \cdots & \frac{\partial^2 l\left(\hat{\phi}|n\right)}{\partial\phi_m^2}\end{pmatrix}.$$

- The inverse of the observed Fisher information matrix, $\left[I\left(\hat{\phi}\right)\right]^{-1}$, is an estimator of the asymptotic covariance matrix (if certain consistency requirements are met). (See Efron and Hinkley [43] for a discussion on using the observed versus expected Fisher information matrix in these estimates). Thus, using the central limit theorem and the law of large numbers it follows that $\hat{\phi}\sim N\left(\phi,\left[I\left(\hat{\phi}\right)\right]^{-1}\right)$.

- As a result, the asymptotic confidence intervals for the parameters can be constructed as follows:

$$\hat{\phi}_i\pm z_{1-\frac{\alpha}{2}}\left[\left[I\left(\hat{\phi}\right)\right]^{-1}\right]^{0.5}_{(i,i)},i=1,2,...,m.$$

For an m-parameter, discrete distribution, a set of sufficient requirements for the maximum likelihood estimates to be consistent (and thus for the above confidence interval method to be implemented) is:

1. The observations $X_1,X_2,...,X_n$ are n i.i.d. random variables with probability density function $f_\phi\left(x\right)=P_\phi\left(X_i=x\right)$.

2. The distributions, $f_\phi\left(x\right)$, are distinct. In other words if $f_{\phi^1}\left(x\right)=f_{\phi^2}\left(x\right)$ it implies that the parameters are the same, $\phi^1=\phi^2$.

3. The parameter space, Φ, is open.

4. The set A on which $f_\phi\left(x\right)$ is positive is independent of ϕ.

5. For all x in A the third partial derivatives $\frac{\partial^3}{\partial\phi_i\partial\phi_j\partial\phi_k} f_\phi(x)$ for $i, j, k \in \{1, 2, ..., m\}$ exist and are continuous, and the third partial derivatives of $\sum_{i=1}^{n} f_\phi(x_i)$ can be obtained by differentiating under the summation sign.

6. If $\phi^0 = (\phi_1^0, \phi_2^0, ..., \phi_k^0)$ denotes the true value of ϕ, there exists a function $M_{ijk}(x)$ and a positive number $c(\phi^0)$ such that

$$\left|\frac{\partial^3}{\partial\phi_i\partial\phi_j\partial\phi_k} ln\left(f_\phi(x)\right)\right| \leq M_{ijk}(x),$$

for all ϕ with $\sum_{i=1}^{k} (\phi_i - \phi_i^0)^2 < c(\phi^0)$, where $E_{\phi^0}(M_{ijk}(x)) < \infty$ for all i, j, k.

7. The information matrix $I(\phi)$ is positive definite and has finite elements.

Note that the requirements for a single parameter distribution are very similar and can be found in Lehmann [89] p469.

Showing that a specific distribution meets the aforementioned consistency requirements can range from trivial to impossible, depending of the makeup and complexity of the probability mass function under consideration. A few examples of weighted Poisson distributions from Chapter 4 that fall on this spectrum are discussed.

It can immediately be established that some of the weighted Poisson distributions do not meet the above sufficient requirements. Namely, all of the distributions that have parameters that affects their respective domains ($w(n;\phi) = \binom{m}{n} p^n (1-p)^{m-n}$, $w(n;\phi) = \frac{ab^a}{n^{a+1}}$, $w(n;\phi) = I(n \geq a)$, $w(n;\phi) = I(n \leq b)$, $w(n;\phi) = I(n \geq a) I(n \leq b)$, $w(n;\phi) = \binom{n}{a}$, $w(n;\phi) = \binom{n}{a}^{-1}$) clearly violate requirement 4 above. Similarly, all of the distributions that have parameters that are defined exclusively on $\mathbb{N}_0$ or $\mathbb{N}_1$ ($w(n;\phi) = n^{-a}$, $w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$, $w(n;\phi) = \binom{m}{n} p^n (1-p)^{m-n}$, $w(n;\phi) = \frac{ab^a}{n^{a+1}}$, $w(n;\phi) = \binom{n}{a}$, $w(n;\phi) = \binom{n}{a}^{-1}$, $w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$) violate requirement 3 since the natural numbers is not an open set.

It can also be demonstrated that some simple weights result in distributions that meet the sufficient set of requirements for consistency specified above. For example when $w(n;\phi) = n$.

Requirements 1 though 4 are clearly met by the definition of the probability mass function.

The third derivative of the probability mass function is:

$$\frac{d^3}{d\lambda^3} \frac{e^{-\lambda}\lambda^{n-1}}{(n-1)!} = \frac{e^{-\lambda}\lambda^{n-4}\left(n^3 - 3n^2(2+\lambda) + n(11+3\lambda(3+\lambda)) - \lambda(6+\lambda(3+\lambda)) - 6\right)}{(n-1)!}.$$

The order of summation and differentiation can be swapped around

$$
\begin{aligned}
& \sum_{n=0}^{\infty} \frac{\partial^3}{\partial\lambda^3} \frac{e^{-\lambda}\lambda^{n-1}}{(n-1)!} \\
= & \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^{n-4}\left(n^3-3n^2(2+\lambda)+n(11+3\lambda(3+\lambda))-\lambda(6+\lambda(3+\lambda))-6\right)}{(n-1)!} \\
= & \frac{\lambda^3+6\lambda^2+7\lambda+1}{\lambda^3} - \frac{3\left(\lambda^4+5\lambda^3+7\lambda^2+2\lambda\right)}{\lambda^4} + \frac{3\lambda^4+12\lambda^3+20\lambda^2+11\lambda}{\lambda^4} - \frac{\lambda^4+3\lambda^3+6\lambda^2+6\lambda}{\lambda^4} \\
= & 0.
\end{aligned}
$$

Thus the 5^{th} requirement has been met.

The log of the probability mass function is given by

$$ln\left(\frac{e^{-\lambda}\lambda^{n-1}}{(n-1)!}\right) = -\lambda + (n-1)\,ln\,(\lambda) - ln\,((n-1)!)\,.$$

The third derivative of the log of the probability mass function is:

$$\frac{d^3}{d\lambda^3}\left(-\lambda + (n-1)\,ln\,(\lambda) - ln\,((n-1)!)\right) = \frac{2(n-1)}{\lambda^3}.$$

If λ^0 denotes the true value of λ and $c\left(\lambda^0\right) = \frac{\lambda^0}{2}$ then

$$
\begin{aligned}
& \max_{\frac{1}{2}\lambda^0<\lambda<\frac{3}{2}\lambda^0} \left|\frac{d^3}{d\lambda^3}\left(-\lambda + (n-1)\,ln\,(\lambda) - ln\,((n-1)!)\right)\right| \\
= & \max_{\frac{1}{2}\lambda^0<\lambda<\frac{3}{2}\lambda^0} \frac{2(n-1)}{\lambda^3} \\
< & \frac{2(n-1)}{\left(\frac{1}{2}\lambda^0\right)^3} \\
= & \frac{16(n-1)}{\left(\lambda^0\right)^3}.
\end{aligned}
$$

It then follows that

$$
\begin{aligned}
& E_{\lambda^0}\left(M_{\lambda^0}(N)\right) \\
= & \sum_{n=1}^{\infty} \frac{16(n-1)}{\left(\lambda^0\right)^3} \frac{e^{-\lambda}\left(\lambda^0\right)^{n-1}}{(n-1)!} \\
= & \frac{16e^{-\lambda_0}}{\left(\lambda^0\right)^3} \sum_{n=1}^{\infty} \frac{(n-1)\left(\lambda^0\right)^{n-1}}{(n-1)!} \\
= & \frac{16e^{-\lambda_0}}{\left(\lambda^0\right)^2} \sum_{n=1}^{\infty} \frac{\left(\lambda^0\right)^{n-2}}{(n-2)!} \\
= & \frac{16e^{-\lambda_0}}{\left(\lambda^0\right)^2} e^{\lambda_0} \\
= & \frac{16}{\left(\lambda^0\right)^2}.
\end{aligned}
$$

Thus the 6^{th} requirement has been met. (Note that the 7^{th} requirement is not applicable for a single parameter distribution.)

As a slightly more complicated example, suppose that $w\,(n;\phi) = n + \varepsilon$.

Requirements 1, 3 and 4 are clearly met by the definition of the probability mass function. However, demonstrating requirement 2 (that the distribution is identifiable) is not trivial.

The probability mass function of the distribution is given by $f_w(n) = \frac{(n+\varepsilon)e^{-\lambda}\lambda^n}{(\lambda+\varepsilon)n!}$. It follows that $f_w(0) = \frac{\varepsilon e^{-\lambda}}{(\lambda+\varepsilon)}$, $f_w(1) = \frac{(1+\varepsilon)e^{-\lambda}\lambda}{(\lambda+\varepsilon)}$ and $f_w(2) = \frac{(2+\varepsilon)e^{-\lambda}\lambda^2}{(\lambda+\varepsilon)2}$. Using these three equalities it follows that $\frac{2f_w(2)-\lambda f_w(1)}{f_w(1)-\lambda f_w(0)} = \lambda$, and consequently that

$$f_w(0)\,\lambda^2 - 2f_w(1)\,\lambda + 2f_w(2) = 0.$$

Using this quadratic equation it is possible to get solutions for λ.

$$\lambda = \frac{f_w(1) \pm \sqrt{(f_w(1))^2 - 2f_w(0)\,f_w(2)}}{f_w(0)}. \tag{5.1}$$

Since $(f_w(1))^2 - 2f_w(0)\,f_w(2) = \frac{e^{-2\lambda}\lambda^2}{(\lambda+\varepsilon)^2} > 0$, Equation (5.1) constitutes a pair of real solutions for λ. Since $f_w(1) - \sqrt{(f_w(1))^2 - 2f_w(0)\,f_w(2)} > 0$, both of the solutions are positive. Solving for ε gives the following expression:

$$\varepsilon = \lambda\left(\frac{e^{-\lambda}}{f_w(1) - \lambda f_w(0)} - 1\right) > 0. \tag{5.2}$$

Since λ has two solutions, ε has two solutions as well. In order to demonstrate identifiability, it needs to be shown that a function (which is monotone in λ) exists, that if λ is given, results in a single, unique value for ε.

From Equation (5.2) it follows that if λ is known, ε will be known as well. Consider $f_w(0) = \frac{\varepsilon e^{-\lambda}}{(\lambda+\varepsilon)}$. By taking its derivative with respect to λ it follows that

$$\begin{aligned}
&\frac{\partial}{\partial\lambda}\frac{\varepsilon e^{-\lambda}}{(\lambda+\varepsilon)} \\
=\;& \frac{\partial}{\partial\lambda}\frac{\lambda\left(\frac{e^{-\lambda}}{f_w(1)-\lambda f_w(0)}-1\right)e^{-\lambda}}{\lambda+\lambda\left(\frac{e^{-\lambda}}{f_w(1)-\lambda f_w(0)}-1\right)e^{-\lambda}} \\
=\;& \frac{-\lambda e^{-\lambda}\left(\frac{e^{-\lambda}}{f_w(1)-\lambda f_w(0)}-1\right)\left(1+\lambda+\lambda\left(\frac{e^{-\lambda}}{f_w(1)-\lambda f_w(0)}-1\right)\right)}{\left(\lambda+\lambda\left(\frac{e^{-\lambda}}{f_w(1)-\lambda f_w(0)}\right)\right)^2}.
\end{aligned}$$

From Equation (5.2) it follows that $\frac{\partial}{\partial\lambda}f_w(0) < 0$. As a result, it has been proven that $f_w(0), f_w(1)$ and $f_w(2)$ uniquely determine λ and ε and thus that 2^{nd} consistency requirement has been met.

The third partial derivatives of the distribution are:

$$
\begin{aligned}
& \frac{\partial^3}{\partial\lambda^3}\frac{(n+\varepsilon)e^{-\lambda}\lambda^n}{(\lambda+\varepsilon)n!} \\
= \ & \frac{(n+\varepsilon)e^{-\lambda}\lambda^{n-3}\big(3\lambda\varepsilon^2(n-\lambda-1)\big(\lambda^2-2n\lambda+n(n-3)\big)\big)}{(\lambda+\varepsilon)^4 n!} \\
- \ & \frac{(n+\varepsilon)e^{-\lambda}\lambda^{n-3}\big(3\lambda^2\varepsilon\big(2\lambda-(n-\lambda-2)\big(n(n-3)-2n\lambda+\lambda^2\big)\big)\big)}{(\lambda+\varepsilon)^4 n!} \\
+ \ & \frac{(n+\varepsilon)e^{-\lambda}\lambda^{n-3}\big(\varepsilon^3\big(n^3-\lambda^3-3n^2(1+\lambda)+n(2+3\lambda(1+\lambda))\big)\big)}{(\lambda+\varepsilon)^4 n!} \\
+ \ & \frac{(n+\varepsilon)e^{-\lambda}\lambda^{n-3}\big(\lambda^3\big(n^3-3n^2(2+\lambda)-6-\lambda(6+\lambda(3+\lambda))+n(11+3\lambda(3+\lambda))\big)\big)}{(\lambda+\varepsilon)^4 n!},
\end{aligned}
$$

$$
\begin{aligned}
& \frac{\partial^3}{\partial\varepsilon^3}\frac{(n+\varepsilon)e^{-\lambda}\lambda^n}{(\lambda+\varepsilon)n!} \\
= \ & \frac{6e^{-\lambda}\lambda^n(n-\lambda)}{(\lambda+\varepsilon)^4 n!},
\end{aligned}
$$

$$
\begin{aligned}
& \frac{\partial^3}{\partial\lambda^2\partial\varepsilon}\frac{(n+\varepsilon)e^{-\lambda}\lambda^n}{(\lambda+\varepsilon)n!} \\
= \ & \frac{e^{-\lambda}\lambda^{n-2}\big(\lambda^2\big(2\lambda-(n-\lambda-2)\big(\lambda^2-2n\lambda+n(n-3)\big)\big)\big)}{(\lambda+\varepsilon)^4 n!} \\
- \ & \frac{e^{-\lambda}\lambda^{n-2}\big(\varepsilon^2(n-\lambda)\big(\lambda^2-2(n+1)\lambda+n(n-1)\big)\big)}{(\lambda+\varepsilon)^4 n!} \\
- \ & \frac{e^{-\lambda}\lambda^{n-2}\big(2\varepsilon\lambda\big(2\lambda+(n-\lambda)\big(\lambda^2-2n\lambda+n(n-3)\big)\big)\big)}{(\lambda+\varepsilon)^4 n!},
\end{aligned}
$$

and

$$
\begin{aligned}
& \frac{\partial^3}{\partial\lambda\partial\varepsilon^2}\frac{(n+\varepsilon)e^{-\lambda}\lambda^n}{(\lambda+\varepsilon)n!} \\
= \ & \frac{2e^{-\lambda}\lambda^{n-1}\Big(\varepsilon\Big((n-\lambda)^2-\lambda\Big)\Big)}{(\lambda+\varepsilon)^4 n!} \\
+ \ & \frac{2e^{-\lambda}\lambda^{n-1}\big(\lambda\big(n^2+\lambda(\lambda+2)-n(3+2\lambda)\big)\big)}{(\lambda+\varepsilon)^4 n!},
\end{aligned}
$$

respectively.

The order of summation and differentiation of the probability mass function can also be swapped around. This is demonstrated for the four cases discussed above.

$$
\begin{aligned}
& \sum_{n=0}^{\infty}\frac{\partial^3}{\partial\lambda^3}\frac{(n+\varepsilon)e^{-\lambda}\lambda^n}{(\lambda+\varepsilon)n!} \\
= \ & \sum_{n=0}^{\infty}\frac{(n+\varepsilon)e^{-\lambda}\lambda^{n-3}\big(3\lambda\varepsilon^2(n-\lambda-1)\big(\lambda^2-2n\lambda+n(n-3)\big)\big)}{(\lambda+\varepsilon)^4 n!} \\
- \ & \sum_{n=0}^{\infty}\frac{(n+\varepsilon)e^{-\lambda}\lambda^{n-3}\big(3\lambda^2\varepsilon\big(2\lambda-(n-\lambda-2)\big(n(n-3)-2n\lambda+\lambda^2\big)\big)\big)}{(\lambda+\varepsilon)^4 n!} \\
+ \ & \sum_{n=0}^{\infty}\frac{(n+\varepsilon)e^{-\lambda}\lambda^{n-3}\big(\varepsilon^3\big(n^3-\lambda^3-3n^2(1+\lambda)+n(2+3\lambda(1+\lambda))\big)\big)}{(\lambda+\varepsilon)^4 n!} \\
+ \ & \sum_{n=0}^{\infty}\frac{(n+\varepsilon)e^{-\lambda}\lambda^{n-3}\big(\lambda^3\big(n^3-3n^2(2+\lambda)-6-\lambda(6+\lambda(3+\lambda))+n(11+3\lambda(3+\lambda))\big)\big)}{(\lambda+\varepsilon)^4 n!} \\
= \ & 0+\frac{6\varepsilon}{(\lambda+\varepsilon)^4}+0-\frac{6\varepsilon}{(\lambda+\varepsilon)^4} \\
= \ & 0.
\end{aligned}
$$

$$
\begin{aligned}
& \sum_{n=0}^{\infty}\frac{\partial^3}{\partial\varepsilon^3}\frac{(n+\varepsilon)e^{-\lambda}\lambda^n}{(\lambda+\varepsilon)n!} \\
= \ & \sum_{n=0}^{\infty}\frac{6e^{-\lambda}\lambda^n(n-\lambda)}{(\lambda+\varepsilon)^4 n!} \\
= \ & 0.
\end{aligned}
$$

$$\begin{aligned}
& \sum_{n=0}^{\infty} \frac{\partial^3}{\partial\lambda^2\partial\varepsilon} \frac{(n+\varepsilon)e^{-\lambda}\lambda^n}{(\lambda+\varepsilon)n!} \\
= & \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^{n-2}\left(\lambda^2\left(2\lambda-(n-\lambda-2)\left(\lambda^2-2n\lambda+n(n-3)\right)\right)\right)}{(\lambda+\varepsilon)^4 n!} \\
- & \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^{n-2}\left(\varepsilon^2(n-\lambda)\left(\lambda^2-2(n+1)\lambda+n(n-1)\right)\right)}{(\lambda+\varepsilon)^4 n!} \\
- & \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^{n-2}\left(2\varepsilon\lambda\left(2\lambda+(n-\lambda)\left(\lambda^2-2n\lambda+n(n-3)\right)\right)\right)}{(\lambda+\varepsilon)^4 n!} \\
= & \ 0-0-0 \\
= & \ 0.
\end{aligned}$$

$$\begin{aligned}
& \sum_{n=0}^{\infty} \frac{\partial^3}{\partial\lambda\partial\varepsilon^2} \frac{(n+\varepsilon)e^{-\lambda}\lambda^n}{(\lambda+\varepsilon)n!} \\
= & \sum_{n=0}^{\infty} \frac{2e^{-\lambda}\lambda^{n-1}\left(\varepsilon\left((n-\lambda)^2-\lambda\right)\right)}{(\lambda+\varepsilon)^4 n!} \\
+ & \sum_{n=0}^{\infty} \frac{2e^{-\lambda}\lambda^{n-1}\left(\lambda\left(n^2+\lambda(\lambda+2)-n(3+2\lambda)\right)\right)}{(\lambda+\varepsilon)^4 n!} \\
= & \ 0+0 \\
= & \ 0.
\end{aligned}$$

Thus this distribution meets the 5^{th} requirement.

The the log of the probability mass function is given by

$$\begin{aligned}
& ln\left(\frac{(n+\varepsilon)e^{-\lambda}\lambda^n}{(\lambda+\varepsilon)n!}\right) \\
= & -\lambda+ln\left(n+\varepsilon\right)+nln\left(\lambda\right)-ln\left(\lambda+\varepsilon\right)-ln\left(n!\right).
\end{aligned}$$

The third partial derivatives of the log of the probability mass function are:

$$\begin{aligned}
& \frac{\partial^3}{\partial\lambda^3}\left(-\lambda+ln\left(n+\varepsilon\right)+nln\left(\lambda\right)-ln\left(\lambda+\varepsilon\right)-ln\left(n!\right)\right) \\
= & \frac{2n}{\lambda^3}-\frac{2}{(\lambda+\varepsilon)^3},
\end{aligned}$$

$$\begin{aligned}
& \frac{\partial^3}{\partial\varepsilon^3}\left(-\lambda+ln\left(n+\varepsilon\right)+nln\left(\lambda\right)-ln\left(\lambda+\varepsilon\right)-ln\left(n!\right)\right) \\
= & \frac{2}{(n+\varepsilon)^3}-\frac{2}{(\lambda+\varepsilon)^3},
\end{aligned}$$

$$\begin{aligned}
& \frac{\partial^3}{\partial\lambda^2\partial\varepsilon}\left(-\lambda+ln\left(n+\varepsilon\right)+nln\left(\lambda\right)-ln\left(\lambda+\varepsilon\right)-ln\left(n!\right)\right) \\
= & -\frac{2}{(\lambda+\varepsilon)^3},
\end{aligned}$$

and

$$\begin{aligned}
& \frac{\partial^3}{\partial\lambda\partial\varepsilon^2}\left(-\lambda+ln\left(n+\varepsilon\right)+nln\left(\lambda\right)-ln\left(\lambda+\varepsilon\right)-ln\left(n!\right)\right) \\
= & -\frac{2}{(\lambda+\varepsilon)^3},
\end{aligned}$$

respectively.

Let λ^0 denote the true value of λ and ε^0 the true value of ε. If $c\left(\lambda^0,\varepsilon^0\right)=\left(\frac{1}{2}min\left(\lambda^0,\varepsilon^0\right)\right)^2$, then $\lambda\in\left(\frac{1}{2}\lambda^0,\frac{3}{2}\lambda^0\right)$ and $\varepsilon\in\left(\frac{1}{2}\varepsilon^0,\frac{3}{2}\varepsilon^0\right)$.

It then follows that

$$\begin{aligned}
& \left|\frac{\partial^3}{\partial\lambda^3}\left(-\lambda + ln\left(n+\varepsilon\right) + nln\left(\lambda\right) - ln\left(\lambda+\varepsilon\right) - ln\left(n!\right)\right)\right| \\
= & \left|\frac{2n}{\lambda^3} - \frac{2}{(\lambda+\varepsilon)^3}\right| \\
\leq & \frac{2n}{\lambda^3} + \frac{2}{(\lambda+\varepsilon)^3}.
\end{aligned}$$

This function is decreasing in both λ and ε. Thus it holds that

$$\begin{aligned}
& \frac{2n}{\lambda^3} + \frac{2}{(\lambda+\varepsilon)^3} \\
\leq & \frac{2n}{\left(\frac{1}{2}\lambda^0\right)^3} + \frac{2}{\left(\frac{1}{2}\lambda^0+\frac{1}{2}\varepsilon^0\right)^3} \\
= & 16\left(\frac{n}{(\lambda^0)^3} + \frac{1}{(\lambda^0+\varepsilon^0)^3}\right).
\end{aligned}$$

It then follows that

$$\begin{aligned}
& E\left(16\left(\frac{n}{(\lambda^0)^3} + \frac{1}{(\lambda^0+\varepsilon^0)^3}\right)\right) \\
= & 16\left(\frac{1}{(\lambda^0+\varepsilon^0)^3} + \frac{\lambda^0\left(1+\lambda^0+\varepsilon^0\right)}{(\lambda^0)^3(\lambda^0+\varepsilon^0)}\right) \\
< & \infty.
\end{aligned}$$

Similarly

$$\begin{aligned}
& \left|\frac{\partial^3}{\partial\lambda^2\partial\varepsilon}\left(-\lambda + ln\left(n+\varepsilon\right) + nln\left(\lambda\right) - ln\left(\lambda+\varepsilon\right) - ln\left(n!\right)\right)\right| \\
= & \frac{2}{(\lambda+\varepsilon)^3}.
\end{aligned}$$

This function is again decreasing in both λ and ε. Thus it holds that

$$\begin{aligned}
& \frac{2}{(\lambda+\varepsilon)^3} \\
\leq & \frac{2}{\left(\frac{1}{2}\lambda^0+\frac{1}{2}\varepsilon^0\right)^3} \\
= & \frac{16}{(\lambda^0+\varepsilon^0)^3},
\end{aligned}$$

and

$$\begin{aligned}
& E\left(\frac{16}{(\lambda^0+\varepsilon^0)^3}\right) \\
= & 16\left(\frac{1}{(\lambda^0+\varepsilon^0)^3} + \frac{\lambda^0\left(1+\lambda^0+\varepsilon^0\right)}{(\lambda^0)^3(\lambda^0+\varepsilon^0)}\right) \\
< & \infty.
\end{aligned}$$

Lastly

$$\begin{aligned}
& \left|\frac{\partial^3}{\partial\varepsilon^3}\left(-\lambda + ln\left(n+\varepsilon\right) + nln\left(\lambda\right) - ln\left(\lambda+\varepsilon\right) - ln\left(n!\right)\right)\right| \\
= & \left|\frac{2}{(n+\varepsilon)^3} - \frac{2}{(\lambda+\varepsilon)^3}\right| \\
\leq & \frac{2}{(n+\varepsilon)^3} + \frac{2}{(\lambda+\varepsilon)^3}.
\end{aligned}$$

This function is again a decreasing in both λ and ε. Thus it holds that

$$
\begin{aligned}
& \tfrac{2}{(n+\varepsilon)^3} + \tfrac{2}{(\lambda+\varepsilon)^3} \\
\leq \;& \tfrac{2}{\left(n+\frac{1}{2}\varepsilon^0\right)^3} + \tfrac{2}{\left(\frac{1}{2}\lambda^0+\frac{1}{2}\varepsilon^0\right)^3} \\
= \;& \tfrac{2}{\left(n+\frac{1}{2}\varepsilon^0\right)^3} + \tfrac{16}{(\lambda^0+\varepsilon^0)^3},
\end{aligned}
$$

and

$$
\begin{aligned}
& E\left(\tfrac{2}{\left(n+\frac{1}{2}\varepsilon^0\right)^3} + \tfrac{16}{(\lambda^0+\varepsilon^0)^3}\right) \\
= \;& \tfrac{16}{(\lambda^0+\varepsilon^0)^3} + E\left(\tfrac{2}{\left(n+\frac{1}{2}\varepsilon^0\right)^3}\right) \\
\leq \;& \tfrac{16}{(\lambda^0+\varepsilon^0)^3} + E\left(\tfrac{2}{\left(\frac{1}{2}\varepsilon^0\right)^3}\right).
\end{aligned}
$$

since $n \geq 0$. Thus

$$
\begin{aligned}
& E\left(\tfrac{2}{\left(n+\frac{1}{2}\varepsilon^0\right)^3} + \tfrac{16}{(\lambda^0+\varepsilon^0)^3}\right) \\
\leq \;& \tfrac{16}{(\lambda^0+\varepsilon^0)^3} + \tfrac{16}{\left(\frac{1}{2}\varepsilon^0\right)^3} \\
< \;& \infty.
\end{aligned}
$$

Consequently the 6^{th} requirement is met.

Proving the 7^{th} requirement is more complicated to demonstrate since it requires that all eigenvalues of the information matrix be positive, and while an expression for the eigen values exists, without observations from an actual dataset it is difficult to determine if this is indeed the case. The expression is given below. Let I denote the Fisher information matrix;

$$
\begin{aligned}
I(\phi) &= -\begin{pmatrix} \frac{\partial^2 l(\phi|n)}{\partial\lambda^2} & \frac{\partial^2 l(\phi|n)}{\partial\lambda\partial\varepsilon} \\ \frac{\partial^2 l(\phi|n)}{\partial\varepsilon\partial\lambda} & \frac{\partial^2 l(\phi|n)}{\partial\varepsilon^2} \end{pmatrix} \\
&= \begin{pmatrix} \frac{(\lambda^2+\varepsilon^2+2\lambda\varepsilon)\sum_{i=1}^n x_i - \lambda^2 n}{\lambda^2(\varepsilon+\lambda)^2} & -\frac{n}{(\varepsilon+\lambda)^2} \\ -\frac{n}{(\varepsilon+\lambda)^2} & \sum_{i=1}^n \frac{1}{(\varepsilon+x_i)^2} - \frac{n}{(\varepsilon+\lambda)^2} \end{pmatrix}.
\end{aligned}
$$

Then

$$
\begin{aligned}
& det\left(I(\phi) - E * I\right) \\
= \;& E^2 - \left(\tfrac{(\lambda^2+\varepsilon^2+2\lambda\varepsilon)\sum_{i=1}^n x_i - \lambda^2 n}{\lambda^2(\varepsilon+\lambda)^2} + \sum_{i=1}^n \tfrac{1}{(\varepsilon+x_i)^2} - \tfrac{n}{(\varepsilon+\lambda)^2}\right) E \\
+ \;& \tfrac{(\lambda^2+\varepsilon^2+2\lambda\varepsilon)\sum_{i=1}^n x_i - \lambda^2 n}{\lambda^2(\varepsilon+\lambda)^2}\left(\sum_{i=1}^n \tfrac{1}{(\varepsilon+x_i)^2} - \tfrac{n}{(\varepsilon+\lambda)^2}\right) - \left(\tfrac{n}{(\varepsilon+\lambda)^2}\right)^2 \\
\\
= \;& E^2 - \left(\tfrac{\sum_{i=1}^n x_i}{\lambda^2 \sum_{i=1}^n (\varepsilon+x_i)^2}\right) E + \tfrac{(\lambda^2+\varepsilon^2+2\lambda\varepsilon)\sum_{i=1}^n x_i - \lambda^2 n}{\lambda^2(\varepsilon+\lambda)^2} \sum_{i=1}^n \tfrac{1}{(\varepsilon+x_i)^2} \\
- \;& \tfrac{(\lambda^2+\varepsilon^2+2\lambda\varepsilon)\sum_{i=1}^n x_i - \lambda^2 n}{\lambda^2(\varepsilon+\lambda)^2} \tfrac{n}{(\varepsilon+\lambda)^2} - \left(\tfrac{n}{(\varepsilon+\lambda)^2}\right)^2 \\
\\
= \;& E^2 - \tfrac{\sum_{i=1}^n x_i}{\lambda^2 \sum_{i=1}^n (\varepsilon+x_i)^2} E \\
+ \;& \tfrac{\sum_{i=1}^n x_i}{\lambda^2 \sum_{i=1}^n (\varepsilon+x_i)^2} - \tfrac{n}{(\varepsilon+\lambda)^2 \sum_{i=1}^n (\varepsilon+x_i)^2} - \tfrac{n \sum_{i=1}^n x_i}{\lambda^2(\varepsilon+\lambda)^2} \\
\\
= \;& E^2 - \tfrac{\sum_{i=1}^n x_i}{\lambda^2 \sum_{i=1}^n (\varepsilon+x_i)^2} E + \tfrac{(\varepsilon+\lambda)^2 \sum_{i=1}^n x_i - n\lambda^2 - n\sum_{i=1}^n (\varepsilon+x_i)^2 \sum_{i=1}^n x_i}{\lambda^2(\varepsilon+\lambda)^2 \sum_{i=1}^n (\varepsilon+x_i)^2}
\end{aligned}
$$

$$E = \frac{\frac{\sum_{i=1}^{n} x_i}{\lambda^2 \sum_{i=1}^{n} (\varepsilon+x_i)^2} \pm \sqrt{\left(\frac{\sum_{i=1}^{n} x_i}{\lambda^2 \sum_{i=1}^{n} (\varepsilon+x_i)^2}\right)^2 - 4\frac{(\varepsilon+\lambda)^2 \sum_{i=1}^{n} x_i - n\lambda^2 - n\sum_{i=1}^{n} (\varepsilon+x_i)^2 \sum_{i=1}^{n} x_i}{\lambda^2 (\varepsilon+\lambda)^2 \sum_{i=1}^{n} (\varepsilon+x_i)^2}}}{2}$$

If it is assumed that $\left(\frac{\sum_{i=1}^{n} x_i}{\lambda^2 \sum_{i=1}^{n} (\varepsilon+x_i)^2}\right)^2 - 4\frac{(\varepsilon+\lambda)^2 \sum_{i=1}^{n} x_i - n\lambda^2 - n\sum_{i=1}^{n} (\varepsilon+x_i)^2 \sum_{i=1}^{n} x_i}{\lambda^2 (\varepsilon+\lambda)^2 \sum_{i=1}^{n} (\varepsilon+x_i)^2} > 0$ and $\left(\frac{\sum_{i=1}^{n} x_i}{\lambda^2 \sum_{i=1}^{n} (\varepsilon+x_i)^2}\right) - \sqrt{\left(\frac{\sum_{i=1}^{n} x_i}{\lambda^2 \sum_{i=1}^{n} (\varepsilon+x_i)^2}\right)^2 - 4\frac{(\varepsilon+\lambda)^2 \sum_{i=1}^{n} x_i - n\lambda^2 - n\sum_{i=1}^{n} (\varepsilon+x_i)^2 \sum_{i=1}^{n} x_i}{\lambda^2 (\varepsilon+\lambda)^2 \sum_{i=1}^{n} (\varepsilon+x_i)^2}} > 0$ the information matrix $I(\phi)$ will be positive definite and the requirements to show that a distribution has consistent maximum likelihood estimates will be met.

There are also some distributions that may adhere to the above regularity conditions, but that have expressions which are so complicated that these results may be infeasible or impossible. For example if $w(n;\phi) = (n+a)(n-b)^2$, then $\frac{\partial^3}{\partial\phi_i \partial\phi_j \partial\phi_k} ln(f_\phi(x))$ results in expressions that consist of 4 or more terms, each of which are longer than half a page. In this situation finding expressions for $M_{ijk}(x)$ is unlikely. Similarly, any of the probability mass functions that contain hypergeometric functions result in partial derivatives that contain digamma functions. In these situations it is impossible to show that the order of differentiation and summation can be swapped.

5.1.2 Non-parametric bootstrap

The second confidence interval method (non-parametric bootstrap approach) can be described as follows:

- Assume that A is a sample consisting of n observations, $A = \{X_1, X_2, ..., X_n\}$.
- Using these n observations the maximum likelihood estimates of the model, $\hat{\phi} = \left[\hat{\phi}_1, ..., \hat{\phi}_m\right]$, can be calculated (as discussed earlier in this chapter).
- It is possible to calculate confidence intervals for $\hat{\phi}$ without making any assumptions regarding the distribution of $\hat{\phi}$; if one assumes that the sample is an accurate representation of the population.
- To do this, create B bootstrap samples, each of size n, by sampling, with replacement, from A.
- For each bootstrap sample calculate the "bootstrap parameter estimates", $\hat{\phi^*}_i, i = 1, 2, ..., B$.
- The empirical distribution of these B estimates represents the uncertainty about the true value of $\hat{\phi}$.
- The confidence interval of ϕ can be obtained by using the quantiles of $\hat{\phi^*}$.
- Confidence interval: $\left[\hat{\phi^*}_{\frac{\alpha}{2}}, \hat{\phi^*}_{1-\frac{\alpha}{2}}\right]$.

5.1.3 Parametric bootstrap

The third confidence interval method (parametric bootstrap approach) can be described as follows:

- Assume that A is a sample consisting of n observations, $A = \{X_1, X_2, ..., X_n\}$.
- Using these n observations the maximum likelihood estimates of the model, $\hat{\phi} = \left[\hat{\phi}_1, ..., \hat{\phi}_m\right]$, can be calculated (as discussed earlier in this chapter).
- The maximum likelihood probability mass function is then given by $f\left(x|\hat{\phi}\right)$.
- Generate B bootstrap samples of size m, $B_i = \{X_{i1}, X_{i2}, ..., X_{im}\}$, by generating pseudo-random data points from $f\left(x|\hat{\phi}\right)$.
- For each bootstrap sample calculate the bootstrap parameter estimates, $\hat{\phi^*}_i, i = 1, 2, ..., B$, as well as the bootstrap parameter difference, $\hat{\delta}^*_i = \hat{\phi}^*_i - \hat{\phi}$.
- The confidence interval of ϕ can be obtained by using the quantiles of $\hat{\delta}^*$.
- Confidence interval: $\left[\hat{\phi} - \hat{\delta^*}_{\frac{\alpha}{2}}, \hat{\phi} - \hat{\delta^*}_{1-\frac{\alpha}{2}}\right]$.

These three methods for constructing confidence intervals all have different pros and cons which have been discussed at length by various authors. However, they can be briefly summarised as follows:

The Fisher information approach gives asymptotic confidence intervals. This means that, in addition to the restrictions discussed earlier in this section, this method is unsuitable for relatively small sample sizes (what constitutes “small” varies from distribution to distribution), since the method relies on the law of large numbers and the central limit theorem. This method also relies on finding the inverse of the Fisher information matrix. In some cases, when two (or more) parameters are highly correlated, calculating these inverses may result in matrices that are numerically unstable, although some authors have proposed methods to mitigate the impact of this (see Gill and King [56]). The final concern that is applicable to this research, is that some derivatives of the log-likelihood function (with respect to the parameters) can become computationally impractical to calculate for large sample sizes (although they exist theoretically).

The non-parametric bootstrap method assumes that the sample is an accurate representation of the population. Consequently, if the sample is small, the non parametric bootstrap samples may underestimate the amount of variation in the population. Additionally, while this method is less prone to experiencing the “computationally impractical calculations” of the Fisher information method, it may still be a computationally intensive method for calculating the confidence intervals.

In contrast, the parametric bootstrap method usually provides more accurate confidence intervals than the non-parametric bootstrap, but it does so while assuming an inherently arbitrary choice of model. This selected model may not necessarily be a good choice to fit to the data.

In Section 5.2, 95% confidence intervals will be provided for the parameter estimates of the weighted Poisson distribution that gives the best fit to the data. All three estimation methods discussed will be implemented (when the restrictions mentioned above do not prohibit their calculation or implementation). For the non-parametric bootstrap, the number of bootstrap samples will be $B = 1000$. Similarly for the parametric bootstrap the number of bootstrap samples will be $B = 1000$ and the bootstrap sample size will be $m = 500$ (when computational intensity allows).

5.2 Novel data fits

The datasets that are used in this section were almost all obtained from www.Kaggle.com. They range in the number of observations from 52 to close to 7 million. Some datasets only contained observations from the discrete variables that will be modelled, while others also had dozens of related explanatory variables.

The purpose of this chapter is not to provide the best possible regression fit for the data, but rather to demonstrate how well the weighted Poisson distribution can model observed discrete data, even when the data do not adhere to the restrictions of the Poisson distribution. As a result, it is almost certain that better model fits could be obtained by using any number of different regression modelling methodologies that take into account the various explanatory variables in the datasets (since using these methods would decrease the unexplained heterogeneity in the response variable). However, as was stated in the introductory chapter, one of the aims of the weighted Poisson distribution is actually to be able to model latent heterogeneity in data. It is for this reason that the datasets are also minimally cleaned. Only observations that had missing or nonsensical values were removed; outliers and other data points that would usually be removed before modelling were kept in the datasets in an attempt to keep the data as noisy as possible. If the outliers were to be removed, the weighted Poisson fits could be substantially better.

It is important to note that the datasets and graphs which are in this thesis have in no way been cherry-picked to shed a favourable light on the weighted Poisson distribution. All of the datasets which were analysed during the writing of this thesis are presented, including those for which the weighted Poisson distributions do not provide impressive measures of fit.

For all of the different datasets that are investigated in this section, at least one graph is provided. In the graphs, the blue dots with the vertical filling represent the empirical probability mass function. The black dots represent the fitted Poisson probability mass function, and the red dots represent the fitted weighted Poisson probability mass function. The dots of the Poisson and weighted Poisson probability mass functions are again joined by lines for clarity/aesthetic reasons. Estimation was done using maximum likelihood as previously described. Parameter estimates were obtained using R, and were verified with Mathematica. To increase the likelihood that estimates converge to the same stable values, the starting values of the optimisation functions were varied, and the programs run multiple times.

It should be noted that some of the fits achieved in this chapter are either near-perfect, or result in two models being visually indistinguishable from each other. In these situations, it might appear as if only a single plot is present on a graph. This, however, is not the case: the two plots are merely so similar that the one overlaps the other.

The process for the fitting of each dataset is as follows:

- The minimum observed value of the dataset is determined. The minimum is crucial since many of the weighted Poisson distributions derived in Chapter 4 are only applicable when only non-zero observations are present. Thus if zeros are observed, only weighted Poisson distributions that are defined on the domain $n \in \mathbb{N}_0$ can be implemented. However, if datasets have a minimum observed value larger than zero, weighted Poisson distributions that are defined on $n \in \mathbb{N}_0$ are still viable. (Transforming the non-compliant data is a viable practical alternative, but was not used in this thesis, in an attempt to demonstrate the flexibility of the weighted Poisson distribution.)

- The maximum likelihood estimate for the Poisson distribution's parameter is calculated as well as the values for the maximum log-likelihood, Akaike information criterion (AIC), small sample Akaike information criterion (AICc), and the Bayesian information criterion (BIC) of the fit.

- For all of the viable weighted Poisson distributions, the maximum likelihood estimates for the parameters are calculated as well as values for the maximum log-likelihood, AIC, AICc and BIC of the fit. (Tables containing these results can be found in Chapter 10.)

- Of all of the fitted models, the one that has the smallest AIC, AICc and BIC values is plotted against the Poisson distribution in the graphs. (While the AIC and AICc can give a different pronouncement on the best model in comparison to the BIC this occurred extremely rarely.)

- The different sets of confidence intervals are calculated for the best weighted Poisson model using the methods described earlier in the chapter.

After all of the datasets have been analysed and presented, a general discussion will be given, which highlights some of the findings and shortcomings of the various weighted Poisson distributions.

The first dataset contains information about mass shootings in the USA between 2013 and 2017, and originated from "Shooting Tracker". While there are many versions of this data available online (an even larger dataset, which spans from 1966 to 2017 can be found at `https://www.kaggle.com/zusmani/us-mass-shootings-last-50-years`) the source of our data is `https://www.shootingtracker.com`.

Shooting Tracker uses the FBI definition of a mass shooting which is "four or more shot and/or killed in a single event [incident], at the same general time and location, not including the shooter". This means that the total number of victims per incident will be at least 4. For this reason it is expected that one of the truncating weighted Poisson distributions will be a suitable model.

For the first variable, the number of victims per incident, of the 27 weight functions that were tested, 12 models perform better than the Poisson, and 1 additional model may (the AIC and AICc disagree with the BIC). The best fit is achieved when $w(n;\phi) = \frac{9\times 4^9}{n^{10}}$ and $\lambda = 19.5315$. This is shown in Figure 5.1. The weighted Poisson distribution gives a substantially better fit than the Poisson. For demonstration purposes, a second graph is also included where the extreme outliers have been removed from the data. This is shown in Figure 5.2.

The confidence intervals for the weighted Poisson parameters are as follows:

	Non-parametric bootstrap	Parametric bootstrap
λ	$(5.4933; 36.6402)$	$(5.71414; 19.5361)$
a	$(2; 14)$	$(1; 9)$

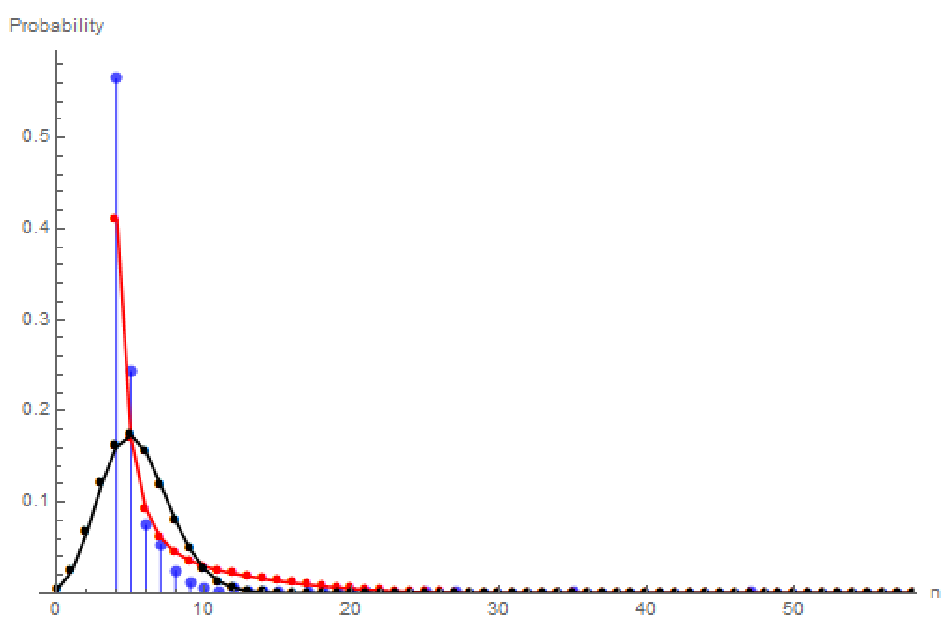


Figure 5.1: Mass shootings - Victims per incident

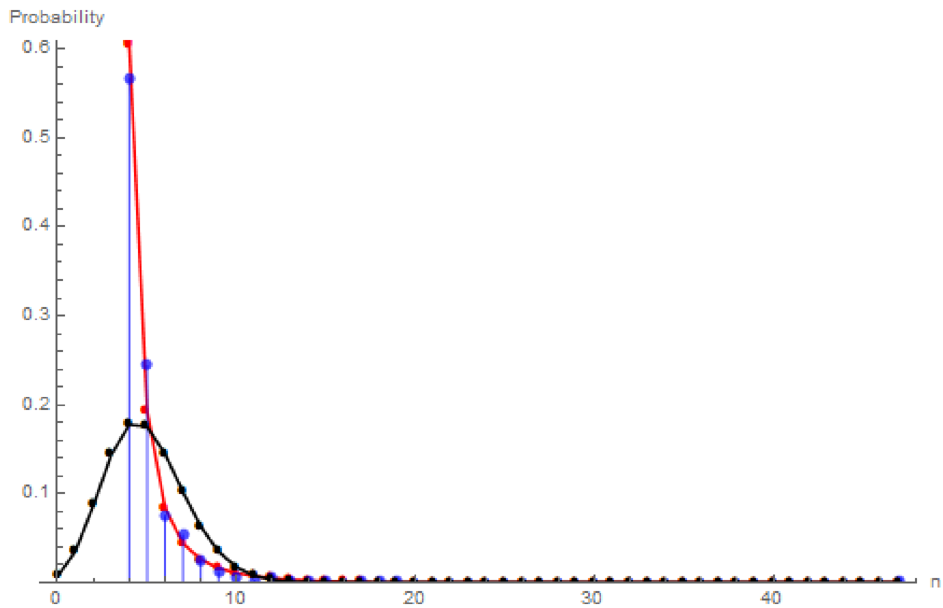


Figure 5.2: Mass shootings - Victims per incident excluding outliers

For the second variable, the number of incidents per day, of the 14 weight functions that were tested, 7 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = (2.08179)_n$ and $\lambda = 0.294394$. This is shown in Figure 5.3. In this case, the weighted Poisson distribution clearly gives a better fit to the data. In fact, the fit is near-perfect.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	$(0.239908; 0.34888)$	$(0.237945; 0.343552)$	$(0.182261; 0.386682)$
a	$(1.54953; 2.61404)$	$(1.677704; 2.74481)$	$(1.3965; 3.79694)$

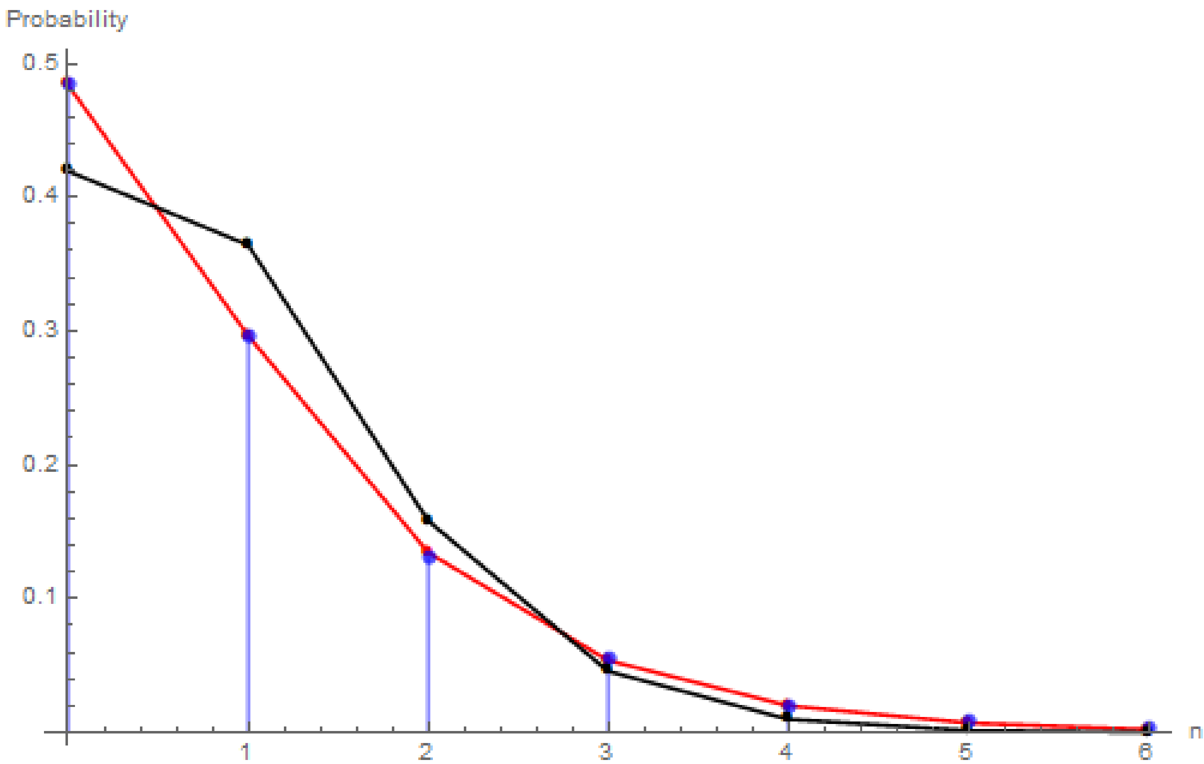


Figure 5.3: Mass shootings - Incidents per day

The next dataset contains information about vehicle accidents in Great Britain in 2015. This dataset can be found at `https://www.kaggle.com/silicon99/dft-accident-data`. The original dataset spanned from 2005 to 2015.

For the first variable, the number of incidents per hour, of the 13 weight functions that were tested, 8 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = \frac{\Gamma(0.887543+n)}{n!\Gamma(0.887543)} \frac{Beta(1352.2475,0.877599+n)}{Beta(1351.36,0.877599)}$ and $\lambda = 1373.38$. This is shown in Figure 5.4. In this case, the weighted Poisson distribution gives a good fit to the data.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	$(1242.59; 1504.17)$	$(834.854; 4714.48)$	$(850.5; 5058.23)$
a	$(1246.66; 1486.06)$	$(803.479; 4887.09)$	$(817.737; 5193.52)$
b	$(0.691629; 1.06357)$	$(0.2776; 1.83454)$	$(0.193936; 2.1056)$
r	$(0.69158; 1.06351)$	$(0.344266; 1.94401)$	$(0.22022; 2.23273)$

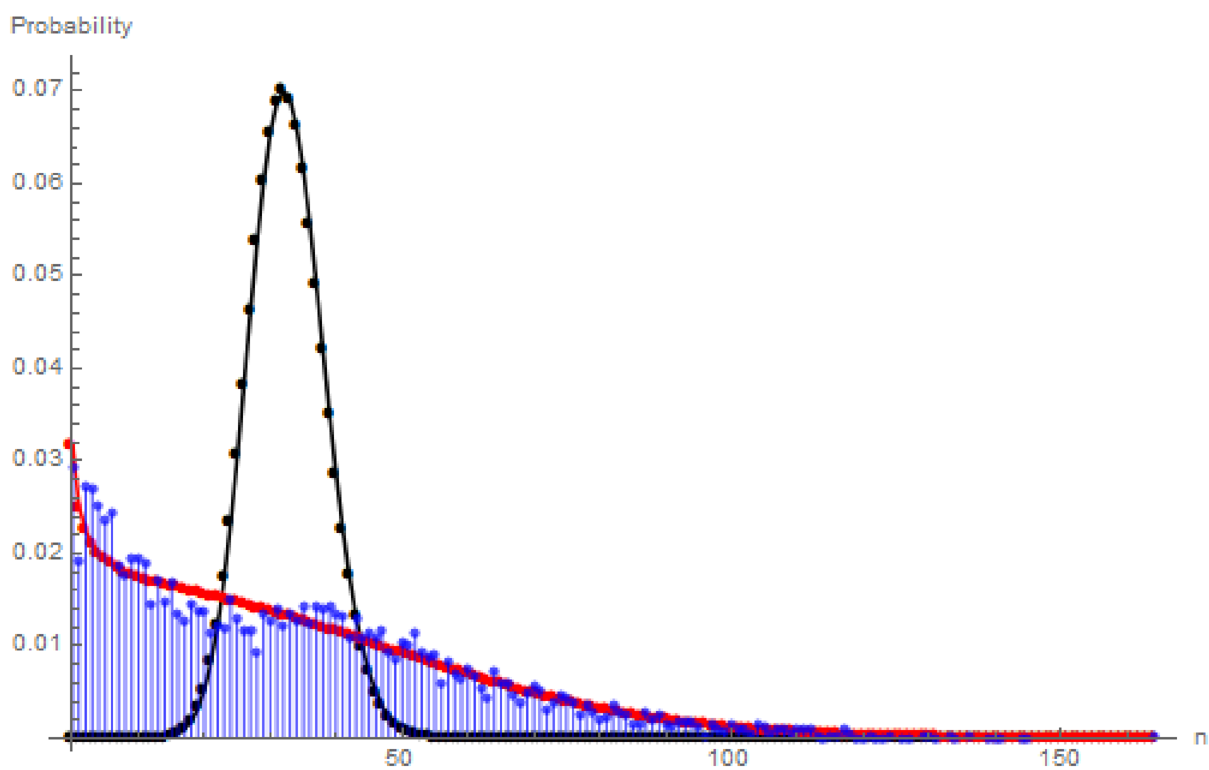


Figure 5.4: Britain accidents - Incidents per hour

For the second variable fit, the number of vehicles per incident, of the 26 weight functions that were tested, 19 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = (n + 0.00000001)(n - 0.655111)^2$ and $\lambda = 0.229596$. This is shown in Figure 5.5. In this case, the weighted Poisson distribution gives a very good fit to the data.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	$(0.228948; 0.230245)$	$(0.164424; 0.364196)$	$(0.137628; 0.330561)$
a	$(0; 0.00000001)$	$(0.0000001; 0.00000013)$	$(0.0000001; 0.00000015)$
b	$(0.653885; 0.656338)$	$(0.37588; 0.74394)$	$(0.425134; 0.810132)$

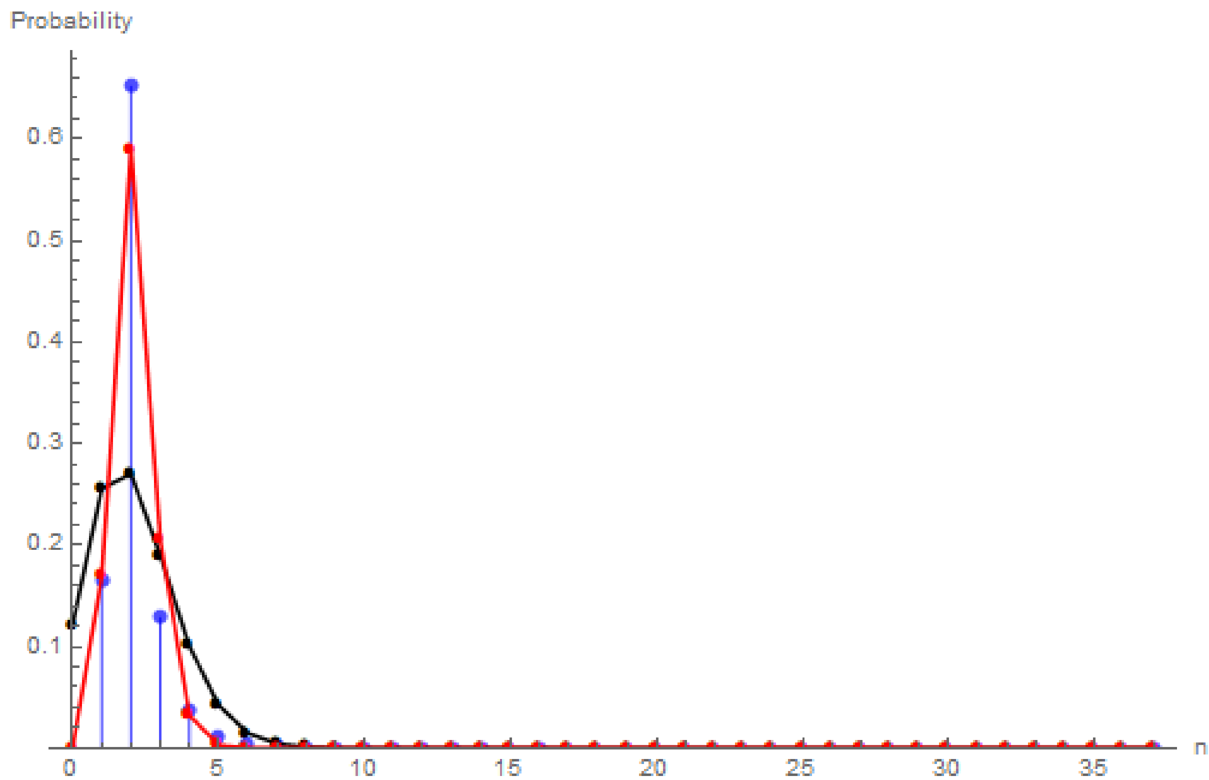


Figure 5.5: Britain accidents - Vehicles per incident

The next dataset contains information about gun violence in the USA from 1 January 2014 until 31 December 2017. This dataset can be found at `https://www.kaggle.com/jameslko/gun-violence-data`.

The first variable, the number of injuries per incident, of the 12 weight functions that were tested, 8 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = \frac{\Gamma(3.07126+n)}{n!\Gamma(3.07126)} \frac{Beta(126.39226,3.1556+n)}{Beta(123.321,3.1556)}$ and $\lambda = 6.93968$. This is shown in Figure 5.6. In this case, both the Poisson and weighted Poisson distributions give very good fits to the data. The AIC of the best weighted Poisson distribution is only 0.38% smaller than that of the Poisson distribution.

The confidence intervals for the weighted Poisson parameters are as follows:

	Non-parametric bootstrap	Parametric bootstrap
λ	$(0.0127158; 997.491)$	$(0.0549575; 112.377)$
a	$(0.0100005; 5128.67)$	$(0.01; 1342.73)$
b	$(1.13517; 99.6114)$	$(0.896717; 123.845)$
r	$(1.24204; 142.748)$	$(0.624673; 128.555)$

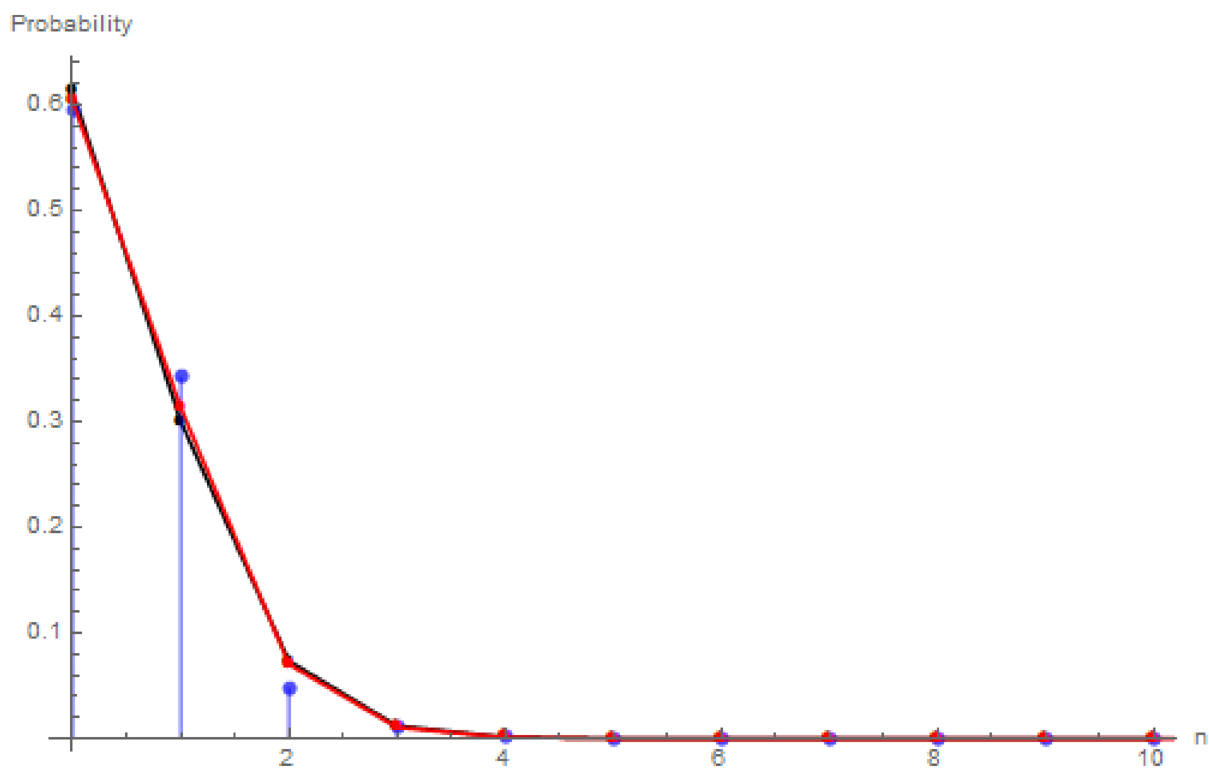


Figure 5.6: USA gun violence - Injured per incident

For the second variable, the number of fatalities per day, of the 28 weight functions that were tested, 9 models perform better than the Poisson, and 1 additional model may (the AIC and AICc disagree with the BIC). The best fit is achieved when $w(n;\phi) = (30.8585)_n$ and $\lambda = 0.557367$. This is shown in Figure 5.7. In this case, the weighted Poisson distribution gives a relatively good fit to the data.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	(0.526156; 0.588577)	(0.505001; 0.598864)	(0.493193; 0.608775)
a	(26.9723; 34.7448)	(25.8952; 37.9053)	(24.9895; 39.8908)

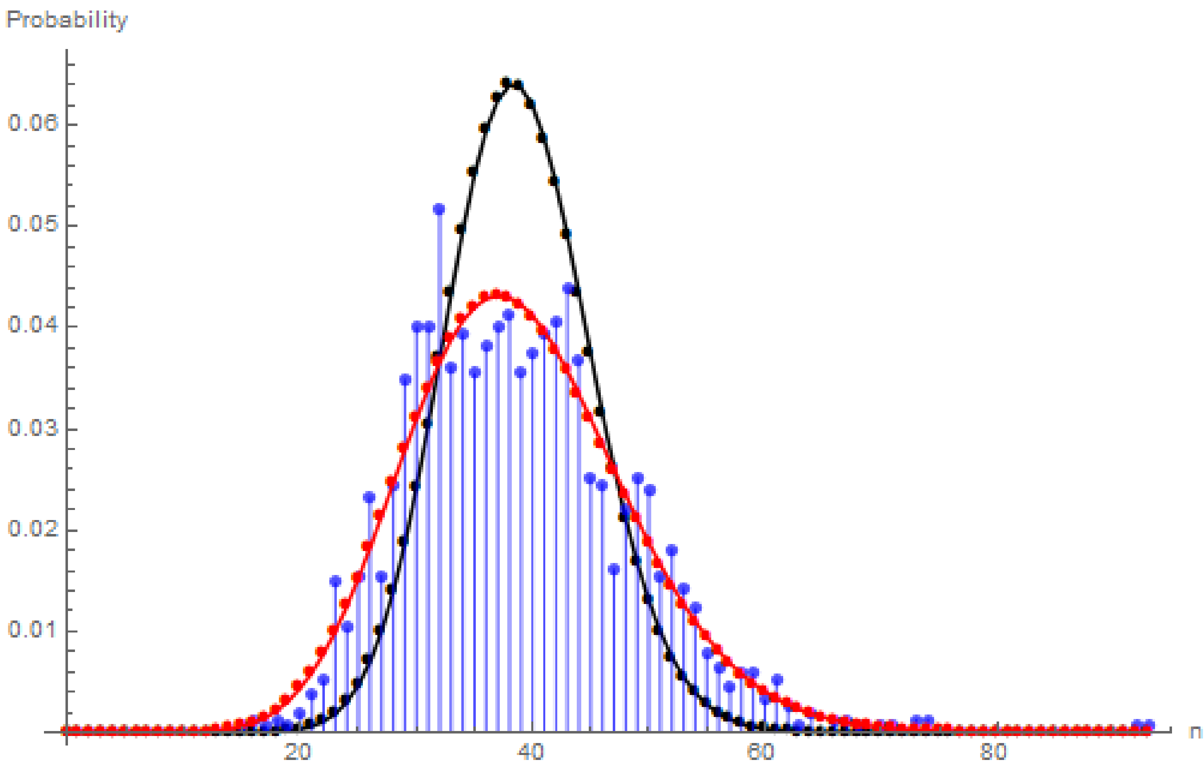


Figure 5.7: USA gun violence - Fatalities per day

The next dataset contains information about all of the goals scored in the English Premier League (EPL) between 2002 and 2016. This dataset cannot be found on Kaggle but was constructed piece by piece from other sources.

For the number of home team goals per game, of the 14 weight functions that were tested, only 1 model performs better than the Poisson, and 5 additional model may (the AIC and AICc disagree with the BIC). The best fit is achieved when $w(n;\phi) = (16.5054)_n$ and $\lambda = 0.084907$. This is shown in Figure 5.8. In this case, both the Poisson and weighted Poisson distributions give a very good fit to the data. The AIC of the weighted Poisson distribution is only 0.11% smaller than that of the Poisson distribution.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	(0.0505357; 0.119278)	(0.0497925; 0.112234)	(0.00710242; 0.188799)
a	(9.21295; 23.7979)	(11.2234; 29.3604)	(6.63044; 210.04)

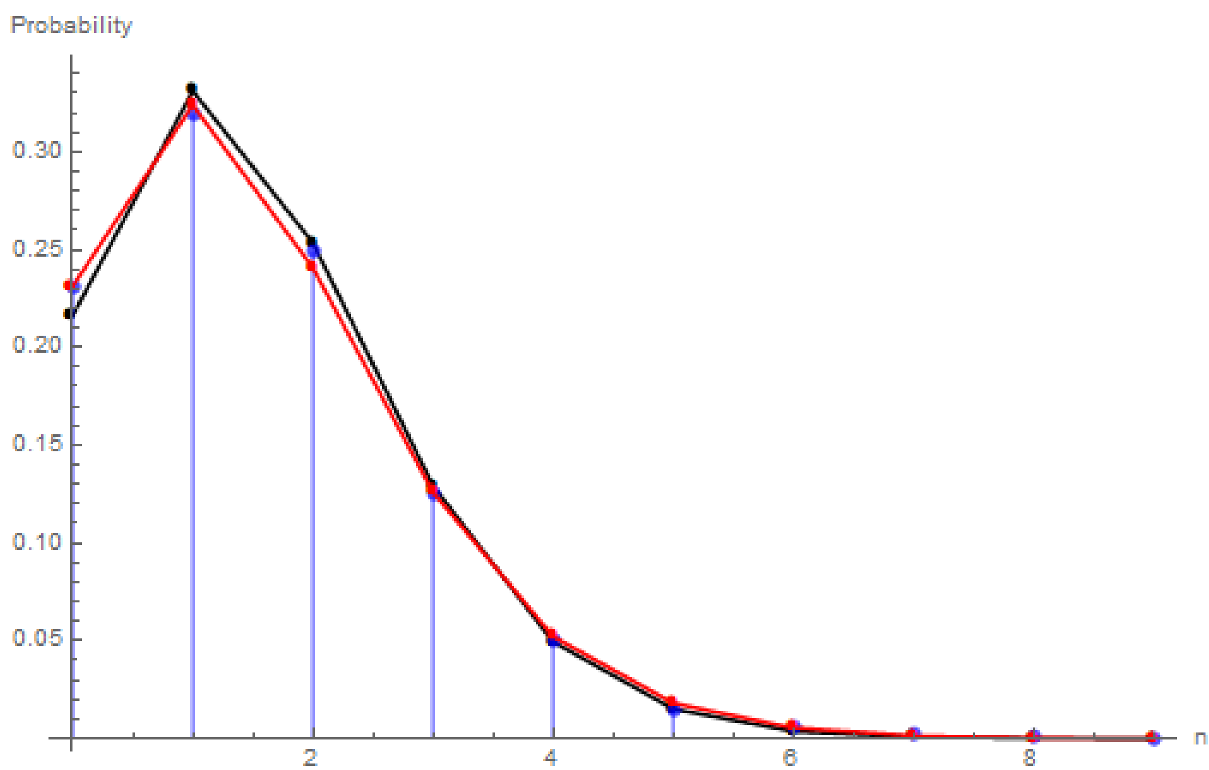


Figure 5.8: EPL games - Home team goals per game

The next dataset contains information about flights in the USA in 2015. This dataset can be found at `https://www.kaggle.com/usdot/flight-delays#flights.csv`.

For the number of minutes that arrivals are delayed, of the 26 weight functions that were tested, 12 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = (0.748927)_n$ and $\lambda = 0.977883$. This is shown in Figure 5.9. In this case, the weighted Poisson distributions gives an acceptable fit to the data, although the number of short delays of up to 20 minutes are underestimated.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	$(0.977832; 0.977934)$	$(0.968609; 0.983902)$	$(0.973909; 0.980871)$
a	$(0.747635; 0.75022)$	$(0.639775; 0.89591)$	$(0.672767; 0.855432)$

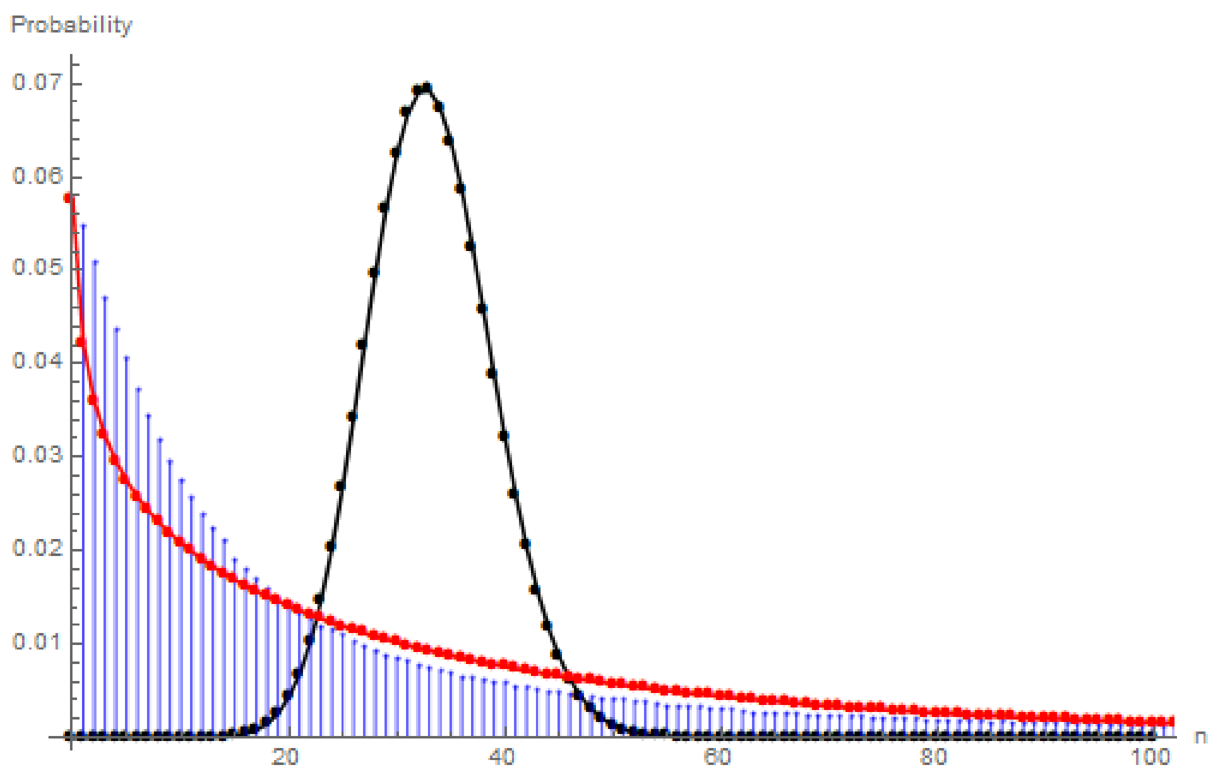


Figure 5.9: USA flights - Arrival delay

From all of the above plots, it is apparent that the weighted Poisson distribution performs well in modelling observed discrete data. This is the case not only when the Poisson distribution is not an appropriate modelling choice, but also when the underlying data appears to be roughly Poisson distributed. In many of the datasets (EPL home team goals - Figures 5.8, US gun violence injuries per incident - Figure 5.6) the Poisson distribution produced very good measures of fit. However, it was nevertheless outperformed by the weighted Poisson distribution, even after penalising the weighted Poisson distributions for the additional parameters.

Out of all of the weight functions that were presented in this thesis, there are two that gave the best fits most often (based on these datasets). These are when $w(n;\phi) = (a)_n$ and $w(n;\phi) = \frac{\Gamma(r+n)}{n!\Gamma(r)} \frac{Beta(a+r.b+n)}{Beta(a,b)}$ (although other weight functions are prevalent as well). That is not to say that the other weight functions are less useful, just that for these specific datasets, they did not result in optimal fit characteristics. It may very well be that for other datasets the other weight functions may be more prevalent. Additionally, the use of specific weight functions may vary depending on practical considerations. For example, when looking at the plot of the arrival delays of US flights (Figure 5.9), an airline may wish to choose a weight function that accurately models the tail of the distribution rather than values close to zero (since passengers, airlines and airports are likely to be dramatically more inconvenienced by longer delays).

Two of the weight functions, $w(n;\phi) = \left(\begin{array}{c} m \\ n \end{array}\right) p^n (1-p)^{m-n}$ and $w(n;\phi) = \frac{ab^a}{n^{a+1}}$, were often excluded from the above fit attempts. This was done because of the excessive computational time required to calculate the maximum likelihood estimates. When the weight function is chosen to be the binomial probability mass function, the resulting weighted Poisson distribution contains an m^{th} order Laguerre polynomial. The parameter m has to be at least as large as the largest observation in the dataset. Since a general closed-form expression does not exist for Laguerre polynomials, this can quickly become computationally impractical to calculate. As a result, the binomial weight function was only used when the maximum observed value in the dataset was relatively small (although no consistent cut-off point for "small" was used). When the weight function is chosen to be the Pareto probability density function, a similar problem occurs. The parameter b is defined to be the minimum observed value. If the value of b is relatively large, the computation time required to calculate maximum likelihood estimates dramatically increases. (In the tables in Chapter 10 these cases are clearly labelled as "Computationally intractable")

Unsurprisingly the Pareto weight function gave the best fit for the "Mass shootings - Victims per incident" dataset (figures 5.1 and 5.2). Based on all of the models presented in this chapter a general statement can be made with regards to the truncating weight functions: These specific weight functions can perform very well, however, for them to provide a significantly better fit than the Poisson distribution the mode of the data has to be close to the truncated part of the distribution. For example, in the "Mass shootings - Victims per incident" dataset, the mean of the data is 5.37894 and the minimum value is 4. For this dataset, all three lower truncated weight functions resulted in models which performed better than Poisson models. In contrast, the "US gun violence - Injuries per day" dataset had a mean of 75.8547 and a minimum value of 17. None of the truncating weight functions provided a better than Poisson distribution in this situation. Since the Poisson distribution is centred around 75.8547 it naturally has a close to zero probability of observing a value of 17 without artificially truncating the distribution at that point.

It should also be mentioned that the lower truncating weight functions might be of questionable practical use. It might be easier, and give a larger set of candidate weight functions, if the data are merely transformed to start at 0. However, there may be specific practical situations where using these lower truncating weights can be of use. (For instance when data values are consistently larger than 0, but there is no theoretical reason why the values cannot be lower.)

For the analyses in this chapter, the models that were applied to specific datasets were based purely on the domains that the various weight functions allowed. This decision was made because the aim was to demonstrate the wide range of potential shapes that the weighted Poisson distribution could assume and accurately model. Practically speaking, however, this approach is unlikely to be implemented in the real world. For example, in the "US flights - Arrival delay" dataset (Figure 5.9), the best weighted Poisson function allows for zero observations, which is practically nonsensical since delays will have observations of at least 1 minute (unless zeros are used to denote flights that are on time). The reverse situation also occurs where models which are strictly defined on $n \in \mathbb{N}_1$ were applied to datasets which, although they did not demonstrate zero observations, could, in practice, have these values. As an example, in the "US gun violence - Fatalities per day" dataset (Figure 5.7), the minimum observation was 15 and the maximum 93. In this case, weight functions which can only model data strictly larger than zero were applied (in addition to those that can accommodate zeros). A practitioner may wish to discard these weight functions since there is no theoretical reason why there might not be a day in which no fatalities occur.

It should also be noted that there were a few limited situations where weight functions experienced convergence problems. In general, these occurred when nonsensical weight functions were fit to data. This happened most often when the zero-inflated Poisson distribution was applied to a dataset that had no zero observations. In a few rare situations, convergence problems also occurred when a pair of dual weight functions were considered. Occasionally if one weight function gave a very good fit, the dual partner of that function would not converge. (These situations are clearly labelled in the tables in Chapter 10 as "Convergence".)

One last caveat must be provided regarding the model fits in the chapter. In the practical modelling environment, if a parameter is estimated to have a value that is not statistically significantly different from zero, it is often assumed to be zero, and the reduced model can be refit to the data. This may lower the respective AIC, AICc and BIC values for the fit since the number of parameter estimates is being reduced. This approach of reducing parameters was not implemented in this thesis, since the changes to the AIC, AICc and BIC were usually negligible and did not affect the model selection. As a result, some of the AIC, AICc and BIC values reported in Chapter 10 are slightly larger than they would be if the models were implemented in a real-world setting.

5.3 Previous method comparisons

In this section, the fit of the various weighted Poisson distributions derived in this thesis will be compared to other modelling methodologies presented in Chapter 2. The scope of this section is limited, however, because many of the papers mentioned in Chapter 2 either did not provide/publish their datasets, did not provide sufficient information to recreate the datasets, or used datasets that are not publicly available. In these cases, direct comparisons are impossible and are omitted from this section. However, if data could be obtained (either through the papers or by other means), the models fitted are compared to the fitted weighted Poisson distributions.

For the plots in this section, similar to the previous sections, the observed proportions in the data will be represented by blue dots and a vertical line and the best weighted Poisson fit will be graphed in red. However, unlike in previous sections, the corresponding Poisson distribution will not be plotted. Instead, the best proposed fitted model from the relevant published paper will be provided in green to allow easy comparison between the two methods.

5.3.1 Generalised Poisson distribution - Bosch and Ryan

Bosch and Ryan [16] proposed using generalised Poisson distributions to model discrete data. In the paper, they used two datasets (as has already been mentioned in Chapter 2). The first dataset concerns the number of sea-urchin eggs fertilised in various time intervals. The second recorded the number of annual doctor visits per patient.

When weighted Poisson distributions are fit to the sea-urchin dataset, it is found that the weighted Poisson distribution outperforms their original generalised Poisson approach. The best fit is achieved when $w(n;\phi) = n + 0.00515226$ and $\lambda = 0.204567$. This is shown in Figure 5.10.

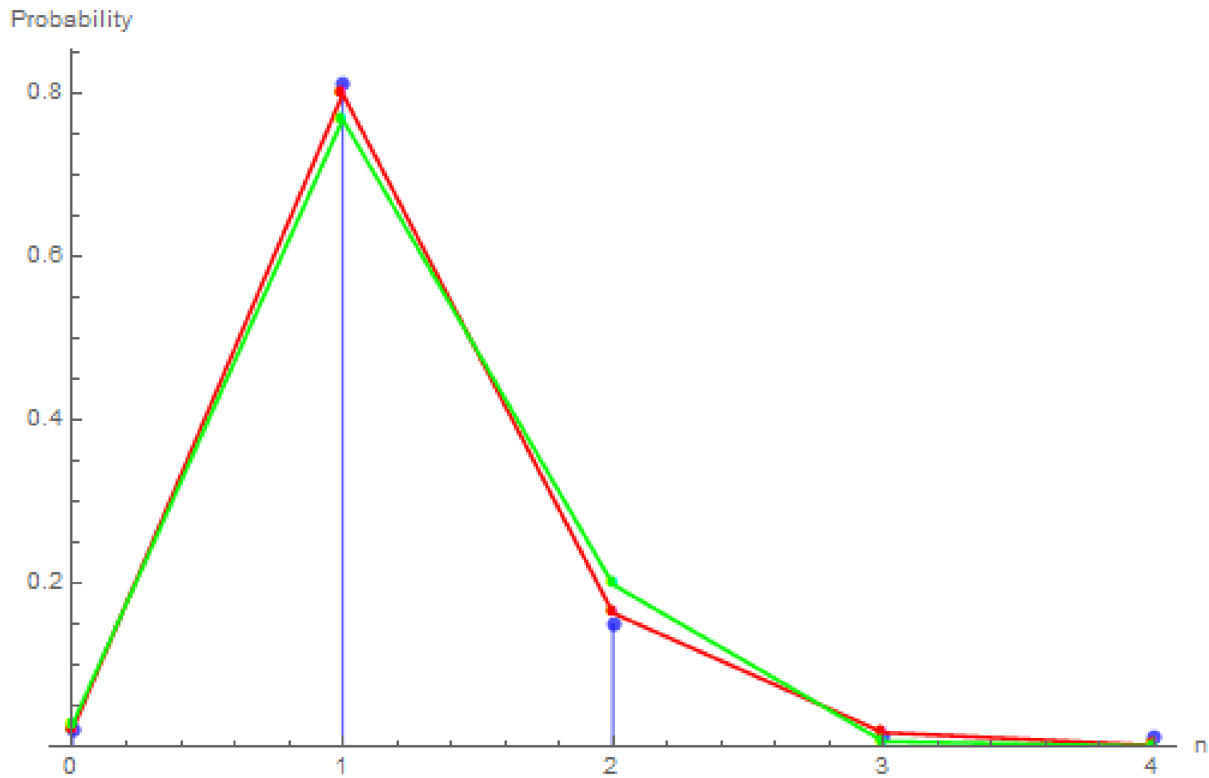


Figure 5.10: Fertilisation of sea-urchin eggs

When weighted Poisson distributions are fit to the doctor visit dataset it is found that the weighted Poisson distribution does not outperform their original generalised Poisson approach. The best weighted Poisson fit is achieved when $w(n;\phi) = \frac{\Gamma(0.541878+n)}{n!\Gamma(0.541878)} \frac{Beta(330.535878,0.541878+n)}{Beta(329.994,0.541878)}$ and $\lambda = 296.025$. This is shown in Figure 5.11.

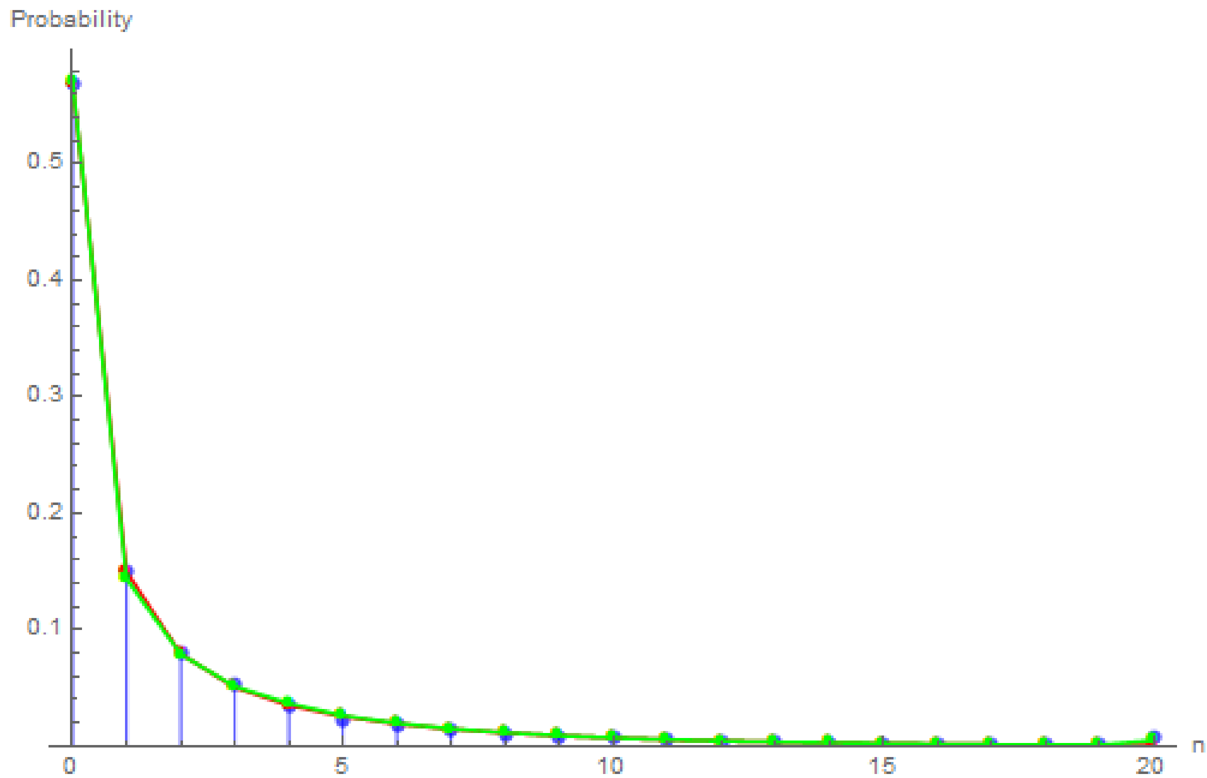


Figure 5.11: Doctor visits

However, it is observed that on the main body of the data, the two models give near-identical, near-perfect fits, with the generalised Poisson model only providing a better fit at the extreme point of the data (20). If one inspects the actual data, it is observed that the value of 20 was chosen as the truncating point. In other words, "20" actually refers to "≥ 20". Additionally, a substantial number of observations fall into this category.

It may well be that using Markov processes to model data with such a "bump" at the end of the support of the distribution gives excellent fit characteristics (as Bosch and Ryan [16] did). However, Sur et al. [134] demonstrated that this could also be achieved with COM-Poisson mixture models.

5.3.2 Generalised Poisson distribution - Consul and Jain

Consul and Jain [29] proposed using an alternative form of generalised Poisson distribution. In the paper, they looked at four datasets to demonstrate that their distribution could closely adhere to the shapes of other discrete distributions (binomial, negative binomial and Poisson). The first dataset recorded the number of accidents of women working on shells, and the second recorded the number of lost papers found in the Bell Telephone and Telegraph Buildings in New York City, both of which are traditionally assumed to be negative binomially distribution. The third dealt with the number of deaths caused by horse kicks in the Prussian army, which is traditionally assumed to be Poisson distributed. The fourth dataset recorded "ten shots fired from a rifle at each of 100 targets" and is assumed in the literature to be binomially distributed.

It should be noted that Consul and Jain [29] performed their analyses using the method of moment estimates for the parameters. For consistency's sake, maximum likelihood estimates are obtained for their model, which enables more accurate comparisons between the models.

For the first dataset, the number of accidents of women working on shells, the weighted Poisson distribution outperforms the generalised Poisson distribution. The best weighted Poisson fit is achieved when $w(n;\phi) = (0.865116)_n$ and $\lambda = 0.349703$. This is shown in Figure 5.12.

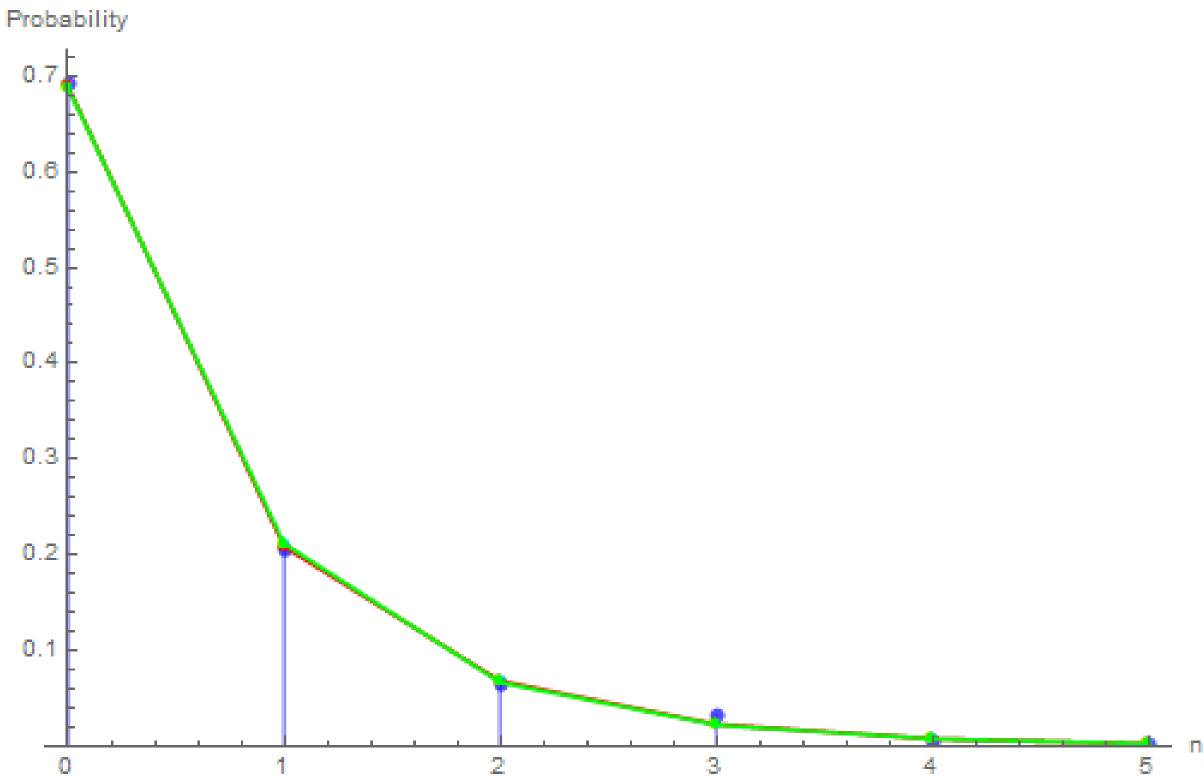


Figure 5.12: Shell accidents

It is observed, however, that the two fits are nearly identically.

For the second dataset, the number of deaths due to horse kicks, the weighted Poisson distribution outperforms the generalised Poisson distribution. The best weighted Poisson fit is achieved when $w(n;\phi) = \mathit{Boole}(n \leq 4)$ and $\lambda = 0.61194$. It is unlikely, however, that there would be a practical reason why the number of horse kick deaths would be limited to 4. Consequently the second best performing weight function was selected, which is achieved when $w(n;\phi) = n + 15.1684$ and $\lambda = 0.573565$. This is shown in Figure 5.13.

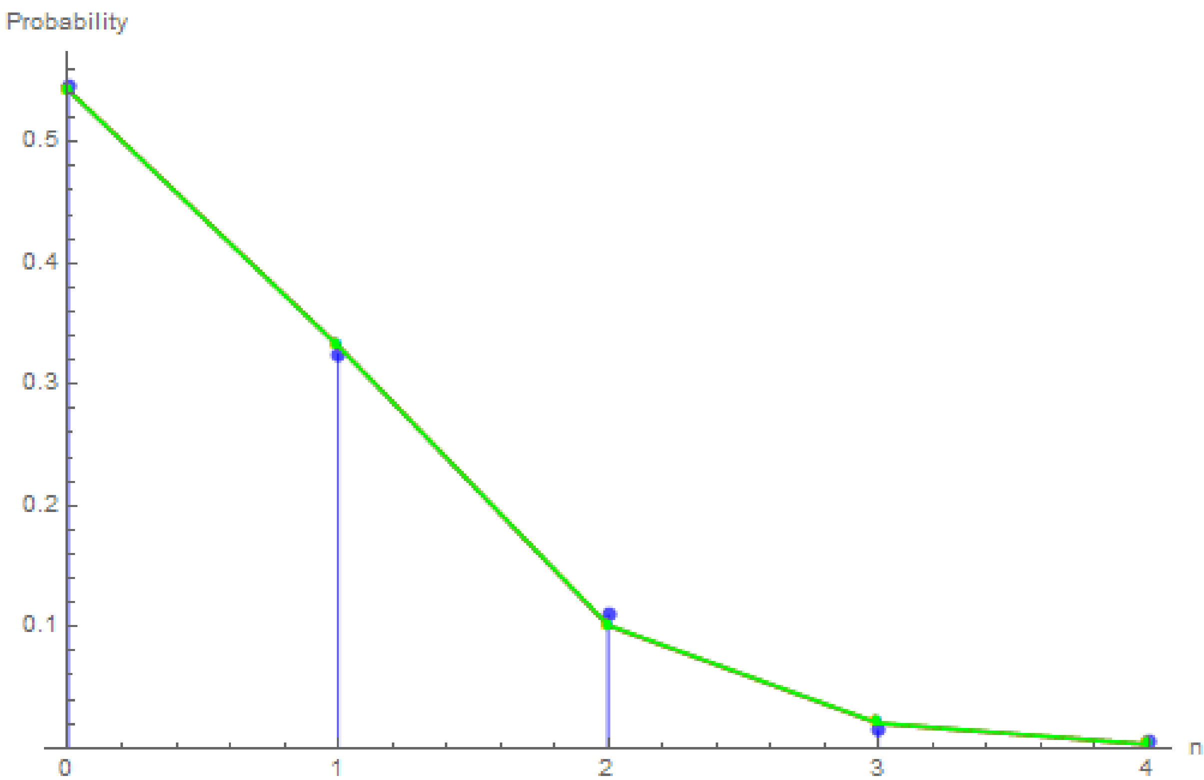


Figure 5.13: Horse kick deaths

It is observed that the two fits are nearly identical.

For the third dataset, the number of lost items in the telephone building, the weighted Poisson distribution outperforms the generalised Poisson distribution. The best weighted Poisson fit is achieved when $w(n;\phi) = (n+0)(n^2 - 2.15404n + 1.2518)$ and $\lambda = 0.884193$. This is shown in Figure 5.14.

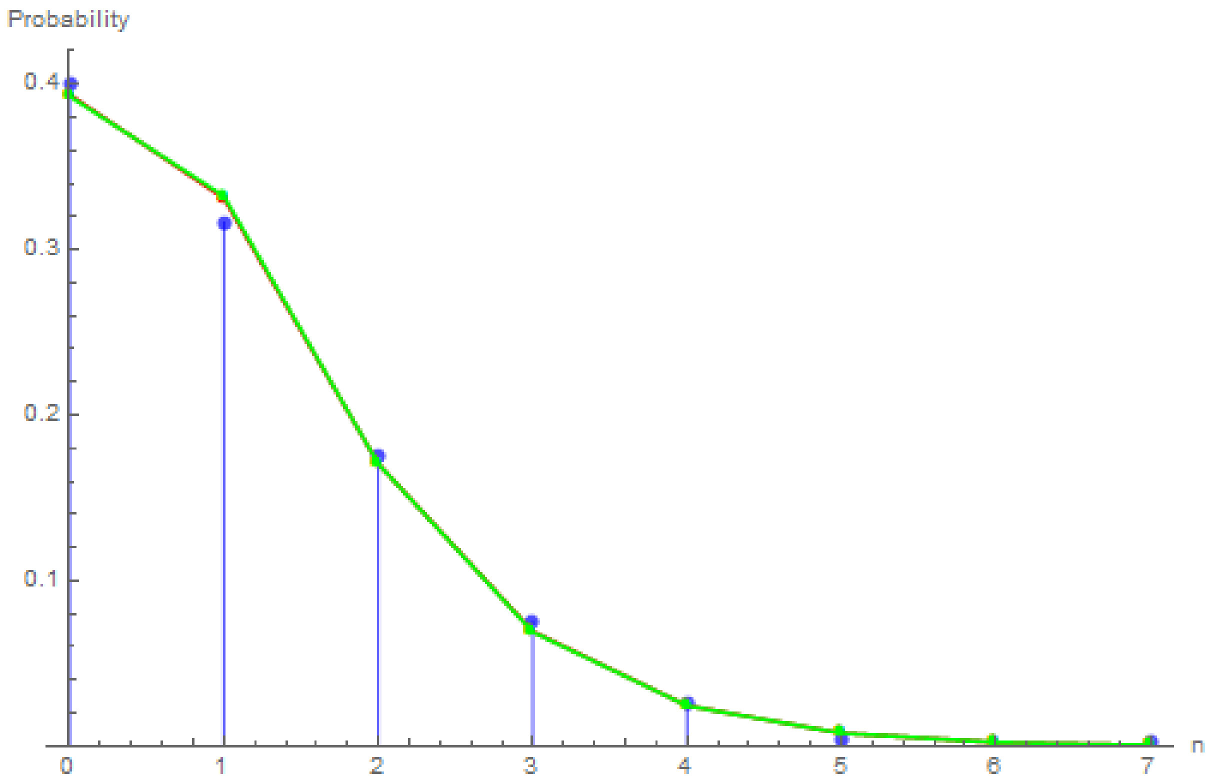


Figure 5.14: Lost items

It is again observed that the two fits are nearly identical.

For the fourth dataset, the number of rifle shots at targets, the weighted Poisson distribution does not outperform the generalised Poisson distribution. The best weighted Poisson fit is achieved when $w(n;\phi) = (n)_{13.3313}$ and $\lambda = 1.123831$. This is shown in Figure 5.15.

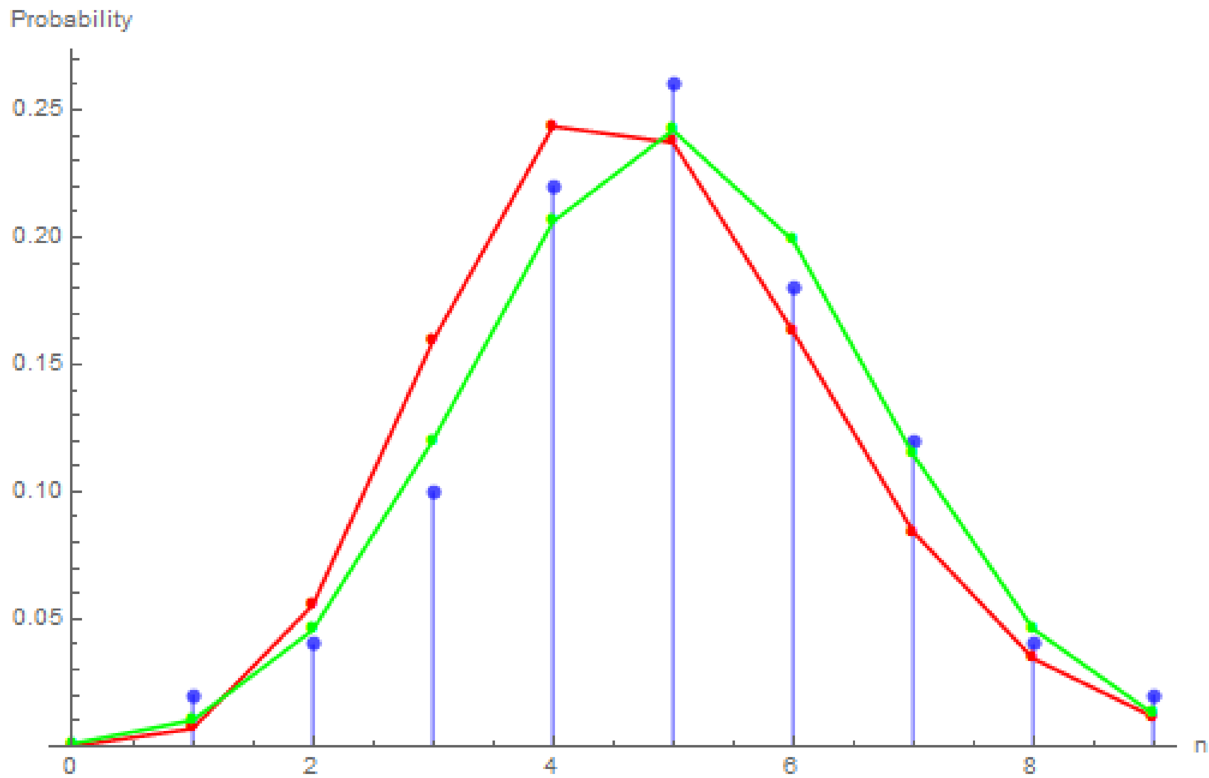


Figure 5.15: Rifle shots on targets

The above four datasets and plots provide some evidence that the weighted Poisson performs roughly the same as the generalised Poisson distribution when the distribution of the data being modelled resembles the binomial, negative binomial or Poisson distributions.

Consul and Jain [29] authored a book that discussed, in much greater detail, the theory and applications of their proposed model. Since the four plots above were intended to demonstrate that the model could accommodate data from well-known distributions, three datasets from the book will be analysed that do not appear to be realised from these distributions. The first dataset concerns the number of home injuries to 122 men between 1937 and 1942. The second, which is also used by Castillo and Perez [38], recorded the number of strike outbreaks in the coal industry in the United Kingdom between 1948 and 1959. The third contains the number of authors who published papers in the journal "Theoretical Statistics and Probability" between 1700 and 1943.

For the first dataset, the number of home injuries to men, the weighted Poisson distribution outperforms the generalised Poisson distribution. The best weighted Poisson fit is achieved when $w(n;\phi) = (4.41713)_n$ and $\lambda = 0.109111$. This is shown in Figure 5.16.

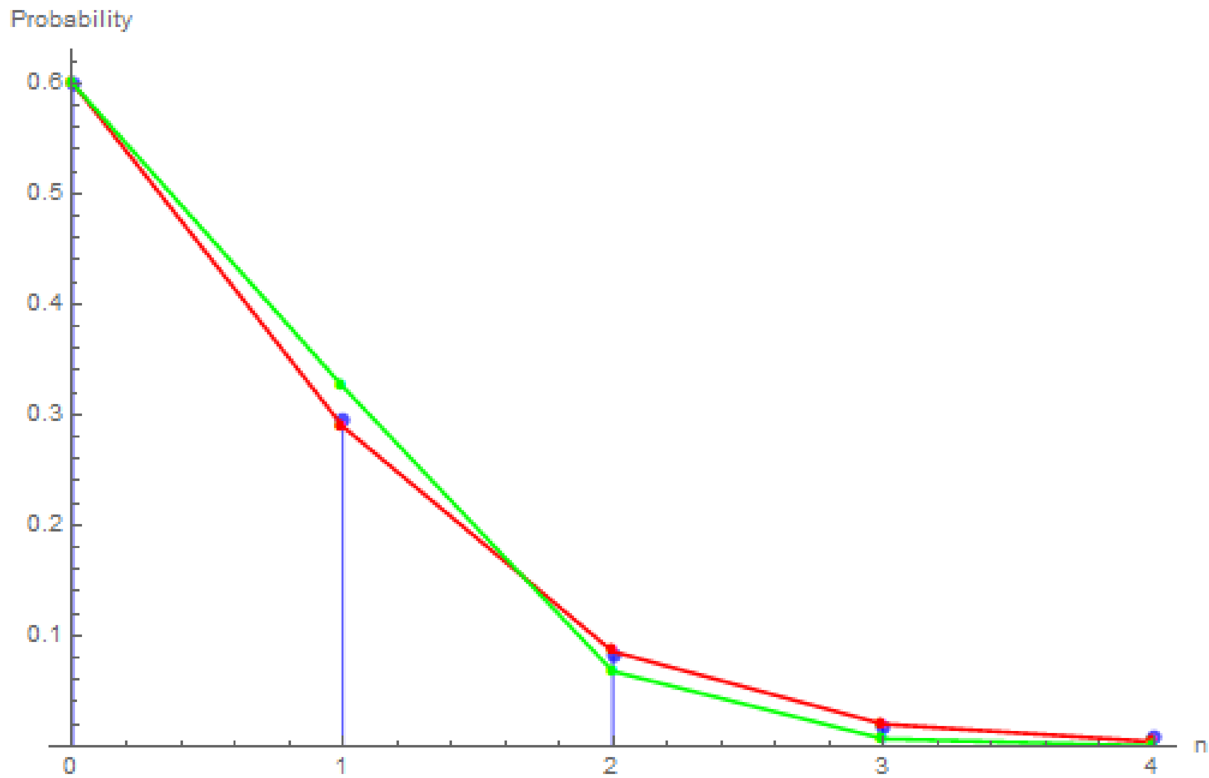


Figure 5.16: Home accidents to men

For the second dataset, the number of strike outbreaks, the weighted Poisson distribution outperforms the generalised Poisson distribution. The best weighted Poisson fit is achieved when $w(n;\phi) = n + 0.418559$ and $\lambda = 0.466504$. This is shown in Figure 5.17.

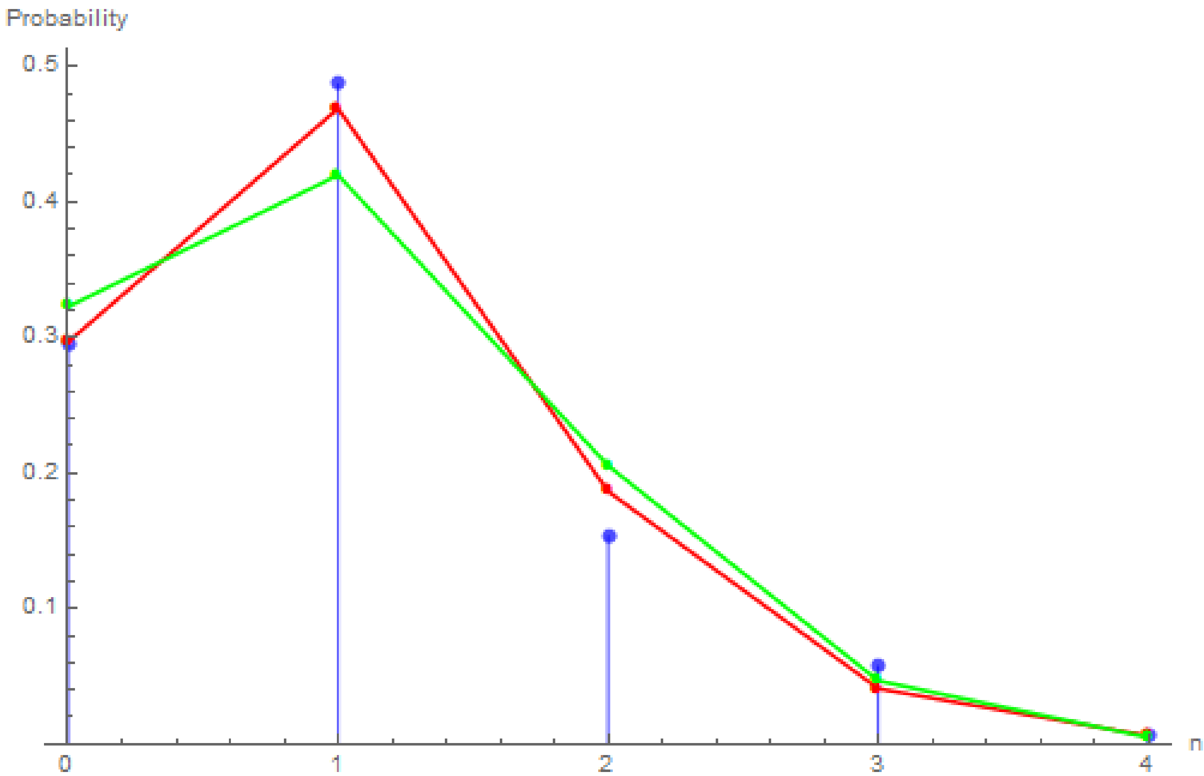


Figure 5.17: Strike outbreaks

For the third dataset, the number of authors in "Theoretical Statistics and Probability", the weighted Poisson distribution outperforms the generalised Poisson distribution. The best weighted Poisson fit is achieved when $w(n;\phi) = \frac{\Gamma(0.449024+n)}{\Gamma(0.449024)n!}\frac{Beta(19.100524,0.449024+n)}{Beta(18.6515,0.449024)}$ and $\lambda = 22.5577$. This is shown in Figure 5.18.

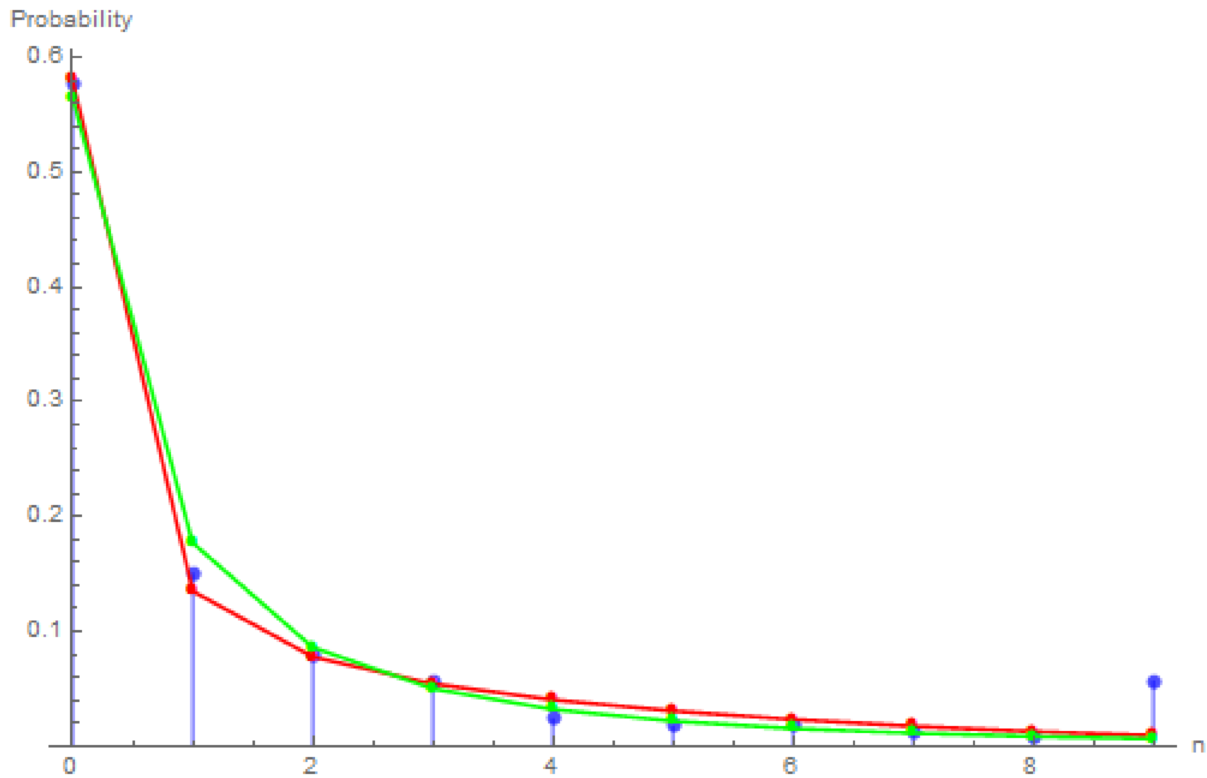


Figure 5.18: Statistical journal authors

5.3.3 COM-Poisson - Shmueli et al.

As has already been discussed, the COM-Poisson distribution is a member of the weighted Poisson family. In fact, it is arguably the most commonly used weighted Poisson distribution. The aim here is to establish how the newly derived distributions perform in comparison to the COM-Poisson. Shmueli et al. [126] applied the COM-Poisson distribution to two discrete datasets. The first dataset contains a count of the number of syllables in the Hungarian words. The second recorded the sales numbers of a specific item of clothing.

Due to the limited number of COM-Poisson data fits available in the literature, some of the datasets discussed in Section 5.2 will be reanalysed with the COM-Poisson distribution. The aim is to more accurately compare the performance of the COM-Poisson against the other weighted Poisson distributions. These specific datasets have not been cherry-picked. They were chosen based on three criteria: First, any peculiar datasets (truncated, bi-modal, zero-inflated), which would probably result in bad fits for the COM-Poisson, are excluded. Second, the chosen datasets include a range of fits, from near-perfect to less than optimal under the weighted Poisson models of this thesis. Lastly, they all include small counts. The range of fit characteristics is important because in both of the datasets in Shmueli et al. [126], the COM-Poisson and weighted Poisson distributions give very good (near-identical) fits. The third condition that only small count levels were analysed was done purely for computational reasons. Since the COM-Poisson does not have a closed-form expression for its probability mass function, calculating maximum likelihood estimates for datasets that have a large number of observations, and a broad range of values can become computationally intensive.

For the first dataset discussed in Shmueli et al. [126], the length of Hungarian words, the weighted Poisson distribution outperforms the COM-Poisson distribution. The best weighted Poisson fit is achieved when $w(n;\phi) = (n+0)(n^2 - 2.15404n + 1.2518)$ and $\lambda = 0.884193$. This is shown in Figure 5.19. While the weighted Poisson fit is better, the difference is minimal.

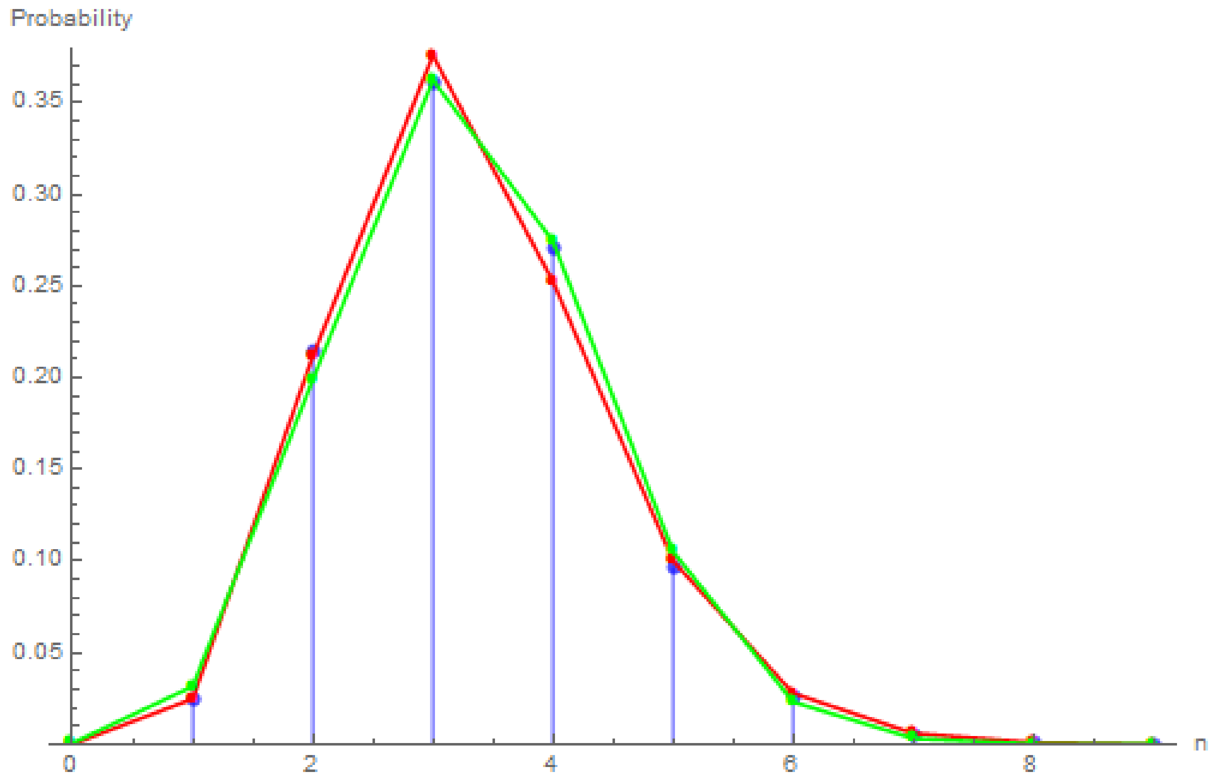


Figure 5.19: Hungarian word length

For the second dataset, the number of clothing sales, the weighted Poisson distribution does not outperform the COM-Poisson distribution. The best weighted Poisson fit is achieved when $w\left(n;\phi\right)=\frac{\Gamma(2.6095+n)}{n!\Gamma(2.6095)}\frac{Beta(1752.71095,0.571388+n)}{Beta(1755.11,0.571388)}$ and $\lambda=1152.17$. This is shown in Figure 5.20. While the COM-Poisson fit is better, the difference is again marginal.

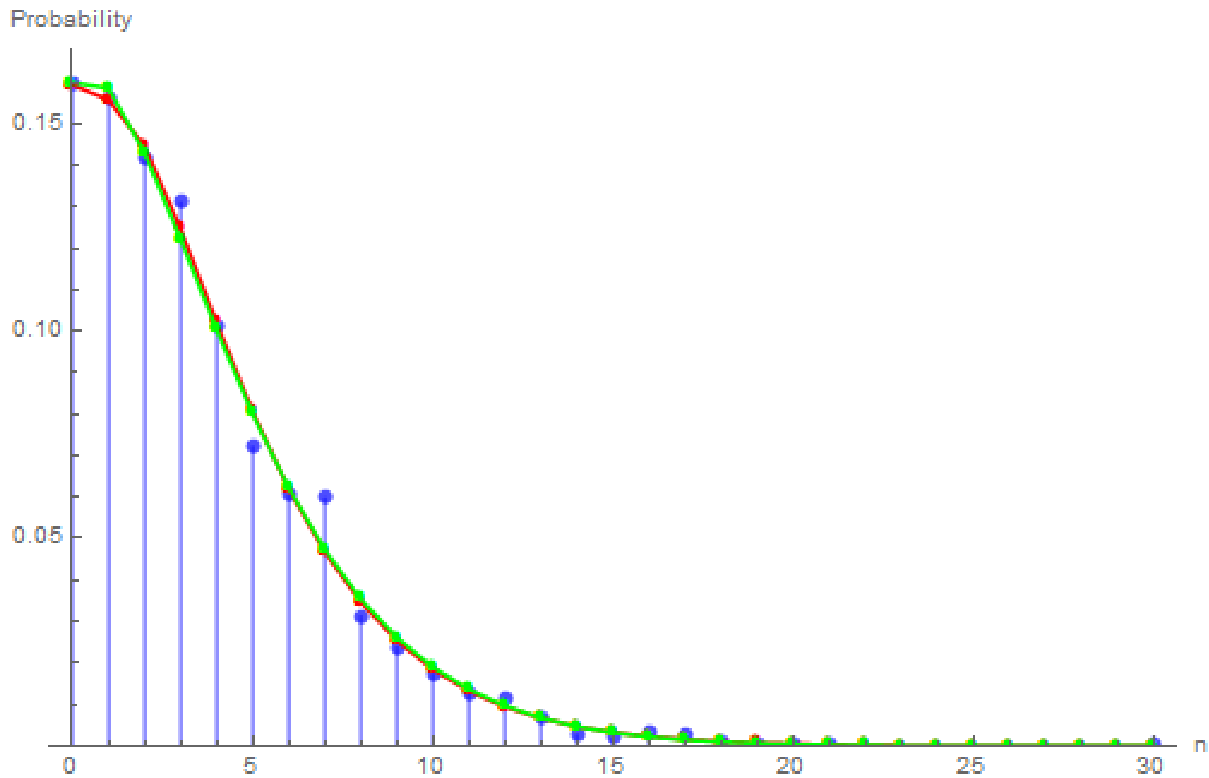


Figure 5.20: Clothing item sales

The first novel dataset to which both the weighted Poisson and the COM-Poisson distributions are fitted is the "Number of airplane accidents per month" (shown in Figure 10.5). As was stated the best weighted Poisson fit occurs when $w(n;\phi) = (34.9643)_n$ and $\lambda = 0.148743$. For the COM-Poisson distribution, the best fit occurs when $w(n;\phi) = (n!)^{1-0.837851}$ and $\lambda = 4.49253$. In this situation, the weighted Poisson outperforms the COM-Poisson, but only marginally. The corresponding plots are given in Figure 5.21.

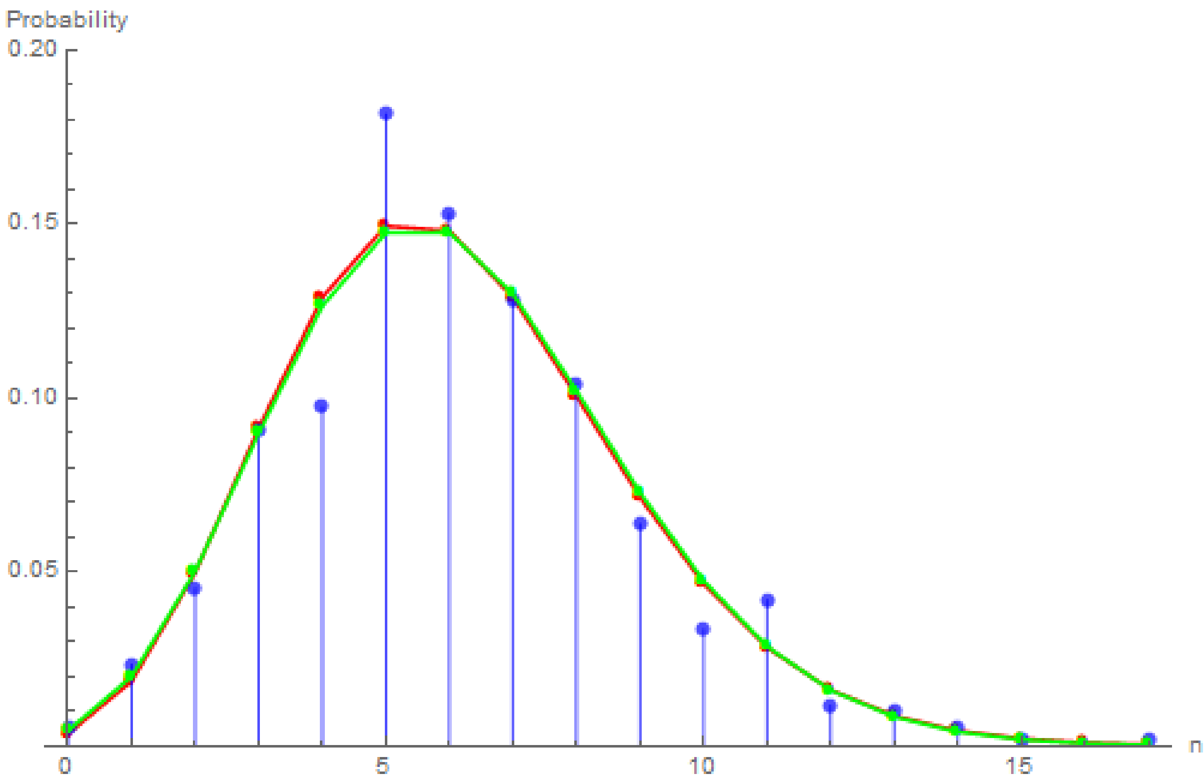


Figure 5.21: Airplane accidents - Incidents per month

The next dataset reanalysed is "US gun violence - Injuries per incident" (Figure 5.6). As stated the best weighted Poisson fit occurs when $w(n;\phi) = \frac{\Gamma(3.07126+n)}{n!\Gamma(3.07126)} \frac{Beta(126.39226,3.1556+n)}{Beta(123.321,3.1556)}$ and $\lambda = 6.93968$. For the COM-Poisson distribution, the best fit occurs when $w(n;\phi) = (n!)^{1-1.12583}$ and $\lambda = 0.59374$. In this situation the weighted Poisson outperforms the COM-Poisson. These plots are given in Figure 5.22. While the weighted Poisson fit is better, the difference is marginal.

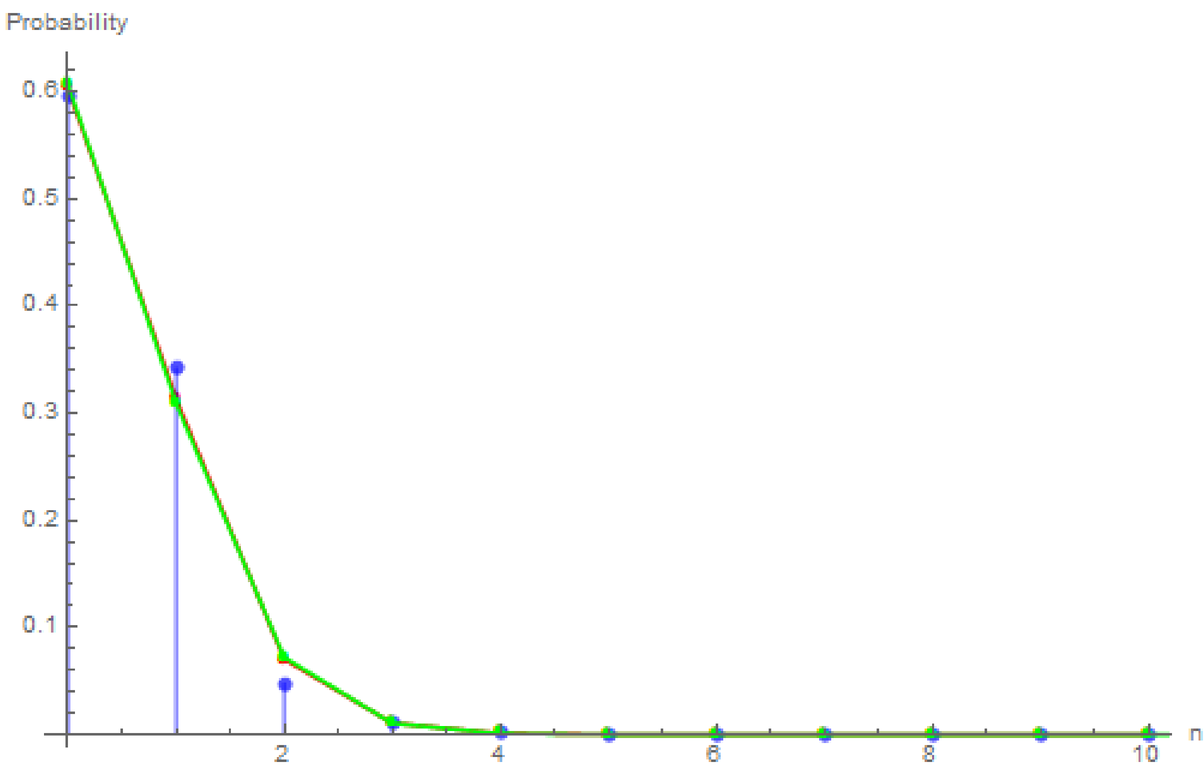


Figure 5.22: US gun violence - Injured per incident

The next dataset reanalysed is "Canada accidents - Vehicles per incident" (Figure 10.16). As stated the best weighted Poisson fit occurs when $w(n;\phi) = (n + 9.1268 \times 10^{-12})(n - 0.324713)^2$ and $\lambda = 0.340524$. For the COM-Poisson distribution, the best fit occurs when $w(n;\phi) = (n!)^{1-1.97762}$ and $\lambda = 5.45389$. In this situation the weighted Poisson substantially outperforms the COM-Poisson. These plots are given in Figure 5.23.

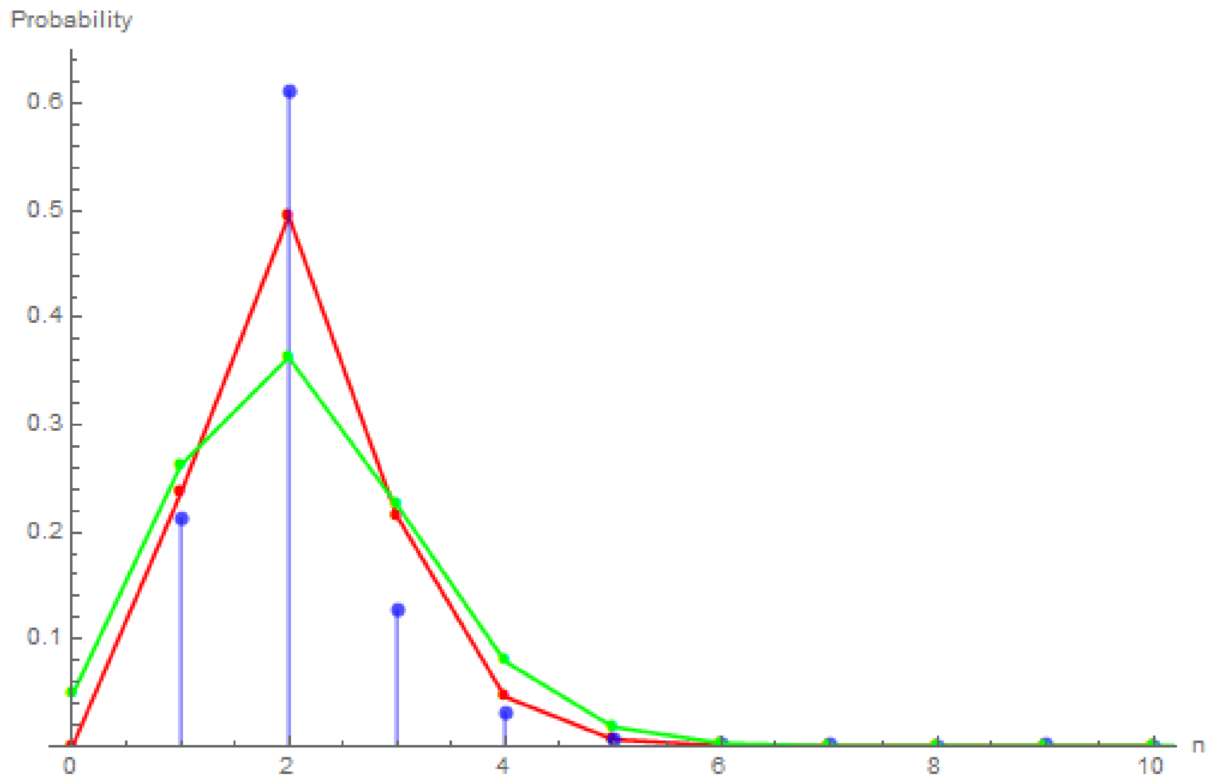


Figure 5.23: Canada accidents - Vehicles per incident

From the above five plots and sets of fits, it appears as if the range of new weighted Poisson distributions performs either better than the COM-Poisson distribution or roughly equally well. (This is, of course, anecdotal, and has not been vigorously tested.)

5.3.4 Weighted Poisson - Castillo and Perez

Castillo and Perez [39] applied four variations of their proposed weighted Poisson distribution: $w(n;\phi) = e^{r\left(-\sqrt{n^2+a}\right)}$, $w(n;\phi) = e^{r\left(-e^{-an}\right)}$, $w(n;\phi) = e^{r\left(\frac{n+1}{n+a}\right)}$ and $w(n;\phi) = e^{rln(n+a)}$. The resulting distributions were fitted to a dataset that contained information about car accidents. All four of these weight functions perform better than the weighted Poisson distributions derived in this thesis. The best variation from Castillo and Perez [39] occurred when $w(n;\phi) = e^{-4.3\left(-\sqrt{n^2+32}\right)}$ and $\lambda = 0.115553$. The best weighted Poisson (from this thesis) fit is achieved when $w(n;\phi) = (0.701512)_n$ and $\lambda = 0.234045$. These two fits are shown in Figure 5.24.

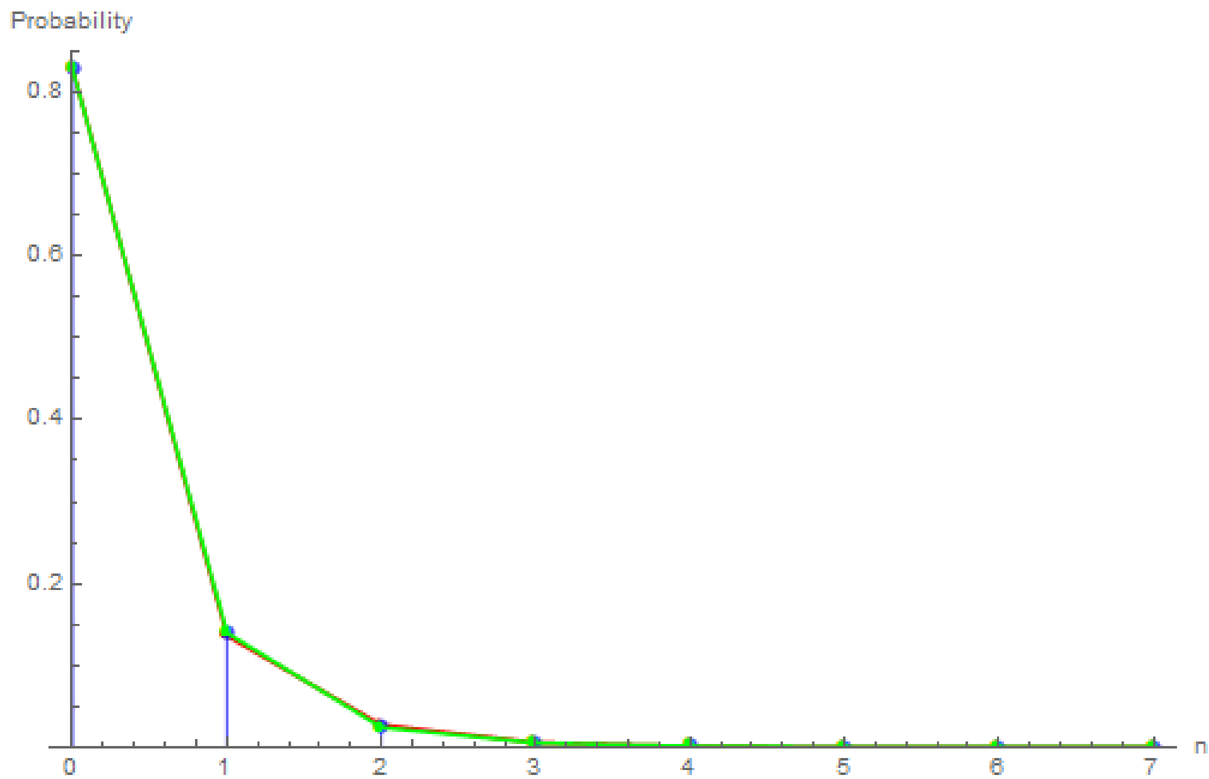


Figure 5.24: Car accidents

Even though the Castillo and Perez weight functions perform better; from a graphical perspective, the two models are indistinguishable.

Castillo and Perez [38] fitted their weighted Poisson distribution to two datasets. The first is the same dataset used by Consul and Jain [29], which recorded the number of accidents to women working on shells and the second recorded the number of strike outbreaks in the United Kingdom.

For the first dataset, number of accidents to women working on shells, the weighted Poisson distributions in this thesis outperform their proposed weighted Poisson distribution. The best weighted Poisson fit is achieved when $w(n;\phi) = (0.865116)_n$ and $\lambda = 0.349703$. This is shown in Figure 5.25. As can be seen, the performance is only marginally better.

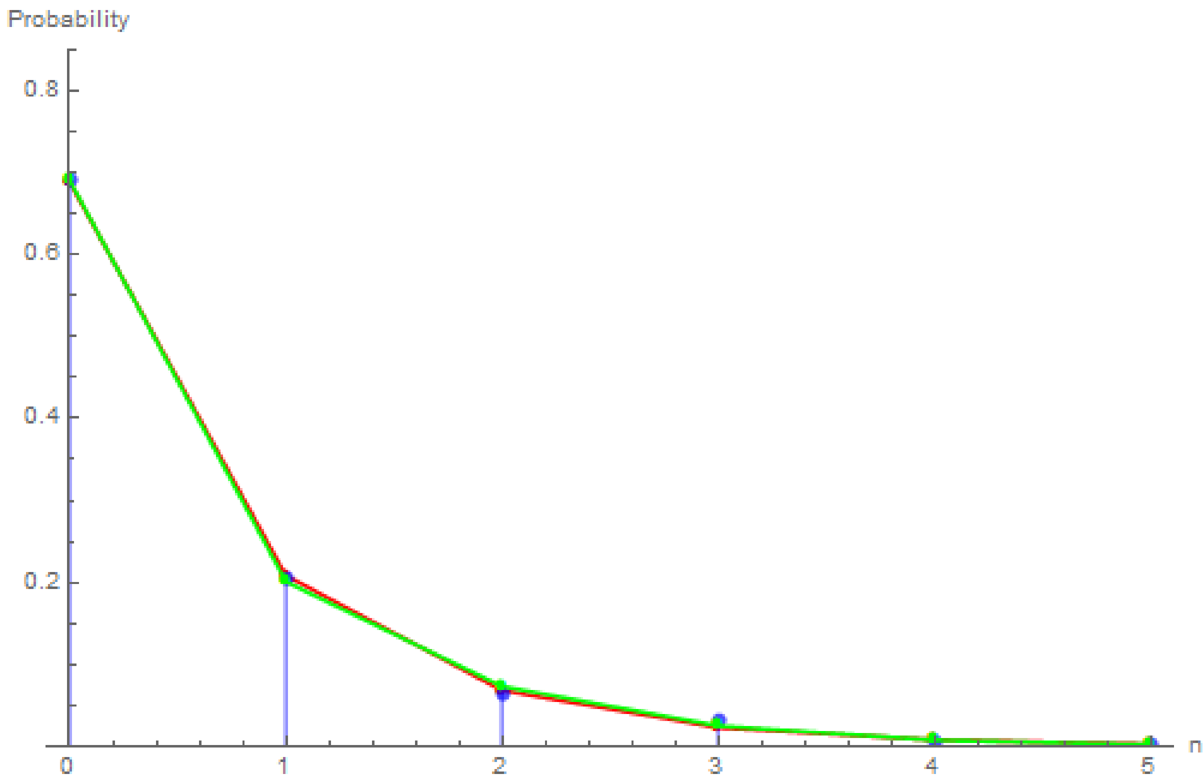


Figure 5.25: Castillo and Perez - Shell accidents

For the second dataset, number of strike outbreaks, the weighted Poisson distributions in this thesis outperform their proposed weighted Poisson distribution. The best weighted Poisson fit is achieved when $w(n;\phi) = n + 0.418559$ and $\lambda = 0.466504$. This is shown in Figure 5.26. As can be seen the new weighted Poisson performs better.

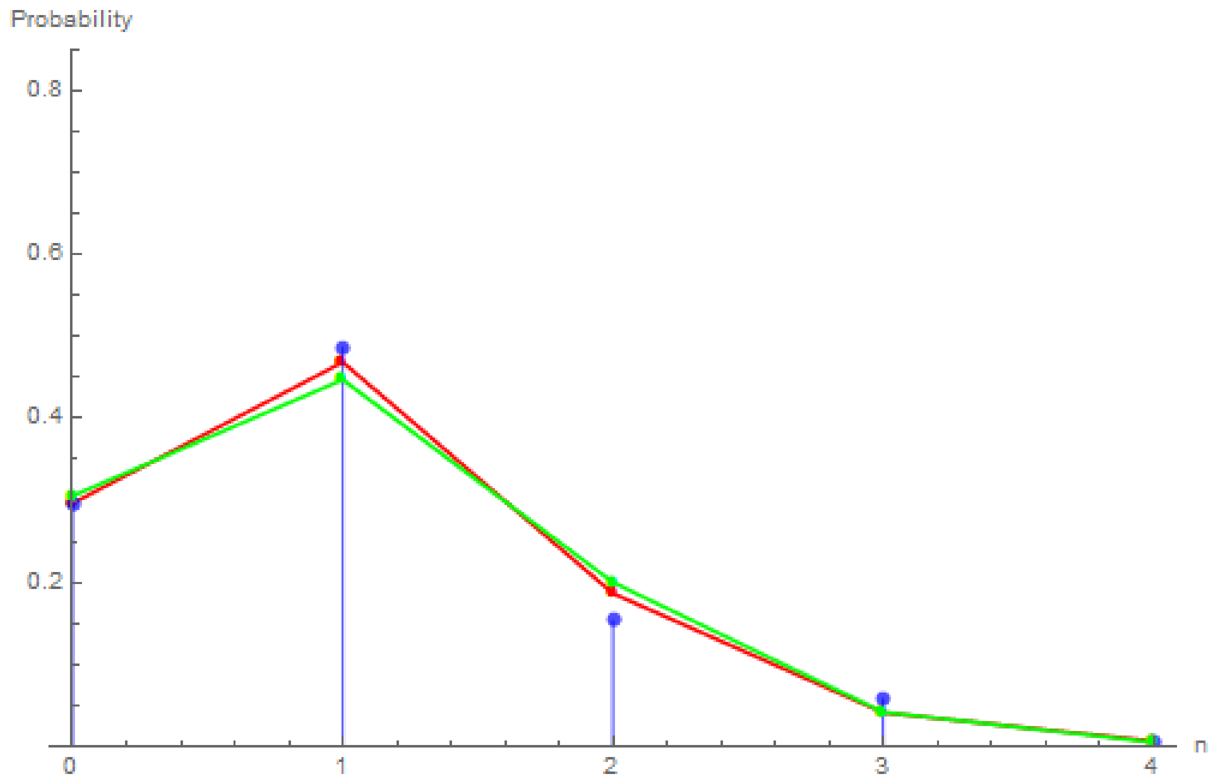


Figure 5.26: Castillo and Perez - Number of strike outbreaks

In addition to the 8 fits presented in Section 5.2, 25 other variables were also modelled. This was done to give a more accurate view of the prevalence of the newly derived weighted Poisson distributions, as well as to get a better sense of the comparative performance of the weighted Poisson distributions against the Poisson. These additional fits can be found in Chapter 10. The modelled variables come from a range of different industries/settings and include:

- Weekly sales figures of items over the span of a year (which was discussed in the introductory chapter). The sales numbers of item 409 and 726 were reanalysed.
- Global airplane accidents from 1960 to 2009. The number of fatalities per incident as well as the number incidents per month were analysed.
- USA mass shooting between 2013 and 2017. In addition to the number of victims per incident (Figure 5.1) and incidents per day (Figure 5.3), the number of injuries per

incident, injuries per day, injuries per month, fatalities per incident, fatalities per day, fatalities per month and incidents per month were also analysed.

- Vehicle accidents in Great Britain in 2015. In addition to the number of incidents per hour (Figure 5.4) and vehicles per incident (Figure 5.5), the number of casualties per incident, casualties per hour and vehicles per hour were also analysed.
- Vehicle accidents in Canada in 2014. The number of vehicles per incident, incidents per hour and vehicles per hour were analysed.
- USA gun violence data between 2014 and 2017. In addition to the number of injuries per incident (Figure 5.6) and fatalities per day (Figure 5.7), the number of fatalities per day and injuries per day were analysed.
- EPL game data between 2002 and 2016. In addition to the number of home team goals per game (Figure 5.8), the number of away team goals per game, home team shots on target per game and away team shots on target per game were analysed.
- Flight delay data for all flights in the USA from 2015. In addition to the arrival delay time (Figure 5.9), the departure delay, the time spent taxiing on takeoff and time spent taxiing on landing were analysed.

Of the 33 datasets that were analysed in Section 5.2, the weighted Poisson distribution outperformed the Poisson in every case.

Of the 17 datasets that were analysed in Section 5.3, the weighted Poisson outperformed the various alternative modelling methodologies 13 times, which is roughly 76%. In total, 45 different datasets were analysed. In Table 5.1 below, the prevalence of the various weight functions that gave the best fit is presented.

Weight function	Frequency	Percentage occurrence
$w(n;\phi) = (a)_n$	21	46.67%
$w(n;\phi) = \frac{\Gamma(r+n)}{n!\Gamma(r)} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	11	24.44%
$w(n;\phi) = n + \varepsilon$	4	8.89%
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	2	4.44%
$w(n;\phi) = (n+a)(n^2 - bn + c)$	2	4.44%
$w(n;\phi) = (n+a)(n-b)^2$	2	4.44%
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{n!\Gamma(r)} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	1	2.22%
$w(n;\phi) = (n)_a$	1	2.22%
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	1	2.22%

Table 5.1: Weight function occurrence

Chapter 6

Poisson Process: Goodness-of-fit Testing

Up to this point of the thesis, the focus has solely been on weighted Poisson distributions and no attention has been paid to how events occur as time progresses. This clearly limits the potential applications of the weighted Poisson framework. Consequently, chapters 6, 7 and 8 will extend the notions of goodness-of-fit (Chapter 3), the newly derived weighted Poisson distributions (Chapter 4) and their applications (Chapter 5) to stochastic processes.

The stochastic process extension of the Poisson distribution is called the Poisson process. This process has received considerable research attention, and as a result has many different applications (see Feller [48]), specifically in the actuarial industry (see Konstantinides [85], Rolski et al. [118] and Willmot and Lin [142]). This large amount of research has led to many different (equivalent) definitions of the Poisson process (definitions 10.25, 10.26 and 10.27). Using these definitions, a list of tests will be discussed that can indicate whether a dataset represents a sample path from a homogeneous Poisson process or not. These tests will be differentiated based on different types of information used in the underlying statistics.

6.1 Tests based on the number of events per interval

In the first category of tests, it is assumed that the number of events that occur in consecutive, non-overlapping, intervals is known, where $\{N_i\}, i = 1, 2, ..., k$, represents the number of events in each interval and $\{t_i\}, i = 1, 2, ..., k$, represents the length of each interval (where $\sum_{i=1}^{k} t_i = T$ is the entire length of time for which the process is observed).

6.1.1 Chi-square test

It is known that if the rate (λ) of a Poisson process does not change with respect to time that $N_i \sim Poi(\lambda t_i), i = 1, 2, ..., k$. Furthermore, the distribution of N_i, when the total number of events, $N = N_1 + N_2 + ... + N_k$, is known, follows a binomial distribution with parameters n and $p_i = \frac{t_i}{T}, i = 1, 2, ..., k$. Written formally: $P(N_i = n_i | N = n) = \binom{n}{n_i} \left(\frac{t_i}{T}\right)^{n_i} \left(1 - \frac{t_i}{T}\right)^{n-n_i}$

(See Ross [119]).

It then follows that the joint probability mass function of the number of events in the k different intervals follow a multinomial distribution: $P\left(N_1 = n_1, N_2 = n_2, ..., N_k = n_k\right) = \frac{n!}{n_1!n_2!...n_k!}\left(\frac{t_1}{T}\right)^{n_1}\left(\frac{t_2}{T}\right)^{n_2}...\left(\frac{t_k}{T}\right)^{n_k}$.

If the assumption is made that all $t_i N$ values are sufficiently large ("large" is usually assumed to be greater than or equal to 5), it is possible to construct a chi-square test statistic:

$$\sum_{i=1}^{k}\frac{\left(N_i - \frac{t_i}{T}N\right)^2}{\frac{t_i}{T}N},$$

which is $\chi^2\left(k-1\right)$ distributed under the null hypothesis that the data are realised from a homogeneous Poisson process.

6.1.2 Fisher index test

If a Poisson process is homogeneous, non-overlapping intervals are independent of each other. This leads to the ability to construct tests based on statistics from each individual increment $\left\{N_i\right\}, i = 1, 2, ..., k$. The Fisher index test for homogeneous Poisson processes is constructed in the following manner:

The mean, $\bar{N} = \sum_{i=1}^{k}\frac{N_i}{k}$, and variance, $S^2 = \frac{\sum_{i=1}^{k}\left(N_i - \bar{N}\right)^2}{k-1}$, of the number of the number of events is calculated, and the Fisher index is calculated, $FI\left(N_1, ..., N_k\right) = \frac{S^2}{\bar{N}}$. As Anderson and Siddiqui [3] discussed, this statistic can be modeled reasonably well by a $\chi^2\left(1\right)$ distribution. This distribution is, however, an approximation and not the exact distribution of $FI\left(N_1, ..., N_k\right)$. As a result Table 6.1 contains the 95% confidence intervals for $FI\left(N_1, ..., N_k\right)$ for various rates and number of intervals. The values chosen for k correspond to values that are likely to occur in practice like days in a week or minutes in an hour.

	$k = 8760$	365	60	52
$\lambda = 5$	$(0.9989; 1)$	$(0.9753; 1.2315)$	$(0.8644; 1.4576)$	$(0.8431; 1.4510)$
10	$(0.9982; 0.9997)$	$(0.9588; 1.2009)$	$(0.7966; 1.4224)$	$(0.7647; 1.4650)$
25	$(0.9961; 1.0659)$	$(0.9176; 1.1875)$	$(0.7119; 1.4100)$	$(0.6925; 1.4510)$
50	$(0.9929; 1.0406)$	$(0.8791; 1.1645)$	$(0.6942; 1.3898)$	$(0.6667; 1.4510)$
100	$(0.9869; 1.0368)$	$(0.8684; 1.1621)$	$(0.6787; 1.3927)$	$(0.6587; 1.4283)$
250	$(0.9736; 1.0331)$	$(0.8626; 1.1524)$	$(0.6728; 1.3975)$	$(0.6545; 1.4236)$
500	$(0.9726; 1.0312)$	$(0.8651; 1.149)$	$(0.6727; 1.3933)$	$(0.6508; 1.4276)$
1000	$(0.9716; 1.0308)$	$(0.8619; 1.1534)$	$(0.6712; 1.3828)$	$(0.6542; 1.4287)$

	$k = 24$	12	7
$\lambda = 5$	$(0.6957; 1.6335)$	$(0.5151; 2)$	$(0.3333; 2.25)$
10	$(0.6087; 1.6547)$	$(0.4126; 2)$	$(0.2; 2.33333)$
25	$(0.5409; 1.6522)$	$(0.3636; 2)$	$(0.2150; 2.3566)$
50	$(0.5217; 1.6530)$	$(0.3482; 2)$	$(0.2138; 2.4)$
100	$(0.5130; 1.6593)$	$(0.3550; 1.9802)$	$(0.2061; 2.3918)$
250	$(0.5109; 1.6633)$	$(0.3478; 1.9910)$	$(0.2028; 2.4027)$
500	$(0.5088; 1.6562)$	$(0.3436; 1.9822)$	$(0.2061; 2.3809)$
1000	$(0.5115; 1.6668)$	$(0.3435; 1.9890)$	$(0.2121; 2.3973)$

Table 6.1: Fisher index confidence intervals

6.1.3 Likelihood ratio test

Fierro and Tapia [49] proposed a test, which, in its simplest form, can detect if a process is a homogeneous Poisson process, or as an alternative, is a nonhomogeneous Poisson process with piece-wise constant intensity functions. Stated more formally, their test is

$$\begin{aligned} H_0 &: \quad \lambda_1 = \lambda_2 = ... = \lambda_k, \\ H_A &: \quad \lambda_i \neq \lambda_j \text{ for at least one } i \neq j, \end{aligned}$$

where λ_i is the constant intensity of events in interval N_i.

What makes their test particularly interesting is that (under a more complicated test formulation given in their paper) it is also possible to test whether certain parts of a process could be considered homogeneous, rather than simply considering the entire process as a whole. In addition to deriving an expression for the likelihood function, they also showed that this test is asymptotically optimal.

6.2 Tests based on the exact time of events

For the second class of test, assume that the times when individual events occur, $0 < \tau_1 < \tau_2 < ... < \tau_N$ is known. From this, a sequence of inter-arrival times, $\{T_i\}, i = 1, 2, ..., N$ can be constructed.

6.2.1 Uniformity tests

Lewis [90] discussed a series of tests based on an idea proposed by Barnard [9]. The tests are based on using the ratio between the time of events occurring and the length of the time interval considered as a sequence of statistics. The statistics are given by

$$U_i = \frac{\tau_i}{T}, i = 1, 2, ..., N,$$

where each U_i is i.i.d with a standard uniform distribution. Since these statistics are independent of λ, a range of distribution-free goodness-of-fit tests can be applied to them. In the paper Lewis [90] discusses using chi-square, Kolmogorov-Smirnov, Cramér-von Mises and Anderson-Darling tests on these statistics. He did note, however, that Durbin [41] proposed a transformation of the data which resulted in higher powers of the tests.

6.2.2 Exponentiality tests

Epstein [45] proposed 12 different tests that can be used to detect deviation from the exponential distribution. Although the paper was written in the context of life data, these tests are particularly applicable to Poisson processes as well, because, if a Poisson process is homogeneous, it follows that $T_i \sim Exp\left(\lambda\right), i = 1, 2, ..., N$ which are independent of each other. The formulation of the proposed tests varied greatly: there are two graphical tests, a chi-square test, a test based on the conditional distribution of the T_i values and the conditional rate of the process, tests to detect if T_1 occurs abnormally early or late, there are tests to detect if the mean of the variables fluctuate or experience consistent shifts, as well as a test that can detect if T_i values are abnormally large.

In addition to discussing some classical and powerful exponential tests that are available in the literature (similar to those in Chapter 3), Oosthuysen et al. [105] proposed a series of new tests based on the L^1, L^2 and L^∞ distances between the empirical mass function of the sample increments and the corresponding Poisson probability mass function. The tests were found to have comparable or improved performance over many other tests for exponentiality.

6.2.3 A test based on waiting times

It is a well known property of exponential distributions that the sum of independent exponential distributions is gamma distributed. It then follows that the sum of all of the inter-arrival times of a homogeneous Poisson process is gamma distributed: $T = \sum_{i=1}^{N} T_i \sim Gamma\left(N, \frac{1}{\lambda}\right)$. Using this result it is possible to construct 95% confidence intervals for T for various values of λ and N. (See Table 6.2 below.)

	$N = 5$	10	25	50
$\lambda = 5$	(0.3247; 2.0483)	(0.9591; 3.4170)	(3.2357; 7.1420)	(7.4221; 12.9561)
10	(0.1623; 1.0242)	(0.4795; 1.7085)	(1.6179; 3.5710)	(3.7111; 6.4781)
25	(0.0649; 0.4097)	(0.1918; 0.6834)	(0.6471; 1.4284)	(1.4844; 2.5912)
50	(0.0325; 0.2048)	(0.0959; 0.3417)	(0.3236; 0.7142)	(0.7422; 1.2956)
100	(0.0162; 0.1024)	(0.0480; 0.1708)	(0.1618; 0.3571)	(0.3711; 0.6478)
250	(0.0065; 0.0410)	(0.0192; 0.0683)	(0.0647; 0.1428)	(0.1484; 0.2591)
500	(0.0032; 0.0205)	(0.0096; 0.0342)	(0.0324; 0.0714)	(0.0742; 0.1296)
1000	(0.0016; 0.0102)	(0.0048; 0.0171)	(0.0162; 0.0357)	(0.0371; 0.0648)

	$N = 100$	250	500	1000
$\lambda = 5$	(16.2728; 24.1058)	(43.9936; 56.3852)	(91.4257; 108.953)	(187.795; 212.584)
10	(8.1364; 12.0529)	(21.9968; 28.1926)	(45.7129; 54.4765)	(93.8973; 106.292)
25	(3.2546; 4.8212)	(8.7987; 11.277)	(18.2851; 21.7906)	(37.5589; 42.5168)
50	(1.6273; 2.4106)	(4.3994; 5.6385)	(9.1426; 10.8953)	(18.7795; 21.2584)
100	(0.8136; 1.2053)	(2.1997; 2.8193)	(4.5713; 5.4477)	(9.3897; 10.6292)
250	(0.3255; 0.4821)	(0.8799; 1.1277)	(1.8285; 2.1791)	(3.7559; 4.2517)
500	(0.1627; 0.2411)	(0.4399; 0.5639)	(0.9143; 1.0895)	(1.8780; 2.1258)
1000	(0.0813; 0.1205)	(0.2199; 0.2819)	(0.4571; 0.5448)	(0.9390; 1.0629)

Table 6.2: Exponential waiting time test

6.2.4 Change point type test

By drawing on the literature from stochastic process control (SPC) it is possible to construct a change point type test to detect if the rate of a process changes. This test is useful since it not only detects if λ changes, but it can also give an indication as to when the change occurs.

Assume that before some time k, $0 < k < T$, the rate of the process is λ, and that the rate experiences a single sustained shift from λ to $a\lambda$ where $a \neq 1$ from time k onward.

While this test is explicitly set up to detect a consistent change after a specific point in time, the test is able to detect a gradual shift in the rate (although at a lower power).

The hypotheses of the proposed test is as follows:

$$\begin{array}{lll} H_0: & T_i \sim Exp(\lambda) & i = 1, 2, ..., N. \\ H_A: & T_i \sim Exp(\lambda) & i = 1, 2, ..., k-1. \\ & T_i \sim Exp(a\lambda) & i = k, ..., N. \end{array}$$

The difficulty in developing this test is that, not only is the size of shift in λ unknown, but the location where the shift occurs, k, is also unknown (although it is assumed to be deterministic.)

To estimate the point where the shift in the rate occurs a series of comparisons need to be made:

T_1	is compared to	$T_2, T_3, ..., T_N$,
T_1, T_2	is compared to	$T_3, T_4, ..., T_N$,
T_1, T_2, T_3	is compared to	$T_4, T_5, ..., T_N$,
	and so forth until	
$T_1, T_2, ..., T_{N-1}$	is compared to	T_N.

If there are N inter-arrival times there will be $N-1$ different comparisons made to determine whether, and, if applicable where the process experiences a change in its rate.

The series of statistics that make up the building blocks of the proposed test are given by

$$U_r = \frac{\frac{\sum_{i=r}^{N} T_i}{(N-r+1)}}{\frac{\sum_{i=1}^{r-1} T_i}{(r-1)}} = \frac{(r-1)\sum_{i=r}^{N} T_i}{(N-r+1)\sum_{i=1}^{r-1} T_i}, r =, 2, ..., N. \tag{6.1}$$

Under the null hypothesis of no shift having occurred, and by using the fact that $\sum_{i=1}^{r-1} T_i \sim Gamma\left(r-1, \frac{1}{\lambda}\right)$ and $\sum_{i=r}^{N} T_i \sim Gamma\left(N-r+1, \frac{1}{\lambda}\right)$ it follows that
$U_r \sim F\left(2\left(N-r+1\right), 2\left(r-1\right)\right), r = 2, 3, ..., N$.

The test statistic is then chosen to be $U = \min\{U_2, U_3, ..., U_N\}$ if $a > 1$ and $U = \max\{U_2, U_3, ..., U_N\}$ if $0 < a < 1$. With the reasoning for this choice given as follows:
Suppose that an increase in the process rate does indeed occur at time k, then:

- The statistic U_k's numerator will only contain values from $Exp\left(\lambda a\right)$ distributions, whereas the denominator will only contain values that come from $Exp\left(\lambda\right)$ distributions.

- If b is some integer value such that $2 \leq b < k$, then statistic U_b will contain $k-b$ values in its numerator that are from $Exp\left(\lambda\right)$ distributions. This will increase the average of the data in U_b's numerator in comparison to the numerator of U_k.

- Similarly, if c is some integer value such that $k < c \leq N$, then statistic U_c will contain c values in its denominator that are from a $Exp\left(\lambda a\right)$ distribution. This will decrease the weighted average of the data in U_c's denominator in comparison to the denominator of U_k.

- Thus, any statistic other than the one immediately following the shift in the process rate, will contain either larger (on average) observations in its numerator, or smaller (on average) observations in its denominator. Either of these scenarios result in all other statistics being larger relative to U_k (assuming that the the denominator is greater than 1).

- This leads to the conclusion that the most probable place where a shift in the process rate will be detected is at the statistic immediately following the shift. The value that this statistic assumes also has a high likelihood of being the minimum value of all the $U_r, r = 2, 3, ..., N$ statistics.

- As such, the most reasonable method of calculating the critical value (to detect an upwards shift in the process rate) is to calculate the minimum order statistic of $U_r, r = 2, 3, ..., N$, (under the null hypothesis) and to set the critical value equal to some percentile of the distribution of the minimum order statistic.

Using a similar but inverted argument, it can be justified that the critical value of the test should be set equal to some percentile of the maximum order statistic of $U_r, r = 2, 3, ..., N$, under the null hypothesis of no shift having occurred, if the detection of a downward shift in the process rate is of concern.

As a result, by calculating the minimum and maximum order statistics of $U_r, r = 2, 3, ..., N$, both a decrease or an increase in λ, as well as the location where the shift occurred, k, can be estimated.

To demonstrate this test, suppose that 20 events occurred and that $\lambda = 100$. After the 10th event the rate doubles to $\lambda = 200$. Table 6.3 below gives a simulated set of the data and statistics for this test. A plot of the statistics is given in Figure 6.1.

Waiting time (T_i)			Statistic (U_r)
1	0.00399765		
2	0.00956458	2	2.29147
3	0.00449458	3	1.34757
4	0.010561	4	1.5636
5	0.00511407	5	1.30539
6	0.00864347	6	1.4261
7	0.0211931	7	1.37214
8	0.0111455	8	0.9697
9	0.0256652	9	0.922033
10	0.013708	10	0.633064
11	0.00358537	11	0.560623
12	0.00726844	12	0.627089
13	0.00291053	13	0.637575
14	0.00667508	14	0.729131
15	0.00838265	15	0.754854
16	0.00626938	16	0.737627
17	0.00361156	17	0.774066
18	0.00061888	18	0.936728
19	0.00611749	19	1.44544
20	0.0185206	20	2.20585

Table 6.3: Example - Change point method

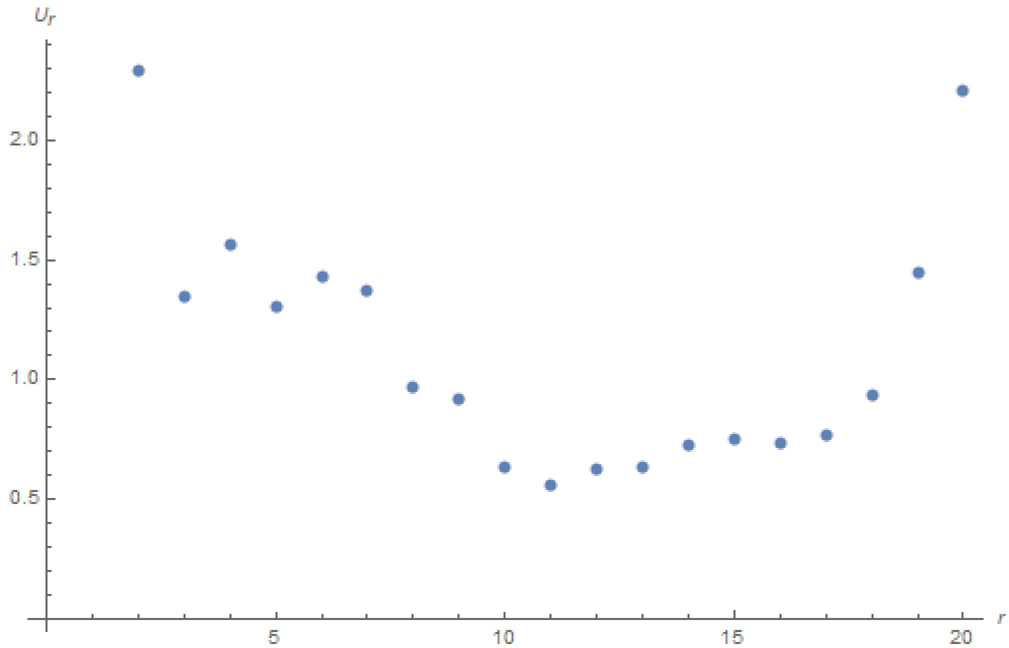


Figure 6.1: Example - Change point method

As can be seen, the minimum of the set of statistics occurs immediately after the change in the rate of the process.

It should be noted that this test has some inherent shortcomings, like the masking of the location where the shift in the rate occurs, as well as the potential for large/small statistics to occur at the endpoints of the series (making this a poor test for detecting rate shifts either early or late in the process, or for very short processes). The relative pros and cons of a similar test, along with a much lengthier theoretical and practical discussion can be found in Mijburgh et al. [95]. Additionally, Balakrishnan et al. [5] proposed a series of tests based on the same change point type setup.

In tables 6.4 and 6.5 below, the 2.5% and 97.5% quantiles of the minimum and maximum order statistics of $U_r, r = 2, 3, ..., N$ will be given. The observed events will be varied, but the rate will not, since under the null hypothesis the tests are independent of the rate. When calculating the order statistics, the first and last 10% of the U_r statistics will be excluded to stabilise the critical values of the test.

	$N = 5$	10	25	50	100	250	500	1000
$Min\,(U_r)_{0.025}$	0.02443	0.02498	0.18571	0.29379	0.43178	0.59166	0.68983	0.76715

Table 6.4: U_r Minimum order statistic - 2.5% quantile

	$N = 5$	10	25	50	100	250	500	1000
$Max\,(U_r)_{0.975}$	41.1653	9.5407	5.4009	3.1084	2.2575	1.6796	1.4491	1.3016

Table 6.5: U_r Maximum order statistic - 97.5% quantile

6.3 Test comparison

In this section, the performance of the four tests (Chi-square, Fisher index, sum of waiting times and change point) that have not received a power analysis in the above mentioned papers will be compared against a range of alternative stochastic processes. The alternate processes have various different parameterisations, but will all have an expected number of events which corresponds to the Poisson process under the null hypothesis. The first group of alternate processes are all nonhomogeneous Poisson processes, where λ is a function of t. The expression relating t to λ will be varied for each alternate process. For the second group of processes, the distribution of the inter-arrival times will be varied. Tables 6.6, 6.7 and 6.8 contain powers in percentages.

	Chi-square	Fisher index	Waiting times	Change point
$Poisson\,(\lambda = 100)$	5	5	5	5
$Poisson\,(\lambda = 95 + 10t)$	5.01	5.07	3.98	5.07
$Poisson\,(\lambda = 90 + 20t)$	5.10	5.33	3.32	5.97
$Poisson\,(\lambda = 85 + 30t)$	5.25	5.80	2.55	7.43
$Poisson\,(\lambda = 80 + 40t)$	5.44	6.48	2.01	9.66
$Poisson\,(\lambda = 75 + 50t)$	5.77	7.56	1.60	12.67
$Poisson\,(\lambda = 70 + 60t)$	6.47	9.10	1.19	16.72
$Poisson\,(\lambda = 65 + 70t)$	6.73	9.66	1.57	18.01
$Poisson\,(\lambda = 60 + 80t)$	8.82	13.22	0.72	29.03
$Poisson\,(\lambda = 55 + 90t)$	10.74	16.22	0.51	37.22
$Poisson\,(\lambda = 50 + 100t)$	12.87	19.6	0.04	46.40
$Poisson\,(\lambda = 45 + 110t)$	16.37	24.09	0.03	56.40
$Poisson\,(\lambda = 40 + 120t)$	20.32	29.21	0.02	66.79
$Poisson\,(\lambda = 35 + 130t)$	24.94	34.84	0.00	76.08
$Poisson\,(\lambda = 30 + 140t)$	30.72	41.59	0.00	84.01
$Poisson\,(\lambda = 25 + 150t)$	37.36	48.77	0.00	90.36
$Poisson\,(\lambda = 20 + 160t)$	44.68	56.30	0.00	94.96
$Poisson\,(\lambda = 15 + 170t)$	52.83	64.32	0.00	97.67
$Poisson\,(\lambda = 10 + 180t)$	60.55	71.41	0.00	99.03
$Poisson\,(\lambda = 5 + 190t)$	68.73	78.54	0.00	99.69
$Poisson\,(\lambda = 200t)$	76.10	84.50	0.00	99.93

Table 6.6: Power of homogeneous Poisson tests 1, $k = 52, N = 100$

	Chi-square	Fisher index	Waiting times	Change point
$Poisson\,(\lambda = 100)$	5	5	5	5
$Poisson\,(\lambda = 105 - 10t)$	5.07	5	6.51	5.37
$Poisson\,(\lambda = 110 - 20t)$	5.13	5.34	7.96	6.37
$Poisson\,(\lambda = 115 - 30t)$	5.15	5.73	9.68	8.20
$Poisson\,(\lambda = 120 - 40t)$	5.60	6.68	11.93	11.0
$Poisson\,(\lambda = 125 - 50t)$	5.85	7.59	14.55	14.98
$Poisson\,(\lambda = 130 - 60t)$	6.43	8.94	17.10	21.0
$Poisson\,(\lambda = 135 - 70t)$	7.34	10.78	20.41	28.56
$Poisson\,(\lambda = 140 - 80t)$	8.73	13.09	23.11	37.66
$Poisson\,(\lambda = 145 - 90t)$	10.46	16.04	25.47	45.56
$Poisson\,(\lambda = 150 - 100t)$	12.96	19.52	26.37	52.34
$Poisson\,(\lambda = 155 - 110t)$	16.22	24.10	27.35	57.71
$Poisson\,(\lambda = 160 - 120t)$	20.02	28.97	28.80	61.63
$Poisson\,(\lambda = 165 - 130t)$	25.13	35.03	31.59	64.63
$Poisson\,(\lambda = 170 - 140t)$	30.96	41.64	36.63	66.67
$Poisson\,(\lambda = 175 - 150t)$	39.44	50.84	42.74	68.38
$Poisson\,(\lambda = 180 - 160t)$	44.76	56.38	49.48	69.22
$Poisson\,(\lambda = 185 - 170t)$	52.69	64.10	55.86	70.10
$Poisson\,(\lambda = 190 - 180t)$	60.76	71.58	62.99	70.37
$Poisson\,(\lambda = 195 - 190t)$	68.69	78.48	69.01	70.52
$Poisson\,(\lambda = 200 - 200t)$	75.97	84.34	75.01	70.91

Table 6.7: Power of homogeneous Poisson tests 2, $k = 52, N = 100$

	Chi-square	Fisher index	Waiting times	Change point
$T_i \sim Gamma\left(0.5, \frac{1}{50}\right)$	83.20	89.27	16.78	27.75
$T_i \sim Gamma\left(0.9, \frac{1}{90}\right)$	7.73	11.82	6.41	6.98
$T_i \sim Exp\,(100)$	5	5	5	5
$T_i \sim Gamma\left(1.1, \frac{1}{110}\right)$	6.46	1.86	4.08	3.53
$T_i \sim Gamma\left(1.5, \frac{1}{150}\right)$	35.66	0.17	1.75	0.99
$T_i \sim Gamma\left(1.75, \frac{1}{175}\right)$	60.74	0	0.98	0.54
$T_i \sim Gamma\left(2, \frac{1}{200}\right)$	79.71	0	0.60	0.23
$T_i \sim Beta\,(0.5, 49.5)$	81.58	88	16.23	26.64
$T_i \sim Beta\,(0.75, 74.25)$	22.38	31.75	8.95	10.82
$T_i \sim Beta\,(1, 99)$	4.95	4.32	4.88	4.71
$T_i \sim Beta\,(1.25, 123.75)$	12.03	0.30	2.76	2.07
$T_i \sim Beta\,(1.5, 148.5)$	37.58	0.01	1.61	0.96

Table 6.8: Power of homogeneous Poisson tests 3, $k = 52, N = 100$

Based on the reported powers in the above tables it appears as if the change point test performs particularly well when the rate of the process varies in a linear manner, but the chi-square test outperforms it if the distribution of the inter-arrival times are varied.

Chapter 7

Weighted Poisson Process: Theory

In this chapter the expansion of weighted Poisson distributions into weighted Poisson processes will be discussed.

Balakrishnan and Kozubowski [6] extended the idea of weighted Poisson distributions to weighted Poisson processes, and derived many of the statistical properties associated with weighted Poisson processes.

Traditionally, Poisson processes are defined either as in Definition 10.25 or Definition 10.26; however, Balakrishnan and Kozubowski [6] used the notion of compound Poisson processes (see Definition 10.28) to define a Poisson process as a special case of the compound Poisson process as follows:

Definition 7.1.

$N(t)$ is said to be a Poisson process with rate $\lambda > 0$, denoted by $\{N(t), t \in [0,1]\}$, if

$$N(t) \overset{d}{=} \sum_{j=1}^{N} I_{[0,t]}(U_j), t \in [0,1],$$

where N is a Poisson random variable with rate λ and U_j are independent standard uniform random variables, which are also independent of N.

Using this definition of a Poisson process, rather than definitions 10.25 or 10.26, gives a convenient and simple way to define a weighted Poisson process. By replacing the Poisson random variable in the upper summation limit with a weighted Poisson version, the stochastic process becomes a weighted Poisson process. Stated more formally:

Definition 7.2.

$N^w(t)$ is said to be a weighted Poisson process with intensity λ and weight function $w(\cdot)$, on the interval $[0,1]$, denoted by $\{N^w(t), t \in [0,1]\}$, if

$$N^{w}(t) \overset{d}{=} \sum_{j=1}^{N^{w}} I_{[0,t]}(U_j), t \in [0,1], \tag{7.1}$$

where N^w is a weighted Poisson random variable with rate λ and weight function $w(\cdot)$, and U_j are standard uniform random variables independent of each other and of N^w.

It should be noted that earlier t was defined on the interval $[0, T]$, and here the interval has been altered to be $[0, 1]$. This apparent change has no real practical impact since the time scale and rate of the process can be redefined as needed. For example, say a process/dataset spans 200 days then, instead of each daily increment being represented by an increase of 1 in t, each daily increment could be represented by an increase of $\frac{1}{200}$ in t with the upper bound set equal to 1. These two situation, after accounting for the change in the process's rate, are equivalent.

The reason for this change in t is that the $I_{[0,t]}(U_j)$ in the above equation are independent and identically distributed, and if $t \in [0, 1]$ each $I_{[0,t]}(U_j)$ is a Bernoulli random variable with parameter t. Thus, the stochastic process observed at a specific point in time can in fact be seen to be a compound random variable, the properties of which are well established in the literature. Consequently, the statistical properties of the weighted Poisson process are simple to determine, assuming that the relevant properties of the corresponding weighted Poisson distribution can be determined.

Additionally, Balakrishnan and Kozubowski [6] reached the following conclusions regarding weighted Poisson processes:

- For each $t > 0$, the random variable $N^w(t)$ has a weighted Poisson distribution with rate parameter λt, and weight function $w_t(n) = \sum_{k=0}^{\infty} \frac{[(1-t)\lambda]^k}{k!} w(n+k), n = 0, 1, 2, ...$ (where t is included in the subscript of the weight function to indicate the reliance of the function on t).

- The dispersion of the entire stochastic process is determined by the dispersion of the related weighted Poisson random variable. This fact is given by the linear equation

  $$\frac{Var(N^w(t))}{E(N^w(t))} = 1 + t\left(\frac{Var(N^w)}{E(N^w)} - 1\right).$$

  In essence, if the dispersion of the weighted Poisson random variable is known, it will also be known for the associated stochastic process for each specific value of t.

- Consecutive, non-overlapping increments of weighted Poisson processes are positively correlated if the weighted Poisson random variable is overdispersed, and negatively correlated if the random variable is underdispersed. This is in contrast to the Poisson processes where the increments are not correlated. It should be noted, however, that the unconditional distributions of individual increments are stationary.

- The direct link between the weighted Poisson process and distribution results in the fact that closed form expressions exist for many of the statistical properties of the weighted Poisson process if the corresponding weighted Poisson distribution properties can be found.

Using the results from Balakrishnan and Kozubowski [6] the weighted Poisson distributions from Chapter 4 will be expanded into weighted Poisson processes. For each weighted Poisson process the following properties will be listed in the sections below:

- The weight function, $w_t(n;\phi)$, for $t \in [0,1]$.
- The probability mass function, $f_{w,t}(n) = P(N^w(t) = n), n = 0, 1, 2, ...$
- The probability generating function, $g_t(z)$.
- The expected value, $E(N^w(t))$.
- The variance, $Var(N^w(t))$.
- The probability generating function of the increments of the process, $g_{t-s}(z)$, where $0 \leq s < t \leq 1$.
- The joint probability mass function $f_{w,(s,t)}(a,b) = P(N^w(s) = a, N^w(t) = b)$, where $0 < a \leq b$.
- The covariance between $N^w(s)$ and $N^w(t), Cov(N^w(s), N^w(t)) = Cov_{s,t}$.
- The covariance between two consecutive increments $N^w(s)$ and $N^w(t) - N^w(s)$, $Cov(N^w(s), N^w(t) - N^w(s)) = Cov_{s,t-s}$

Derivations of the various results will not be provided, since their proofs are relatively simple to obtain after taking the results of Balakrishnan and Kozubowski [6] into account, and by applying similar methods to those demonstrated in Chapter 4.

7.1 Polynomial weight functions

7.1.1 $w(n) = n$

Theorem 7.1. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = n$ *then*

$$
\begin{aligned}
w_t(n) &= e^{\lambda(1-t)}\left(n+\lambda(1-t)\right). \\
f_{w,t}(n) &= \tfrac{e^{-\lambda t}(\lambda t)^n}{n!}\left(\tfrac{n}{\lambda}+1-t\right). \\
g_t(z) &= e^{\lambda t(z-1)}\left(1+t(z-1)\right). \\
E\left(N^w(t)\right) &= t(\lambda+1). \\
Var\left(N^w(t)\right) &= t(\lambda+1-t). \\
g_{t-s}(z) &= e^{\lambda(t-s)(z-1)}\left(1+s(1-z)+t(z-1)\right). \\
f_{w,(s,t)}(a,b) &= \begin{pmatrix} b \\ a \end{pmatrix} \tfrac{e^{-\lambda t}\left(1-\frac{s}{t}\right)^{b-a}\left(\frac{s}{t}\right)^a t^b \lambda^{b-1}(b+\lambda(1-t))}{b!}. \\
Cov_{s,t} &= s(1-t+\lambda). \\
Cov_{s,t-s} &= s(s-t).
\end{aligned}
$$

7.1.2 $w(n;\phi) = n^{-a}$

Theorem 7.2. *If the weight function used in the weighted Poisson probability mass function is chosen as $w(n;\phi) = n^{-a}$ then*

$$
\begin{aligned}
w_t(n;\phi) &= \frac{{}_aF_a(n,\ldots,n;1+n,\ldots,1+n;\lambda(1-t))}{n^a}.\\
f_{w,t}(n) &= \frac{t^n\lambda^{n-1}\Gamma(n)^{a-1}{}_aF_a(n,\ldots,n;1+n,\ldots,1+n;\lambda(1-t))}{n\Gamma(n+1)^a{}_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)}.\\
g_t(z) &= \frac{(1+t(z-1))_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda(1+t(z-1)))}{{}_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)}.\\
E(N^w(t)) &= \frac{t_aF_a(1,\ldots,1;2,\ldots,2;\lambda)}{{}_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)}.\\
Var(N^w(t)) &= \frac{t^2({}_{a-1}F_{a-1}(1,\ldots,1;2,\ldots,2;\lambda)-{}_aF_a(1,\ldots,1;2,\ldots,2;\lambda))}{{}_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)}\\
&+ E(N^w(t)) - (E(N^w(t)))^2.\\
g_{t-s}(z) &= \frac{(1+(t-s)(z-1))_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;(1+(t-s)(z-1))\lambda)}{{}_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)}.\\
f_{w,(s,t)}(a,b) &= \begin{pmatrix} b \\ a \end{pmatrix}\frac{(1-\frac{s}{t})^{b-a}(\frac{s}{t})^a t^b\lambda^{b-1}\Gamma(b)^{a-1}{}_aF_a(b,\ldots,b;1+b,\ldots,1+b;\lambda(1-t))}{b\Gamma(1+b)^a{}_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)}.\\
Cov_{s,t} &= -\frac{s(-t_{a-1}F_{a-1}(1,\ldots,1;2,\ldots,2;\lambda)+(t-1)_aF_a(1,\ldots,1;2,\ldots,2;\lambda))_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)}{{}_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)^2}\\
&- \frac{st_aF_a(1,\ldots,1;2,\ldots,2;\lambda)^2}{{}_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)^2}.\\
Cov_{s,t-s} &= \frac{s(s-t)({}_aF_a(1,\ldots,1;2,\ldots,2;\lambda)-{}_{a-1}F_{a-1}(1,\ldots,1;2,\ldots,2;\lambda))_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)}{{}_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)^2}\\
&+ \frac{s(s-t)_aF_a(1,\ldots,1;2,\ldots,2;\lambda)^2}{{}_{a+1}F_{a+1}(1,\ldots,1;2,\ldots,2;\lambda)^2}.
\end{aligned}
$$

7.1.3 $w(n;\phi) = n + \varepsilon$

Theorem 7.3. *If the weight function used in the weighted Poisson probability mass function is chosen as $w(n;\phi) = n + \varepsilon$ then*

$$
\begin{aligned}
w_t(n;\phi) &= e^{\lambda(1-t)}\left(n+\varepsilon+\lambda(1-t)\right). \\
f_{w,t}(n) &= \tfrac{e^{-\lambda t}(\lambda t)^n}{n!}\tfrac{n+\varepsilon+\lambda(1-t)}{\varepsilon+\lambda}. \\
g_t(z) &= \tfrac{e^{\lambda t(z-1)}(\varepsilon+\lambda(1+t(z-1)))}{\varepsilon+\lambda}. \\
E\left(N^w(t)\right) &= \tfrac{t\lambda(1+\varepsilon+\lambda)}{\varepsilon+\lambda}. \\
Var\left(N^w(t)\right) &= \tfrac{(t\lambda)^2(2+\varepsilon+\lambda)}{\varepsilon+\lambda} \\
&+ E\left(N^w(t)\right) - \left(E\left(N^w(t)\right)\right)^2. \\
g_{t-s}(z) &= \tfrac{e^{\lambda(t-s)(z-1)}(\varepsilon+\lambda(1+s(1-z)+t(z-1)))}{\varepsilon+\lambda}. \\
f_{w,(s,t)}(a,b) &= \begin{pmatrix} b \\ a \end{pmatrix} \tfrac{e^{-\lambda t}\left(1-\frac{s}{t}\right)^{b-a}\left(\frac{s}{t}\right)^a(t\lambda)^b(b+\varepsilon+\lambda(1-t))}{b!(\varepsilon+\lambda)}. \\
Cov_{s,t} &= \tfrac{s\lambda\left(\varepsilon+\varepsilon^2+2\varepsilon\lambda+\lambda(1-t+\lambda)\right)}{(\varepsilon+\lambda)^2}. \\
Cov_{s,t-s} &= \tfrac{s(s-t)\lambda^2}{(\varepsilon+\lambda)^2}.
\end{aligned}
$$

7.1.4 $w(n;\phi) = an^3 + bn^2 + cn$

Theorem 7.4. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = an^3 + bn^2 + cn$ *then*

$$\begin{aligned}
w_t(n;\phi) &= e^{\lambda(1-t)}\left(\lambda(1-t)\left(b+c+2bn+a\left(1+3n\left(1+n\right)\right)\right)\right) \\
&+ e^{\lambda(1-t)} n\left(c+n\left(b+an\right)\right) \\
&+ e^{\lambda(1-t)}\left(\lambda^2(t-1)^2\left(b+3a\left(1+n\right)\right)-a\lambda^3(t-1)^3\right).
\end{aligned}$$

$$f_{w,t}(n) = \frac{e^{-\lambda t}(n+1)t^n\lambda^{n-1}}{(b(1+\lambda)+c+a(1+\lambda(\lambda+3)))\Gamma(n+2)}\frac{w_t(n)}{e^{\lambda(1-t)}}.$$

$$g_t(z) = \frac{e^{t\lambda(z-1)}(1+t(z-1))\left(a+b+c+\lambda(3a+b)(1+t(z-1))+a(1+t(z-1))^2\lambda^2\right)}{b(1+\lambda)+c+a(1+\lambda(\lambda+3))}.$$

$$E\left(N^w(t)\right) = \frac{t\left(b+c+3b\lambda+c\lambda+b\lambda^2+a\left(1+7\lambda+6\lambda^2+\lambda^3\right)\right)}{b+c+b\lambda+a(1+3\lambda+\lambda^2)}.$$

$$\begin{aligned}
Var\left(N^w(t)\right) &= \frac{t^2\lambda\left(c(2+\lambda)+b\left(4+5\lambda+\lambda^2\right)+a\left(8+19\lambda+9\lambda^2+\lambda^3\right)\right)}{b+c+b\lambda+a(1+3\lambda+\lambda^2)} \\
&+ E\left(N^w(t)\right)-\left(E\left(N^w(t)\right)\right)^2.
\end{aligned}$$

$$\begin{aligned}
g_{t-s}(z) &= \frac{e^{\lambda(t-s)(z-1)}(1+s(1-z)+t(z-1))}{b+c+b\lambda+a(1+\lambda(3+\lambda))} \\
&\times \frac{\left(a+b+c+\lambda(3a+b)(1+s(1-z)+t(z-1))+a(1+s(1-z)+t(z-1))^2\lambda^2\right)}{b+c+b\lambda+a(1+\lambda(3+\lambda))}.
\end{aligned}$$

$$\begin{aligned}
f_{w,(s,t)}(a,b) &= \binom{b}{a}\frac{(1+b)e^{-\lambda t}\left(1-\frac{s}{t}\right)^{b-a}\left(\frac{s}{t}\right)^a t^b\lambda^{b-1}}{(b+c+b\lambda+a(1+\lambda(3+\lambda)))\Gamma(b+2)} \\
&\times \left((1+a)b^3+bc-a(t-1)^3\lambda^3\lambda(t-1)\right. \\
&+ \left.(a+b+3ab+(2+3a)b^2+c)+(b+3a(1+b))(t-1)^3\lambda^2\right).
\end{aligned}$$

$$\begin{aligned}
Cov_{s,t} &= s\frac{(1-t)(b+c+b\lambda+a(1+\lambda(3+\lambda)))(b+c+\lambda(c+b(3+\lambda))+a(1+\lambda(7+\lambda(6+\lambda))))}{(b+c+b\lambda+a(1+\lambda(3+\lambda)))^2} \\
&\times \left(s\frac{t\lambda\left(c^2+bc(3+2\lambda)+b^2(2+\lambda(2+\lambda))+a^2(4+\lambda(10+\lambda(14+\lambda(6+\lambda))))\right)}{(b+c+b\lambda+a(1+\lambda(3+\lambda)))^2}\right. \\
&+ \left.s\frac{t\lambda(a(c(5+2\lambda(5+\lambda))+b(6+\lambda(12+\lambda(9+2\lambda)))))}{(b+c+b\lambda+a(1+\lambda(3+\lambda)))^2}\right).
\end{aligned}$$

$$Cov_{s,t-s} = \frac{s(s-t)\left((a+b+c)^2+2(3a+b)(a+b+c)\lambda+2(3a+b)^2\lambda^2+4a(3a+b)\lambda^3+3a^2\lambda^4\right)}{(b+c+b\lambda+a(1+\lambda(3+\lambda)))^2}.$$

7.1.5 $w(n;\phi) = (n+a)(n-b)^2$

Theorem 7.5. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = (n+a)(n-b)^2$ *then*

$$\begin{aligned}
w_t(n;\phi) &= e^{\lambda(1-t)}\left((b-n)^2(a+n) + (3+a-2b+3n)(t-1)^2\lambda^2\right)\\
&+ e^{\lambda(1-t)}\left(\lambda(t-1)\left(1+b^2+3n(1+n)\right) - (t-1)^3\lambda^3\right)\\
&+ e^{\lambda(1-t)}\left(\lambda(t-1)\left(a(1-2b+2n) - 2b(1+2n)\right)\right).
\end{aligned}$$

$$f_{w,t}(n) = \frac{e^{-\lambda t}(n+1)t^n\lambda^n}{\left(a\left((b-\lambda)^2+\lambda\right)+\lambda(1+b^2-2b(1+\lambda)+\lambda(3+\lambda))\right)\Gamma(n+2)}\frac{w_t(n;\phi)}{e^{\lambda(1-t)}}.$$

$$\begin{aligned}
g_t(z) &= \frac{e^{t\lambda(z-1)}\left(ab^2+\lambda\left(a+(b-1)^2-2ab\right)(1+t(z-1))\right)}{a\left((b-\lambda)^2+\lambda\right)+\lambda(1+b^2-2b(1+\lambda)+\lambda(3+\lambda))}\\
&+ \frac{e^{t\lambda(z-1)}\left((3+a-2b)(1+t(z-1))^2\lambda^2+(1+t(z-1))^3\lambda^3\right)}{a\left((b-\lambda)^2+\lambda\right)+\lambda(1+b^2-2b(1+\lambda)+\lambda(3+\lambda))}.
\end{aligned}$$

$$E\left(N^w(t)\right) = \frac{t\lambda\left(1+7\lambda+6\lambda^2+\lambda^3+b^2(1+\lambda)-2b\left(1+3\lambda+\lambda^2\right)+a\left(1+b^2+3\lambda+\lambda^2-2b(1+\lambda)\right)\right)}{a(b^2+\lambda-2b\lambda+\lambda^2)+\lambda(1+b^2+3\lambda+\lambda^2-2b(1+\lambda))}.$$

$$\begin{aligned}
Var\left(N^w(t)\right) &= \frac{(t\lambda)^2\left(8+19\lambda+9\lambda^2+\lambda^3+b^2(2+\lambda)-2b\left(4+5\lambda+\lambda^2\right)+a\left(4+b^2+5\lambda+\lambda^2-2b(2+\lambda)\right)\right)}{a(b^2+\lambda-2b\lambda+\lambda^2)+\lambda(1+b^2+3\lambda+\lambda^2-2b(1+\lambda))}\\
&+ E\left(N^w(t)\right) - \left(E\left(N^w(t)\right)\right)^2.
\end{aligned}$$

$$\begin{aligned}
g_{t-s}(z) &= \frac{e^{\lambda(t-s)(z-1)}\left(\left(a+(b-1)^2-2ab\right)\lambda(1+s(1-z)+t(z-1))+\lambda^3(1+s(1-z)+t(z-1))^3\right)}{a\left((b-\lambda)^2+\lambda\right)+\lambda(1+b^2-2b(1+\lambda)+\lambda(3+\lambda))}\\
&\times \frac{e^{\lambda(t-s)(z-1)}\left(ab^2+\lambda^2(1+s(1-z)+t(z-1))^2(3+a-2b)\right)}{a\left((b-\lambda)^2+\lambda\right)+\lambda(1+b^2-2b(1+\lambda)+\lambda(3+\lambda))}.
\end{aligned}$$

$$\begin{aligned}
f_{w,(s,t)}(a,b) &= \binom{b}{a}\frac{(1+b)e^{-\lambda t}\left(1-\frac{s}{t}\right)^{b-a}\left(\frac{s}{t}\right)^a(\lambda t)^b}{\Gamma(b+2)\left(a\left((b-\lambda)^2+\lambda\right)+\lambda(1+b^2-2b(1+\lambda)+\lambda(3+\lambda))\right)}\\
&\times \left(\lambda(1+a+b)(1-t)+\lambda^2(3+a+b)(t-1)^2-\lambda^3(t-1)^3\right).
\end{aligned}$$

$$
\begin{aligned}
Cov_{s,t} \quad = \quad & \frac{s\lambda\left(a(1+a)(b-1)^2b^2+\lambda 2a\left(1+(b-4)(b-1)^2b-t-2(b-2)bt\right)\right)}{\left(a\left((b-\lambda)^2+\lambda\right)+\lambda(1+b^2-2b(1+\lambda)+\lambda(3+\lambda))\right)^2} \\
- \quad & \frac{s\lambda^2\left((t-1)(b-1)^4+a^2\left(4(b-1)^2b+t+2(b-2)bt-1\right)\right)}{\left(a\left((b-\lambda)^2+\lambda\right)+\lambda(1+b^2-2b(1+\lambda)+\lambda(3+\lambda))\right)^2} \\
+ \quad & \frac{s\lambda^3\left(2a^2(2-t+b(3b+2t-5))+(b-1)^2(10-6t+b(b-8+4t))\right)}{\left(a\left((b-\lambda)^2+\lambda\right)+\lambda(1+b^2-2b(1+\lambda)+\lambda(3+\lambda))\right)^2} \\
- \quad & \frac{s\lambda^3\left(2a\left(4t-7+b\left(19-17b+4b^2+2t(b-5)\right)\right)\right)}{\left(a\left((b-\lambda)^2+\lambda\right)+\lambda(1+b^2-2b(1+\lambda)+\lambda(3+\lambda))\right)^2} \\
+ \quad & \frac{2s\lambda^4\left(2a(6+a-18b)-b\left(24+2a^2-13b-6ab^2+2b^2\right)-t(3+a-2b)^2+14\right)}{\left(a\left((b-\lambda)^2+\lambda\right)+\lambda(1+b^2-2b(1+\lambda)+\lambda(3+\lambda))\right)^2} \\
+ \quad & \frac{s\lambda^5\left(16+13a+a^2-26b-8ab+6b^2-4(3+a-2b)t\right)}{\left(a\left((b-\lambda)^2+\lambda\right)+\lambda(1+b^2-2b(1+\lambda)+\lambda(3+\lambda))\right)^2} \\
+ \quad & \frac{s\lambda^6(9+2a-4b-3t)+s\lambda^7}{\left(a\left((b-\lambda)^2+\lambda\right)+\lambda(1+b^2-2b(1+\lambda)+\lambda(3+\lambda))\right)^2}.
\end{aligned}
$$

$$
\begin{aligned}
Cov_{s,t-s} \quad = \quad & \frac{s(s-t)\lambda^2\left((b-1)^4-\lambda 2(b-1)^2(2b-3)+2(3-2b)^2\lambda^2+4(3-2b)\lambda^3\right)}{\left(a\left((b-\lambda)^2+\lambda\right)+\lambda(1+b^2-2b(1+b)+\lambda(3+\lambda))\right)^2} \\
+ \quad & \frac{s(s-t)\lambda^2 2a\left(\left(1+2b^2(1+\lambda)+2\lambda(1+\lambda)(2+\lambda)-2b(2+\lambda)(1+2\lambda)\right)\right)}{\left(a\left((b-\lambda)^2+\lambda\right)+\lambda(1+b^2-2b(1+b)+\lambda(3+\lambda))\right)^2} \\
+ \quad & \frac{s(s-t)\lambda^2\left(3\lambda^4+a^2\left(1+2b^2-4b(1+\lambda)+2\lambda(1+\lambda)\right)\right)}{\left(a\left((b-\lambda)^2+\lambda\right)+\lambda(1+b^2-2b(1+b)+\lambda(3+\lambda))\right)^2}.
\end{aligned}
$$

7.1.6 $w(n;\phi) = (n+a)\left(n^2 - bn + c\right)$

Theorem 7.6. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = (n+a)(n^2 - bn + c)$ *then*

$$
\begin{aligned}
w_t(n;\phi) &= e^{\lambda(1-t)}\left((a+n)(c+n(n-b))\right) \\
&+ e^{\lambda(1-t)}\lambda\left((t-1)(1+c+3n(1+n))\right) \\
&+ e^{\lambda(1-t)}\lambda\left((t-1)(a(1-b+2n)-b(1+2n))\right) \\
&+ e^{\lambda(1-t)}\lambda^2\left((t-1)^2(3+a-b+3n)\right) \\
&- e^{\lambda(1-t)}\lambda^3\left((t-1)^3\right).
\end{aligned}
$$

$$
\begin{aligned}
f_{w,t}(n) &= \frac{e^{-\lambda t}(1+n)(\lambda t)^n(a+n)(c+n(n-b))}{(\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda))+a(c+\lambda(1-b+\lambda)))\Gamma(2+n)} \\
&- \frac{e^{-\lambda t}(1+n)(\lambda t)^n\lambda(t-1)(1+c+3n(1+n)-b(1+2n)+a(1-b+2n))}{(\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda))+a(c+\lambda(1-b+\lambda)))\Gamma(2+n)} \\
&+ \frac{e^{-\lambda t}(1+n)(\lambda t)^n\lambda^2(t-1)^2(3+a-b+3n)}{(\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda))+a(c+\lambda(1-b+\lambda)))\Gamma(2+n)} \\
&- \frac{e^{-\lambda t}(1+n)(\lambda t)^n\lambda^3(t-1)^3}{(\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda))+a(c+\lambda(1-b+\lambda)))\Gamma(2+n)}.
\end{aligned}
$$

$$
\begin{aligned}
g_t(z) &= \frac{e^{\lambda t(z-1)}(ac+\lambda(1+t(z-1))(c-(b-1)(1+a)))}{\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda)+a(c+\lambda(1-b+\lambda)))} \\
&+ \frac{e^{\lambda t(z-1)}\left(\lambda^2(1+t(z-1))^2(3+a-b)+\lambda^3(1+t(z-1))^3\right)}{\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda)+a(c+\lambda(1-b+\lambda)))}.
\end{aligned}
$$

$$
E(N^w(t)) = \frac{\lambda t\left(1+c+7\lambda+c\lambda+6\lambda^2+\lambda^3-b\left(1+3\lambda+\lambda^2\right)+a\left(1+c+3\lambda+\lambda^2-b(1+\lambda)\right)\right)}{\lambda(1+c+3\lambda+\lambda^2-b(1+\lambda))+a(c+\lambda(1-b+\lambda))}.
$$

$$
\begin{aligned}
Var(N^w(t)) &= \frac{(\lambda t)^2\left(8+2c+19\lambda+c\lambda+9\lambda^2+\lambda^3-b\left(4+5\lambda+\lambda^2\right)+a\left(4+c+5\lambda+\lambda^2-b(2+\lambda)\right)\right)}{\lambda(1+c+3\lambda+\lambda^2-b(1+\lambda))+a(c+\lambda(1-b+\lambda))} \\
&+ E(N^w(t)) - (E(N^w(t)))^2.
\end{aligned}
$$

$$
\begin{aligned}
g_{t-s}(z) &= \frac{e^{\lambda(t-s)(z-1)}(ac+\lambda(1+s(1-z)+t(z-1))(1-b)(1+a))}{\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda))+a(c+\lambda(1-b+\lambda))} \\
&+ \frac{e^{\lambda(t-s)(z-1)}\left(\lambda^2(1+s(1-z)+t(z-1))^2(3+a-b)+\lambda^3(1+s(1-z)+t(z-1))^3\right)}{\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda))+a(c+\lambda(1-b+\lambda))}.
\end{aligned}
$$

$$
\begin{aligned}
f_{w,(s,t)}(a,b) &= \begin{pmatrix} y \\ x \end{pmatrix} \frac{e^{-\lambda t}\left(1-\frac{s}{t}\right)^{y-x}\left(\frac{s}{t}\right)^x(\lambda t)^y(1+y)\left((a+y)(c+y(y-b))-\lambda^3(t-1)^3\right)}{\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda))+a(c+\lambda(1-b+\lambda))} \\
&\times \frac{\lambda(1-t)(1+c+3y(1+y)-b(1+2y)+a(1-b+2y))+\lambda^2(t-1)^2(3+a-b+3y)}{\Gamma(2+y)}.
\end{aligned}
$$

$$\begin{aligned}
Cov_{s,t} &= \frac{st\lambda((a(1+a)c(1-b+c)+2\lambda ac(7-a(b-3)-3b+c)))}{(\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda))+a(c+\lambda(1-b+\lambda)))^2} \\
&+ \frac{st\lambda^3\left(((1+a)(2+a)(b-2)(b-1)+c(5+2a(9+a-2b)-3b)+c^2)\right)}{(\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda))+a(c+\lambda(1-b+\lambda)))^2} \\
&- \frac{st\lambda^4(2(1+a)(5+a-b)(b-1)+(b-2a-5))}{(\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda))+a(c+\lambda(1-b+\lambda)))^2} \\
&+ \frac{st\lambda^5\left(a^2+a(9-4b)+(b-9)b+2(7+c)\right)+2st\lambda^6(3+a-b)+st\lambda^7}{(\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda))+a(c+\lambda(1-b+\lambda)))^2} \\
&+ \frac{st\lambda(1-t)(1+c+7\lambda-b(1+\lambda(3+\lambda))+a(1+c-b(1+\lambda)+\lambda(3+\lambda))+\lambda(c+\lambda(6+\lambda)))}{(\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda))+a(c+\lambda(1-b+\lambda)))^2}.
\end{aligned}$$

$$\begin{aligned}
Cov_{s,t-s} &= \frac{s(s-t)\lambda^2\left((1-b+c)^2+2\lambda(b-3)(b-c-1)+2\lambda^2(b-3)^2-4\lambda^3(b-3)+3\lambda^4\right)}{(\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda))+a(c+\lambda(1-b+\lambda)))^2} \\
&+ \frac{s(s-t)\lambda^2\left(a^2\left(1+b^2-2c-2b(1+\lambda)+2\lambda(1+\lambda)\right)\right)}{(\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda))+a(c+\lambda(1-b+\lambda)))^2} \\
&+ \frac{s(s-t)\lambda^2\left(2a\left(1+b^2(1+\lambda)-2c(1+\lambda)+2\lambda(1+\lambda)(2+\lambda)-b(2+\lambda)(1+2\lambda)\right)\right)}{(\lambda(1+c-b(1+\lambda)+\lambda(3+\lambda))+a(c+\lambda(1-b+\lambda)))^2}.
\end{aligned}$$

7.1.7 $w(n;\phi) = \frac{a^*n+b^*}{c^*n+d^*} = a + \frac{b-ac}{n+c}$

Theorem 7.7. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = a + \frac{b-ac}{n+c}$ *then*

$$
\begin{aligned}
w_t(n;\phi) &= ae^{\lambda(1-t)} + \frac{(b-ac)\gamma(c+n,\lambda(t-1))}{(\lambda(t-1))^{c+n}}.\\
f_{w,t}(n) &= \frac{(\lambda t)^n(-\lambda)^c\left(ae^{\lambda(1-t)}(\lambda(t-1))^{c+n}+(b-ac)\gamma(c+n,\lambda(t-1))\right)}{(\lambda(t-1))^{c+n}\Gamma(n+1)\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)}.\\
g_t(z) &= \frac{(-\lambda)^c\left(ae^{\lambda(1+t(z-1))}(-1+t(1-z))^c+(b-ac)\gamma(c,\lambda(-1+t(1-z)))\right)}{(-1+t(1-z))^c\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)}.\\
E(N^w(t)) &= \frac{t\left(ae^{\lambda}(-\lambda)^c(b-ac+a\lambda)-(b-ac)\Gamma(1+c)+c(b-ac)\Gamma(c,-\lambda)\right)}{ae^{\lambda}(-\lambda)^c+b\gamma(c,-\lambda)-a\Gamma(1+c)+ac\Gamma(c,-\lambda)}.\\
Var(N^w(t)) &= \frac{t^2\left((-\lambda)^c e^{\lambda}\left(-b+ac-bc+ac^2+\lambda b\right)+(-\lambda)^{c+1}e^{\lambda}ac+(-\lambda)^{c+2}e^{\lambda}a\right)}{ae^{\lambda}(-\lambda)^c+b\gamma(c,-\lambda)-a\Gamma(1+c)+ac\Gamma(c,-\lambda)}\\
&+ \frac{t^2\left((b-ac+bc)\Gamma(1+c)+\Gamma(c,-\lambda)\left(ac^2(c+1)-bc(c+1)\right)-ac^3\Gamma(c)\right)}{ae^{\lambda}(-\lambda)^c+b\gamma(c,-\lambda)-a\Gamma(1+c)+ac\Gamma(c,-\lambda)}\\
&+ E(N^w(t)) - (E(N^w(t)))^2.\\
g_{t-s}(z) &= \frac{(-\lambda)^c\left(ae^{\lambda(1+s(1-z)+t(z-1))}(\lambda(-1+t(1-z)+s(z-1)))^c\right)}{(\lambda(-1+t(1-z)+s(z-1)))^c\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)}\\
&+ \frac{(-\lambda)^c(b-ac)\gamma(c,\lambda(-1+t(1-z)+s(z-1)))}{(\lambda(-1+t(1-z)+s(z-1)))^c\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)}.\\
f_{w,(s,t)}(a,b) &= \binom{y}{x}\frac{\left(1-\frac{s}{t}\right)^{y-x}\left(\frac{s}{t}\right)^x(\lambda t)^y(-\lambda)^c\left(ae^{\lambda(1-t)}(\lambda(t-1))^{c+y}+(b-ac)\gamma(c+y,\lambda(t-1))\right)}{(\lambda(t-1))^{c+y}\Gamma(1+y)\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)}.\\
Cov_{s,t} &= \frac{se^{2\lambda}(-\lambda)^{2c}((ac-b)(a(t-1)+bt)+a\lambda(a-bt+act))}{\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)^2}\\
&+ \frac{s(b-ac)\left(e^{\lambda}(-\lambda)^c\gamma(c,-\lambda)\left(b(1+t(c+\lambda-1))+a\left(\lambda+\lambda^2t+c(t(2+\lambda)-2)\right)\right)\right)}{\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)^2}\\
&+ \frac{s(b-ac)^2(t-1)(\Gamma(c)\Gamma(1+c)+\Gamma(c,-\lambda)(c\Gamma(c,-\lambda)-2\Gamma(1+c)))}{\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)^2}.\\
Cov_{s,t-s} &= \frac{s(s-t)(b-ac)\left(e^{2\lambda}(-\lambda)^{2c}(a+b+a\lambda)-e^{\lambda}(-\lambda)^c\Gamma(c)(2ac+b(c+\lambda-1)+a\lambda(c+\lambda))\right)}{\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)^2}\\
&+ \frac{s(s-t)(b-ac)\left((-\lambda)^c E_{1-c}(-\lambda)\left(e^{\lambda}(-\lambda)^c(2ac+b(c+\lambda-1)+a\lambda(c+\lambda))+2(b-ac)\Gamma(1+c)\right)\right)}{\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)^2}\\
&+ \frac{s(s-t)(b-ac)c(ac-b)\left(\Gamma(c)^2-\Gamma(c,-\lambda)^2\right)}{\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)^2}.
\end{aligned}
$$

7.2 Probability mass/density function weight functions

7.2.1 $w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$

Theorem 7.8. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$ *then*

$$
\begin{aligned}
w_t(n;\phi) &= \binom{n+r-1}{n} p^n (1-p)^r \, {}_1F_1(n+r;1+n;\lambda p(1-t)).\\
f_{w,t}(n) &= \binom{n+r-1}{n} (\lambda pt)^n \frac{{}_1F_1(n+r;1+n;\lambda p(1-t))}{{}_1F_1(r;1;\lambda p)\Gamma(n+1)}.\\
g_t(z) &= \frac{{}_1F_1(r;1;\lambda p(1+t(z-1)))}{{}_1F_1(r;1;\lambda p)}.\\
E(N^w(t)) &= \frac{\lambda prt {}_1F_1(r+1;2;\lambda p)}{{}_1F_1(r;1;\lambda p)}.\\
Var(N^w(t)) &= \frac{(\lambda pt)^2 r(r+1) {}_1F_1(r+2;3;\lambda p)}{2 {}_1F_1(r;1;\lambda p)} + E(N^w(t)) - (E(N^w(t)))^2.\\
g_{t-s}(z) &= \frac{{}_1F_1(r;1;\lambda p(1+s(1-z)+t(z-1)))}{{}_1F_1(r;1;\lambda p)}.\\
f_{w,(s,t)}(a,b) &= \binom{b+r-1}{n}\binom{b}{a} \frac{\left(1-\frac{s}{t}\right)^{b-a}\left(\frac{s}{t}\right)^a (\lambda pt)^b {}_1F_1(b+r;1+b;\lambda p(1-t))}{\Gamma(b+1) {}_1F_1(r;1;\lambda p)}.\\
Cov_{s,t} &= \frac{prs\lambda(-2prt\lambda {}_1F_1(1+r;2;\lambda p))^2}{2 {}_1F_1(r;1;\lambda p)^2}\\
&+ {}_1F_1(r;1;\lambda p)\left(2 {}_1F_1(1+r;2;\lambda p) + p(1+r)t\lambda {}_1F_1(2+r;3;\lambda p)\right).\\
Cov_{s,t-s} &= \frac{(\lambda p)^2 rs(t-s)\Big((1+r) {}_1F_1(r;1;\lambda p) {}_1F_1(2+r;3;\lambda p) - 2r {}_1F_1(1+r;2;\lambda p)^2\Big)}{2 {}_1F_1(r;1;\lambda p)^2}.
\end{aligned}
$$

7.2.2 $w(n;\phi) = \begin{pmatrix} m \\ n \end{pmatrix} p^n (1-p)^{m-n}$

Theorem 7.9. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \begin{pmatrix} m \\ n \end{pmatrix} p^n (1-p)^{m-n}$ *then*

$$
\begin{aligned}
w_t(n;\phi) &= \begin{pmatrix} m \\ n \end{pmatrix} p^n (1-p)^{m-n} \, {}_1F_1\left(n-m; 1+n; \tfrac{\lambda p(1-t)}{p-1}\right). \\
f_{w,t}(n) &= \begin{pmatrix} m \\ n \end{pmatrix} \left(\frac{\lambda pt}{1-p}\right)^n \frac{{}_1F_1\left(n-m;1+n;\frac{\lambda p(1-t)}{p-1}\right)}{L_m\left(\frac{p\lambda}{p-1}\right)\Gamma(1+n)}. \\
g_t(z) &= \frac{L_m\left(\frac{p\lambda(1+t(z-1))}{p-1}\right)}{L_m\left(\frac{p\lambda}{p-1}\right)}. \\
E(N^w(t)) &= \frac{\lambda pt L^1_{m-1}\left(\frac{p\lambda}{p-1}\right)}{(p-1)L_m\left(\frac{p\lambda}{p-1}\right)}. \\
Var(N^w(t)) &= \frac{(\lambda pt)^2 L^2_{m-2}\left(\frac{p\lambda}{p-1}\right)}{(p-1)^2 L_m\left(\frac{p\lambda}{p-1}\right)} + E(N^w(t)) - (E(N^w(t)))^2. \\
g_{t-s}(z) &= \frac{L_m\left(\frac{p\lambda(1+s(1-z)+t(z-1))}{p-1}\right)}{L_m\left(\frac{p\lambda}{p-1}\right)}. \\
f_{w,(s,t)}(a,b) &= \begin{pmatrix} m \\ b \end{pmatrix}\begin{pmatrix} b \\ a \end{pmatrix} \frac{\left(1-\frac{s}{t}\right)^{b-a}\left(\frac{s}{t}\right)^a (\lambda pt)^b {}_1F_1\left(b-m;1+b;\frac{p\lambda(1-t)}{p-1}\right)}{\Gamma(1+b)(1-p)^b L_m\left(\frac{p\lambda}{p-1}\right)}. \\
Cov_{s,t} &= \frac{ps\lambda\left(L_m\left(\frac{p\lambda}{p-1}\right)\left(pt\lambda L^2_{m-2}\left(\frac{p\lambda}{p-1}\right)+(1-p)L^1_{m-1}\left(\frac{p\lambda}{p-1}\right)\right)-pt\lambda L^1_{m-1}\left(\frac{p\lambda}{p-1}\right)^2\right)}{\left((p-1)L_m\left(\frac{p\lambda}{p-1}\right)\right)^2}. \\
Cov_{s,t-s} &= \frac{(\lambda p)^2 s(t-s)\left(L^1_{m-1}\left(\frac{p\lambda}{p-1}\right)^2 - L_m\left(\frac{p\lambda}{p-1}\right)L^2_{m-2}\left(\frac{p\lambda}{p-1}\right)\right)}{\left((p-1)L_m\left(\frac{p\lambda}{p-1}\right)\right)^2}.
\end{aligned}
$$

7.2.3 $w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$

Theorem 7.10. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$ *then*

$$
\begin{aligned}
w_t(n;\phi) &= a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\,{}_1F_1(n;n+a+1;\lambda(1-t)).\\
f_{w,t}(n) &= \frac{t^n\lambda^{n-1}}{n}\frac{{}_1F_1(n;n+a+1;\lambda(1-t))}{{}_2F_2(1,1;2,2+a;\lambda)}\frac{\Gamma(2)\Gamma(2+a)}{\Gamma(n+a+1)}.\\
g_t(z) &= \frac{(1+t(z-1))\,{}_2F_2(1,1;2,2+a;\lambda(1+t(z-1)))}{{}_2F_2(1,1;2,2+a;\lambda)}.\\
E(N^w(t)) &= \frac{(1+a)e^{\lambda}t\lambda^{-1-a}\gamma(1+a,\lambda)}{{}_2F_2(1,1;2,2+a;\lambda)}.\\
Var(N^w(t)) &= \frac{t^2\lambda^{-1-a}\left(e^{\lambda}(\lambda-a-1)\Gamma(2+a)+(1+a)\left(\lambda^{1+a}-e^{\lambda}(\lambda-a-1)\Gamma(1+a,\lambda)\right)\right)}{{}_2F_2(1,1;2,2+a;\lambda)}\\
&+ E(N^w(t))-(E(N^w(t)))^2.\\
g_{t-s}(z) &= \frac{(1+s(1-z)+t(z-1))\,{}_2F_2(1,1;2,2+a;\lambda(1+s(1-z)+t(z-1)))}{{}_2F_2(1,1;2,2+a;\lambda)}.\\
f_{w,(s,t)}(a,b) &= \begin{pmatrix} b\\ a\end{pmatrix}\frac{\left(1-\frac{s}{t}\right)^{b-a}\left(\frac{s}{t}\right)^{a}t^b\lambda^{b-1}\,{}_1F_1(b;1+a+b;\lambda(1-t))}{b\,{}_2F_2(1,1;2,2+a;\lambda)}\frac{\Gamma(2)\Gamma(2+a)}{\Gamma(1+a+b)}.\\
Cov_{s,t} &= \frac{(1+a)s\lambda^{-2(1+a)}\left(-(1+a)e^{2\lambda}t\gamma(1+a,\lambda)^2\right)}{{}_2F_2(1,1;2,2+a;\lambda)^2}\\
&+ \frac{(1+a)s\lambda^{-2(1+a)}\left(\lambda^{1+a}\left(t\lambda^{1+a}+e^{\lambda}(1+t(\lambda-1-a))\gamma(1+a,\lambda)\right){}_2F_2(1,1;2,2+a;\lambda)\right)}{{}_2F_2(1,1;2,2+a;\lambda)^2}.\\
Cov_{s,t-s} &= \frac{(1+a)s(s-t)\lambda^{-2(1+a)}\left((1+a)e^{2\lambda}\gamma(1+a,\lambda)^2\right)}{{}_2F_2(1,1;2,2+a;\lambda)^2}\\
&+ \frac{(1+a)s(s-t)\lambda^{-2(1+a)}\left(\lambda^{1+a}\left(-\lambda^{1+a}+e^{\lambda}(1+a-\lambda)\gamma(1+a,\lambda)\right){}_2F_2(1,1;2,2+a;\lambda)\right)}{{}_2F_2(1,1;2,2+a;\lambda)^2}.
\end{aligned}
$$

7.2.4 $w(n;\phi) = \frac{ab^a}{n^{a+1}}$

Theorem 7.11. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \frac{ab^a}{n^{a+1}}$ *then*

$$\begin{aligned}
w_t(n;\phi) &= \frac{ab^a {}_{a+1}F_{a+1}(n,\ldots,n;n+1,\ldots,n+1;\lambda(1-t))}{n^{a+1}}. \\
f_{w,t}(n) &= \frac{t^n\lambda^{n-b} {}_{a+1}F_{a+1}(n,\ldots,n;n+1,\ldots,n+1;\lambda(1-t))}{n {}_{a+2}F_{a+2}(1,b,\ldots b;b+1,\ldots,b+1;\lambda)} \frac{\Gamma(n)^a\Gamma(b+1)^{a+2}}{\Gamma(n+1)^{a+1}\Gamma(b)^{a+1}}. \\
g_t(z) &= \frac{(1+t(z-1))^b {}_{a+2}F_{a+2}(1,b,\ldots b;b+1,\ldots,b+1;(1+t(z-1))\lambda)}{{}_{a+2}F_{a+2}(1,b,\ldots b;b+1,\ldots,b+1;\lambda)}. \\
E(N^w(t)) &= \frac{tb {}_{a+1}F_{a+1}(1,b,\ldots b;b+1,\ldots,b+1;\lambda)}{{}_{a+2}F_{a+2}(1,b,\ldots b;b+1,\ldots,b+1;\lambda)}. \\
Var(N^w(t)) &= \frac{t^2b(b {}_aF_a(1,b,\ldots b;b+1,\ldots,b+1;\lambda) - {}_{a+1}F_{a+1}(1,b,\ldots b;b+1,\ldots,b+1;\lambda))}{{}_{a+2}F_{a+2}(1,b,\ldots b;b+1,\ldots,b+1;\lambda)} \\
&+ E(N^w(t)) - (E(N^w(t)))^2. \\
g_{t-s}(z) &= \frac{(1+s(1-z)+t(z-1))^b {}_{a+2}F_{a+2}(1,b,\ldots b;b+1,\ldots,b+1;(1+s(1-z)+t(z-1))\lambda)}{{}_{a+2}F_{a+2}(1,b,\ldots b;b+1,\ldots,b+1;\lambda)}. \\
f_{w,(s,t)}(x,y) &= \binom{y}{x} \frac{(1-\frac{s}{t})^{y-x}(\frac{s}{t})^x t^y\lambda^{y-b} {}_{a+1}F_{a+1}(y,\ldots,y;1+y,\ldots,1+y;\lambda(1-t))}{y {}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,\ldots,1+b;\lambda)} \\
&\times \frac{\Gamma(y)^a\Gamma(1+b)^{a+2}}{\Gamma(1+y)^{a+1}\Gamma(b)^{a+1}}. \\
Cov_{s,t} &= \frac{bs\left(-bt {}_{a+1}F_{a+1}(1,b,\ldots,b;1+b,\ldots,1+b;\lambda)^2\right)}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,\ldots,1+b;\lambda)^2} \\
&- \frac{bs((t-1) {}_{a+1}F_{a+1}(1,b,\ldots,b;1+b,\ldots,1+b;\lambda) {}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,\ldots,1+b;\lambda))}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,\ldots,1+b;\lambda)^2} \\
&+ \frac{bs(bt {}_aF_a(1,b,\ldots,b;1+b,\ldots,1+b;\lambda) {}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,\ldots,1+b;\lambda))}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,\ldots,1+b;\lambda)^2}. \\
Cov_{s,t-s} &= \frac{bs(s-t)b {}_{a+1}F_{a+1}(1,b,\ldots,b;1+b,\ldots,1+b;\lambda)^2}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,\ldots,1+b;\lambda)^2} \\
&+ \frac{bs(s-t)({}_{a+1}F_{a+1}(1,b,\ldots,b;1+b,\ldots,1+b;\lambda) {}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,\ldots,1+b;\lambda))}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,\ldots,1+b;\lambda)^2} \\
&- \frac{bs(s-t)(b {}_aF_a(1,b,\ldots,b;1+b,\ldots,1+b;\lambda) {}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,\ldots,1+b;\lambda))}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,\ldots,1+b;\lambda)^2}.
\end{aligned}$$

7.2.5 $w(n;\phi) = \frac{-1}{\ln(1-p)}\frac{p^n}{n}$

Theorem 7.12. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \frac{-1}{\ln(1-p)}\frac{p^n}{n}$ *then*

$$
\begin{aligned}
w_t(n;\phi) &= -\frac{p^n \gamma(n,p(t-1)\lambda)}{(p(t-1)\lambda)^n \ln(1-p)}.\\
f_{w,t}(n) &= \frac{(\lambda p)^{n-1} t^n {}_1F_1(n;n+1;p(1-t)\lambda)}{n(n!){}_2F_2(1,1;2,2;p\lambda)}.\\
g_t(z) &= \frac{(1+t(z-1)){}_2F_2(1,1;2,2;p(1+t(z-1))\lambda)}{{}_2F_2(1,1;2,2;p\lambda)}.\\
E(N^w(t)) &= \frac{t\,{}_1F_1(1;2;p\lambda)}{{}_2F_2(1,1;2,2;p\lambda)}.\\
Var(N^w(t)) &= \frac{t^2\left(e^{p\lambda}-{}_1F_1(1;2;p\lambda)\right)}{{}_2F_2(1,1;2,2;p\lambda)} + E(N^w(t)) - \left(E(N^w(t))\right)^2.\\
g_{t-s}(z) &= \frac{(1+s(1-z)+t(z-1)){}_2F_2(1,1;2,2;p(1+s(1-z)+t(z-1))\lambda)}{{}_2F_2(1,1;2,2;p\lambda)}.\\
f_{w,(s,t)}(x,b) &= \begin{pmatrix} y \\ x \end{pmatrix} \frac{\left(1-\frac{s}{t}\right)^{y-x}\left(\frac{s}{t}\right)^x t^y (\lambda p)^{y-1} {}_1F_1(y;1+y;p(1-t)\lambda)}{y(y!){}_2F_2(1,1;2,2;p\lambda)}.\\
Cov_{s,t} &= \frac{s\left(\left(e^{p\lambda}t+(1-t){}_1F_1(1;2;p\lambda)\right){}_2F_2(1,1;2,2;p\lambda)-t\,{}_1F_1(1;2;p\lambda)^2\right)}{{}_2F_2(1,1;2,2;p\lambda)^2}.\\
Cov_{s,t-s} &= \frac{s(s-t)\left(\left({}_1F_1(1;2;p\lambda)-e^{p\lambda}\right){}_2F_2(1,1;2,2;p\lambda)+{}_1F_1(1;2;p\lambda)^2\right)}{{}_2F_2(1,1;2,2;p\lambda)^2}.
\end{aligned}
$$

7.2.6 $w(n;\phi) = \frac{\Gamma(r+n)}{n!\Gamma(r)}\frac{Beta(a+r,b+n)}{Beta(a,b)}$

Theorem 7.13. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \frac{\Gamma(r+n)}{n!\Gamma(r)}\frac{Beta(a+r,b+n)}{Beta(a,b)}$ *then*

$$
\begin{aligned}
w_t(n;\phi) &= \frac{\Gamma(r+n)Beta(a+r,b+n)\,{}_2F_2(b+n,r+n;1+n,a+b+n+r;\lambda(1-t))}{\Gamma(r)\Gamma(1+n)Beta(a,b)}.\\
f_{w,t}(n) &= \frac{(\lambda t)^n\Gamma(r+n)Beta(a+r,b+n)\,{}_2F_2(b+n,r+n;1+n,a+b+n+r;\lambda(1-t))}{\Gamma(r)\Gamma(1+n)^2Beta(a+r,b)\,{}_2F_2(b,r;1,a+b+r;\lambda)}.\\
g_t(z) &= \frac{{}_2F_2(b,r;1,a+b+r;(1+t(z-1))\lambda)}{{}_2F_2(b,r;1,a+b+r;\lambda)}.\\
E(N^w(t)) &= \frac{brt\lambda\,{}_2F_2(1+b,1+r;2,1+a+b+r;\lambda)}{(a+b+r)\,{}_2F_2(b,r;1,a+b+r;\lambda)}.\\
Var(N^w(t)) &= \frac{b(1+b)r(1+r)(t\lambda)^2\,{}_2F_2(2+b,2+r;3,2+a+b+r;\lambda)}{2(a+b+r)(1+a+b+r)\,{}_2F_2(b,r;1,a+b+r;\lambda)} + E(N^w(t)) - (E(N^w(t)))^2.\\
g_{t-s}(z) &= \frac{{}_2F_2(b,r;1,a+b+r;(1+s(1-z)+t(z-1))\lambda)}{{}_2F_2(b,r;1,a+b+r;\lambda)}.\\
f_{w,(s,t)}(x,y) &= \begin{pmatrix} y\\ x\end{pmatrix}\frac{\left(1-\frac{s}{t}\right)^{y-x}\left(\frac{s}{t}\right)^x(\lambda t)^yBeta(a+r,y+b)}{Beta(a+r,b)}\\
&\times \frac{\Gamma(y+r)\,{}_2F_2(y+b,y+r;1+y,1+b+y+r;\lambda(1-t))}{\Gamma(1+y)^2\Gamma(r)\,{}_2F_2(b,r;1,a+b+r;\lambda)}.\\
Cov_{s,t} &= \left(\frac{brs\lambda\left(-brt\lambda\,{}_2F_2(1+b,1+r;2,1+a+b+r;\lambda)^2\right)}{\Gamma(1+c+d+r)^2} + \frac{brs\lambda({}_2F_2(b,r;1,a+b+r;\lambda))}{\Gamma(c+d+r)}\times\right.\\
&\left.\left(\frac{{}_2F_2(1+b,1+r;2,1+a+b+r;\lambda)}{\Gamma(1+a+b+r)} + \frac{(1+b)(1+r)t\lambda\,{}_2F_2(2+b,2+r;3,2+a+b+r;\lambda)}{\Gamma(3)\Gamma(2+a+b+r)}\right)\right)\\
&\times \frac{\Gamma(a+b+r)^2}{{}_2F_2(b,r;1,a+b+r;\lambda)^2}.\\
Cov_{s,t-s} &= \left(\left(\frac{s(s-t)(\lambda br)^2\,{}_2F_2(1+b,1+r;2,1+a+b+r;\lambda)^2}{\Gamma(1+a+b+r)^2}\right)\right.\\
&\left.-\frac{brs(s-t)\lambda^2(1+b)(1+r)\,{}_2F_2(b,r;1,a+b+r;\lambda)\,{}_2F_2(2+b,2+r;3,2+a+b+r;\lambda)}{\Gamma(a+b+r)\Gamma(3)\Gamma(2+a+b+r)}\right)\\
&\times \frac{\Gamma(a+b+r)^2}{{}_2F_2(b,r;1,a+b+r;\lambda)^2}.
\end{aligned}
$$

7.3 Truncating weight functions

7.3.1 $w(n;\phi) = I(n \geq a)$

Theorem 7.14. *If the weight function used in the weighted Poisson probability mass function is chosen as $w(n;\phi) = I(n \geq a)$ then*

$$
\begin{aligned}
w_t(n;\phi) &= e^{\lambda(1-t)}. \\
f_{w,t}(n) &= \frac{(t\lambda)^n e^{-\lambda t}\Gamma(a)}{n!\gamma(a,\lambda)}. \\
g_t(z) &= \frac{e^{\lambda t(z-1)}\gamma(a,\lambda(1+t(z-1)))}{\gamma(a,\lambda)}. \\
E(N^w(t)) &= t\left(\lambda + \frac{e^{-\lambda}\lambda^a}{\gamma(a,\lambda)}\right). \\
Var(N^w(t)) &= t\left(\lambda - \frac{te^{-2\lambda}\lambda^{2a}}{\gamma(a,\lambda)^2} + \frac{e^{-\lambda}(1+t(a-\lambda-1))\lambda^a}{\gamma(a,\lambda)}\right). \\
g_{t-s}(z) &= \frac{e^{(t-s)\lambda(1-z)}\gamma(a,1+s(1-z)+t(z-1))}{\gamma(a,\lambda)}. \\
f_{w,(s,t)}(x,y) &= \begin{pmatrix} y \\ x \end{pmatrix} \frac{\left(1-\frac{s}{t}\right)^{y-x}\left(\frac{s}{t}\right)^x(\lambda t)^y e^{-\lambda t}\Gamma(a)}{\Gamma(1+y)\gamma(a,\lambda)}. \\
Cov_{s,t} &= s\left(\lambda - \frac{te^{-2\lambda}\lambda^{2c}}{\gamma(a,\lambda)^2} + \frac{e^{-\lambda}(1+t(a-\lambda-1))\lambda^a}{\gamma(a,\lambda)}\right). \\
Cov_{s,t-s} &= \frac{s(s-t)\lambda^a e^{-2\lambda}\left(\lambda^a + e^{\lambda}(\lambda+1-a)\gamma(a,\lambda)\right)}{\gamma(a,\lambda)^2}.
\end{aligned}
$$

7.3.2 $w(n;\phi) = I(n \leq b)$

Theorem 7.15. *If the weight function used in the weighted Poisson probability mass function is chosen as $w(n;\phi) = I(n \leq b)$ then*

$$
\begin{aligned}
w_t(n;\phi) &= \tfrac{e^{\lambda(1-t)}\Gamma(1+b-n,\lambda(1-t))}{(b-n)!}.\\
f_{w,t}(n) &= \tfrac{e^{-\lambda t}(\lambda t)^n b!\Gamma(1+b-n,\lambda(1-t))}{n!(b-n)!\Gamma(1+b,\lambda)}.\\
g_t(z) &= \tfrac{e^{-\lambda t(z-1)}\Gamma(1+b,\lambda(1+t(z-1)))}{\Gamma(1+b,\lambda)}.\\
E(N^w(t)) &= \lambda t\left(1-\tfrac{e^{-\lambda}\lambda^b}{\Gamma(1+b,\lambda)}\right).\\
Var(N^w(t)) &= -\tfrac{e^{-\lambda}t^2\lambda\left(b\lambda^b+\lambda^{b+1}-\lambda e^{\lambda}\Gamma(1+b,\lambda)\right)}{\Gamma(1+b,\lambda)} + E(N^w(t)) - (E(N^w(t)))^2.\\
g_{t-s}(z) &= \tfrac{e^{\lambda(t-s)(z-1)}\Gamma(1+b,\lambda(1+s(1-z)+t(z-1)))}{\Gamma(1+b,\lambda)}.\\
f_{w,(s,t)}(x,y) &= \binom{y}{x}\tfrac{\left(1-\frac{s}{t}\right)^{y-x}\left(\frac{s}{t}\right)^x(\lambda t)^y e^{-\lambda t}b!\Gamma(1+b-y,\lambda(1-t))}{y!(b-y)!\Gamma(1+b,\lambda)}.\\
Cov_{s,t} &= s\left(\lambda-\tfrac{e^{-2\lambda}t}{E_{-b}(\lambda)^2}+\tfrac{e^{-\lambda}(\lambda t-tb-1)}{E_{-b}(\lambda)}\right).\\
Cov_{s,t-s} &= \tfrac{e^{-2\lambda}s(s-t)\left(1-e^{\lambda}(\lambda-b)E_{-b}(\lambda)\right)}{E_{-b}(\lambda)^2}.
\end{aligned}
$$

Where $E_n(.)$ is the exponential integral (Definition 10.22).

7.3.3 $w(n;\phi) = I(n \geq a)\, I(n \leq b)$

Theorem 7.16. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = I(n \geq a)\, I(n \leq b)$ *then*

$$
\begin{aligned}
w_t(n;\phi) &= \frac{e^{\lambda(1-t)}\Gamma(1+b-n,\lambda(1-t))}{(b-n)!}. \\
f_{w,t}(n) &= \frac{e^{-\lambda t}(\lambda t)^n b!\Gamma(a)\Gamma(1+b-n,\lambda(1-t))}{n!(b-n)!(\Gamma(1+b,\lambda)\Gamma(a)-b!\Gamma(a,\lambda))}. \\
g_t(z) &= \frac{e^{\lambda t(z-1)}(b!\Gamma(a,\lambda(1+t(z-1)))-\Gamma(a)\Gamma(1+b,\lambda(1+t(z-1))))}{b!\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)}. \\
E(N^w(t)) &= \frac{e^{-\lambda}t\left(\Gamma(1+b)\left(\lambda e^{\lambda}\Gamma(a,\lambda)-\lambda^a\right)+\lambda\Gamma(a)\left(\lambda^b-e^{\lambda}\Gamma(1+b,\lambda)\right)\right)}{b!\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)}. \\
Var(N^w(t)) &= -\frac{e^{-\lambda}t^2\left(\lambda^{1+b}\Gamma(a)\left(\lambda-e^{\lambda}\lambda^2E_{-b}(\lambda)+b\right)+\lambda^a\Gamma(1+b)\left(1-\lambda-a+e^{\lambda}\lambda^2E_{1-a}(\lambda)\right)\right)}{b!\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)} \\
&+ E(N^w(t)) - (E(N^w(t)))^2. \\
g_{t-s}(z) &= \frac{e^{\lambda(t-s)(z-1)}(b!\Gamma(a,\lambda(1+s(1-z)+t(z-1)))-\Gamma(a)\Gamma(1+b,\lambda(1+s(1-z)+t(z-1))))}{b!\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)}. \\
f_{w,(s,t)}(x,y) &= \begin{pmatrix} y \\ x \end{pmatrix} \frac{\left(1-\frac{s}{t}\right)^{y-x}\left(\frac{s}{t}\right)^x(\lambda t)^y e^{-\lambda t}b!\Gamma(a)\Gamma(1+b-y,\lambda(1-t))}{y!(b-y)!(\Gamma(a)\Gamma(1+b,\lambda)-b!\Gamma(a,\lambda))}. \\
Cov_{s,t} &= \frac{e^{-2\lambda}s\lambda^{2a}(b!)^2\left(e^{\lambda}E_{1-a}(\lambda)\left(\lambda t+t-1+\lambda e^{\lambda}E_{1-a}(\lambda)\right)-t\right)}{(b!\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda))^2} \\
&+ \frac{e^{-2\lambda}s\lambda^{1+a+b}b!\Gamma(a)\left(2t-e^{\lambda}E_{-b}(\lambda)(\lambda t+t-1)+e^{\lambda}E_{1-a}(\lambda)\left(1-\lambda t+tb-2\lambda e^{\lambda}E_{-b}(\lambda)\right)\right)}{(b!\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda))^2} \\
&+ \frac{e^{-2\lambda}s\lambda^{2(1+b)}\Gamma(a)^2\left(\lambda e^{2\lambda}(E_{-b}(\lambda))^2+e^{\lambda}E_{-b}(\lambda)(\lambda t-tb-1)-t\right)}{(b!\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda))^2} \\
&- \frac{e^{-2\lambda}se^{\lambda}t\lambda^a b!(ab!\Gamma(a,\lambda)-a!\Gamma(1+b,\lambda))}{(b!\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda))^2}. \\
Cov_{s,t-s} &= \frac{e^{-2\lambda}s(s-t)\lambda^{1+a+b}b!\Gamma(a)\left(e^{\lambda}(1+\lambda-a)E_{-b}(\lambda)+e^{\lambda}(\lambda-b)E_{1-a}(\lambda)-2\right)}{(b!\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda))^2} \\
&+ \frac{e^{-2\lambda}s(s-t)\lambda^{2(1+b)}\Gamma(a)^2\left(1-e^{\lambda}(\lambda-b)E_{-b}(\lambda)\right)}{(b!\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda))^2} \\
&+ \frac{e^{-2\lambda}s(s-t)\lambda^a(b!)^2\Gamma(a,\lambda)\left(\lambda^a-e^{\lambda}(1+\lambda-a)\right)}{(b!\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda))^2}.
\end{aligned}
$$

7.3.4 $w(n;\phi)=\begin{pmatrix} n \\ a \end{pmatrix}$

Theorem 7.17. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi)=\begin{pmatrix} n \\ a \end{pmatrix}$ *then*

$$\begin{aligned}
w_t(n;\phi) &= \begin{pmatrix} n \\ a \end{pmatrix} {}_1F_1\left(1+n,1-a+n;\lambda(1-t)\right). \\
f_{w,t}(n) &= \begin{pmatrix} n \\ a \end{pmatrix} \frac{e^{-\lambda}t^n\lambda^{n-a}a!\,{}_1F_1(1+n,1-a+n;\lambda(1-t))}{n!}. \\
g_t(z) &= e^{\lambda t(z-1)}\left(1+t(z-1)\right)^a. \\
E\left(N^w(t)\right) &= t(a+\lambda). \\
Var\left(N^w(t)\right) &= t\left(a(1-t)+\lambda\right). \\
g_{t-s}(z) &= e^{\lambda(t-s)(z-1)}\left(1+s(1-z)+t(z-1)\right)^a. \\
f_{w,(s,t)}(x,y) &= \begin{pmatrix} y \\ x \end{pmatrix}\begin{pmatrix} y \\ a \end{pmatrix} \frac{\left(1-\frac{s}{t}\right)^{y-x}\left(\frac{s}{t}\right)^x e^{-\lambda}\lambda^{y-a}t^y a!\,{}_1F_1(1+y,1-a+y;\lambda(1-t))}{y!}. \\
Cov_{s,t} &= s\left(a(1-t)+\lambda\right). \\
Cov_{s,t-s} &= as(s-t).
\end{aligned}$$

7.4 Miscellaneous weight functions

7.4.1 $w(n;\phi) = \varepsilon \frac{n!}{e^{-\lambda}\lambda^n} I(n=0) + (1-\varepsilon)$

Theorem 7.18. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \varepsilon \frac{n!}{e^{-\lambda}\lambda^n} I(n=0) + (1-\varepsilon)$ *then*

$$
\begin{aligned}
w_t(n;\phi) &= e^{\lambda(1-t)}(1-\varepsilon). \\
f_{w,t}(n) &= \frac{e^{-\lambda t}(\lambda t)^n (1-\varepsilon)}{n!}. \\
g_t(z) &= e^{\lambda t(z-1)}(1-\varepsilon)+\varepsilon. \\
E(N^w(t)) &= \lambda t(1-\varepsilon). \\
Var(N^w(t)) &= \lambda t(1-\varepsilon)(1+\varepsilon\lambda t). \\
g_{t-s}(z) &= e^{\lambda(t-s)(z-1)}(1-\varepsilon)+\varepsilon. \\
f_{w,(s,t)}(x,y) &= \begin{pmatrix} y \\ x \end{pmatrix} \frac{\left(1-\frac{s}{t}\right)^{y-x}\left(\frac{s}{t}\right)^x e^{-\lambda t}(\lambda t)^y (1-\varepsilon)}{y!}. \\
Cov_{s,t} &= s(1-\varepsilon)\lambda(1+\varepsilon\lambda t). \\
Cov_{s,t-s} &= s(t-a)(1-\varepsilon)\varepsilon\lambda^2.
\end{aligned}
$$

7.4.2 $w(n;\phi) = (a)_n = \frac{\Gamma(a+n)}{\Gamma(a)}$

Theorem 7.19. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = (a)_n$ *then*

$$\begin{aligned}
w_t(n;\phi) &= \frac{(a)_n}{(1+\lambda(t-1))^{a+n}}.\\
f_{w,t}(n) &= \frac{(\lambda t)^n (1-\lambda)^a (a)_n}{(1+\lambda(t-1))^{a+n}}.\\
g_t(z) &= \frac{(1-\lambda)^a}{(1+\lambda(t(1-z)-1))}.\\
E(N^w(t)) &= \frac{at\lambda}{1-\lambda}.\\
Var(N^w(t)) &= \frac{at\lambda(1+\lambda(t-1))}{(\lambda-1)^2}.\\
g_{t-s}(z) &= \frac{(1-\lambda)^a}{(1+\lambda(s(z-1)+t(1-z)-1))^a}.\\
f_{w,(s,t)}(x,y) &= \begin{pmatrix} y \\ x \end{pmatrix} \frac{(1-\frac{s}{t})^{y-x}(\frac{s}{t})^x (\lambda t)^y (1-\lambda)^a (a)_y}{y!(1+\lambda(t-1))^{y+a}}.\\
Cov_{s,t} &= \frac{as\lambda(1+\lambda(t-1))}{(\lambda-1)^2}.\\
Cov_{s,t-s} &= \frac{as(t-s)\lambda^2}{(\lambda-1)^2}.
\end{aligned}$$

7.4.3 $w(n;\phi) = (n)_a = \frac{\Gamma(a+n)}{\Gamma(n)}$

Theorem 7.20. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = (n)_a$ *then*

$$
\begin{aligned}
w_t(n;\phi) &= (n)_a\, {}_1F_1(a+n, n; \lambda(1-t)). \\
f_{w,t}(n) &= \frac{t^n \lambda^{n-1} \Gamma(n+a)\, {}_1F_1(a+n,n;\lambda(1-t))}{\Gamma(n)\Gamma(1+a)\Gamma(1+n)\, {}_1F_1(1+a,2;\lambda)}. \\
g_t(z) &= \frac{(1+t(z-1))\, {}_1F_1(1+a,2;\lambda(1+t(z-1)))}{{}_1F_1(1+a,2;\lambda)}. \\
E(N^w(t)) &= t + \frac{(1+a)\lambda t\, {}_1F_1(2+a,3;\lambda)}{2\, {}_1F_1(1+a,2;\lambda)}. \\
Var(N^w(t)) &= \frac{t^2 \lambda (1+a)(6\, {}_1F_1(2+a,3;\lambda) + \lambda(2+a)\, {}_1F_1(3+a,4;\lambda))}{6\, {}_1F_1(1+a,2;\lambda)} \\
&+ E(N^w(t)) - (E(N^w(t)))^2. \\
g_{t-s}(z) &= \frac{(1+s(1-z)+t(z-1))\, {}_1F_1(1+a,2;\lambda(1+s(1-z)+t(z-1)))}{{}_1F_1(1+a,2;\lambda)}. \\
f_{w,(s,t)}(x,y) &= \binom{y}{x} \frac{\left(1-\frac{s}{t}\right)^{y-x} \left(\frac{s}{t}\right)^x t^y \lambda^{y-1}\, {}_1F_1(a+y,y;\lambda(1-t))}{{}_1F_1(1+a,2;\lambda)} \frac{\Gamma(a+y)}{\Gamma(1+a)\Gamma(1+y)\Gamma(y)}. \\
Cov_{s,t} &= \frac{s\left(12(t-1)\, {}_1F_1(1+a,2;\lambda)^2 + 3(1+a)^2 t \lambda^2\, {}_1F_1(2+a,3;\lambda)^2\right)}{12\, {}_1F_1(1+a,2;\lambda)^2} \\
&- \frac{s2(1+a)\lambda\, {}_1F_1(1+a,2;\lambda)(3\, {}_1F_1(2+a,3;\lambda) + (2+a)t\lambda\, {}_1F_1(3+a,4;\lambda))}{12\, {}_1F_1(1+a,2;\lambda)^2}. \\
Cov_{s,t-s} &= \frac{s(s-t)\left(12\, {}_1F_1(1+a,2;\lambda)^2 + 3(1+a)^2 \lambda^2\, {}_1F_1(2+a,3;\lambda)^2\right)}{12\, {}_1F_1(1+a,2;\lambda)^2} \\
&- \frac{s(s-t)2(1+a)(2+a)\lambda^2\, {}_1F_1(1+a,2;\lambda)\, {}_1F_1(3+a,4;\lambda)}{12\, {}_1F_1(1+a,2;\lambda)^2}.
\end{aligned}
$$

7.5 Inverted weight functions

7.5.1 $w(n;\phi)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$

Theorem 7.21. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$ *then*

$$
\begin{aligned}
w_t(n;\phi) &= \frac{(n+c)\,{}_2F_2\left(\frac{b}{a}+n,1+c+n;1+\frac{b}{a}+n,c+n;\lambda(1-t)\right)}{n+an}.\\
f_{w,t}(n) &= \frac{b(c+n)(\lambda t)^n\,{}_2F_2\left(\frac{b}{a}+n,1+c+n;1+\frac{b}{a}+n,c+n;\lambda(1-t)\right)}{c(b+an)n!\,{}_2F_2\left(\frac{b}{a},1+c;1+\frac{b}{a},c;\lambda\right)}.\\
g_t(z) &= \frac{{}_2F_2\left(\frac{b}{a},1+c;1+\frac{b}{a},c;\lambda(1+t(z-1))\right)}{{}_2F_2\left(\frac{b}{a},1+c;1+\frac{b}{a},c;\lambda\right)}.\\
E\left(N^w(t)\right) &= \frac{bt\left(ae^{\lambda}(-\lambda)^{\frac{b}{a}}(a(c+\lambda)-b)+a(b-ac)\Gamma\left(\frac{a+b}{a}\right)+b(ac-b)\Gamma\left(\frac{b}{a},-\lambda\right)\right)}{(-\lambda)^{\frac{b}{a}}a^3c\,{}_2F_2\left(\frac{b}{a},1+c;1+\frac{b}{a},c;\lambda\right)}.\\
Var\left(N^w(t)\right) &= \frac{t^2ae^{\lambda}(-\lambda)^{\frac{b}{a}}\left(b^2-ab(\lambda+c-1)+a^2\left(\lambda^2+c(\lambda-1)\right)\right)}{a^2\left(ae^{\lambda}(-\lambda)^{\frac{b}{a}}+(ac-b)\gamma\left(\frac{b}{a},-\lambda\right)\right)}\\
&- \frac{t^2ae^{\lambda}(-\lambda)^{\frac{b}{a}}\left(b(a+b)(b-ac)\gamma\left(\frac{b}{a},-\lambda\right)\right)}{a^2\left(ae^{\lambda}(-\lambda)^{\frac{b}{a}}+(ac-b)\gamma\left(\frac{b}{a},-\lambda\right)\right)}+E\left(N^w(t)\right)-\left(E\left(N^w(t)\right)\right)^2.\\
g_{t-s}(z) &= \frac{{}_2F_2\left(\frac{b}{a},1+c;1+\frac{b}{a},c;\lambda(1+s(1-z)+t(z-1))\right)}{{}_2F_2\left(\frac{b}{a},1+c;1+\frac{b}{a},c;\lambda\right)}.\\
f_{w,(s,t)}(x,y) &= \binom{y}{x}\frac{\left(1-\frac{s}{t}\right)^{y-x}\left(\frac{s}{t}\right)^x(\lambda t)^y(c+y)b\,{}_2F_2\left(\frac{b}{a}+y,1+c+y;1+\frac{b}{a}+y,c+y;\lambda(1-t)\right)}{c(a+ay)y!\,{}_2F_2\left(\frac{b}{a},1+c;1+\frac{b}{a},c;\lambda\right)}.\\
Cov_{s,t} &= \frac{st}{a^2}\frac{\left(ae^{\lambda}(-\lambda)^{\frac{b}{a}}\left(b^2-ab(\lambda+c-1)+a^2\left(\lambda^2+c(\lambda-1)\right)\right)-b(a+b)(b-ac)\gamma\left(\frac{b}{a},-\lambda\right)\right)}{ae^{\lambda}(-\lambda)^{\frac{b}{a}}-(b-ac)\gamma\left(\frac{b}{a},-\lambda\right)}\\
&- \frac{s}{a^6}b(-\lambda)^{\frac{2b}{a}}\left(ae^{\lambda}(-\lambda)^{\frac{b}{a}}(a(c+\lambda)-b)+b(b-ac)\gamma\left(\frac{b}{a},-\lambda\right)\right)\\
&\times \frac{\left(b^2(b-ac)t\gamma\left(\frac{b}{a},-\lambda\right)\right)+a(-\lambda)^{\frac{b}{a}}\left(-be^{\lambda}t(b-a(c+\lambda))-a^2c\,{}_2F_2\left(\frac{b}{a},1+c;1+\frac{b}{a},c;\lambda\right)\right)}{c^2\,{}_2F_2\left(\frac{b}{a},1+c;1+\frac{b}{a},c;\lambda\right)^2}.
\end{aligned}
$$

Note that the expression for $Cov\left(N^w(s),N^w(t)-N^w(s)\right)$ is extremely lengthy, and can barely be simplified beyond the given definition. Thus it is excluded.

7.5.2 $w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$

Theorem 7.22. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$ *then*

$$w_t(n;\phi) = \frac{{}_1F_1\left(1+n, n+r; \frac{\lambda(1-t)}{p}\right)}{p(1-p)\binom{n+r-1}{n}}.$$

$$f_{w,t}(n) = \frac{(\lambda t)^n \Gamma(r-1)\, {}_1F_1\left(1+n, n+r; \frac{\lambda(1-t)}{p}\right)}{e^{\frac{\lambda}{p}} p^n \Gamma(n+r) \gamma\left(r-1, \frac{\lambda}{p}\right)}.$$

$$g_t(z) = \frac{e^{\frac{\lambda t(z-1)}{p}} \left(\frac{\lambda}{p}\right)^{r-1} \gamma\left(r-1, \frac{1+\lambda t(z-1)}{p}\right)}{\left(\frac{1+\lambda t(z-1)}{p}\right)^{r-1} \gamma\left(r-1, \frac{\lambda}{p}\right)}.$$

$$E(N^w(t)) = \frac{e^{-\frac{t}{p}} t \left(p^2 (r-1) \left(\frac{\lambda}{p}\right)^r + e^{\frac{\lambda}{p}} \lambda (p - pr + \lambda) \left(\Gamma(r) - (r-1)\Gamma\left(r-1, \frac{\lambda}{p}\right)\right)\right)}{p(r-1)\lambda \gamma\left(r-1, \frac{\lambda}{p}\right)}.$$

$$\begin{aligned} Var(N^w(t)) &= \frac{e^{-\frac{2t}{p}} t \left(e^{\frac{\lambda}{p}} p\lambda (r-1) \left(\frac{\lambda}{p}\right)^r (p + pt(r-2) - t\lambda) \left(\Gamma(r) - (r-1)\Gamma\left(r-1, \frac{\lambda}{p}\right)\right)\right)}{p\lambda^2 (r-1)^2 \gamma\left(r-1, \frac{\lambda}{p}\right)^2} \\ &+ \frac{e^{-\frac{2t}{p}} t \left(e^{\frac{2t}{p}} \lambda^2 (p(r-1)(t-1) + \lambda) \left(\Gamma(r) - (r-1)\Gamma\left(r-1, \frac{\lambda}{p}\right)\right)^2 - p^3 (r-1)^2 t \left(\frac{\lambda}{p}\right)^{2r}\right)}{p\lambda^2 (r-1)^2 \gamma\left(r-1, \frac{\lambda}{p}\right)^2}. \end{aligned}$$

$$g_{t-s}(z) = \frac{e^{\frac{\lambda(t-s)(z-1)}{p}} \left(\frac{\lambda}{p}\right)^{r-1} \gamma\left(r-1, \frac{\lambda(1+s(1-z)+t(z-1))}{p}\right)}{\left(\frac{\lambda(1+s(1-z)+t(z-1))}{p}\right)^{r-1} \gamma\left(r-1, \frac{\lambda}{p}\right)}.$$

$$f_{w,(s,t)}(x,y) = \binom{y}{x} \frac{\left(1-\frac{s}{t}\right)^{y-x} \left(\frac{s}{t}\right)^x (\lambda t)^y \left(\frac{\lambda}{p}\right)^{r-1} \Gamma(r-1)\, {}_1F_1\left(1+y, y+r; \frac{\lambda(1-t)}{p}\right)}{e^{\frac{\lambda}{p}} p^y \Gamma(r+y) \gamma\left(r-1, \frac{\lambda}{p}\right)}.$$

$$
\begin{aligned}
Cov_{s,t} &= \frac{e^{-\frac{2t}{p}} s\left(e^{\frac{\lambda}{p}} p\lambda(r-1)\left(\frac{\lambda}{p}\right)^r (p+pt(r-2)-t\lambda)\left(\Gamma(r)-(r-1)\Gamma\left(r-1,\frac{\lambda}{p}\right)\right)\right)}{p\lambda^2(r-1)^2\gamma\left(r-1,\frac{\lambda}{p}\right)^2} \\
&+ \frac{e^{-\frac{2t}{p}} s\left(e^{\frac{2t}{p}} \lambda^2(p(r-1)(t-1)+\lambda)\left(\Gamma(r)-(r-1)\Gamma\left(r-1,\frac{\lambda}{p}\right)\right)^2 - p^3(r-1)^2 t\left(\frac{\lambda}{p}\right)^{2r}\right)}{p\lambda^2(r-1)^2\gamma\left(r-1,\frac{\lambda}{p}\right)^2}.
\end{aligned}
$$

$$
\begin{aligned}
Cov_{s,t-s} &= \frac{e^{-\frac{2t}{p}} s(t-s)\left(e^{\frac{\lambda}{p}} \lambda\left(\Gamma(r)-(r-1)\Gamma\left(r-1,\frac{\lambda}{p}\right)\right)(p(r-2)-\lambda)\left(\frac{\lambda}{p}\right)^r\right)}{\lambda^2(r-1)\gamma\left(r-1,\frac{\lambda}{p}\right)^2} \\
&+ \frac{e^{-\frac{2t}{p}} s(t-s)\left(e^{\frac{\lambda}{p}} \lambda\left(\Gamma(r)-(r-1)\Gamma\left(r-1,\frac{\lambda}{p}\right)\right) e^{\frac{\lambda}{p}} \lambda\left(\Gamma(r)-(r-1)\Gamma\left(r-1,\frac{\lambda}{p}\right)\right)\right)}{\lambda^2(r-1)\gamma\left(r-1,\frac{\lambda}{p}\right)^2} \\
&- \frac{e^{-\frac{2t}{p}} s(t-s)\left(\left(p^2(r-1)\left(\frac{\lambda}{p}\right)^{2r}\right)(p(r-2)-\lambda)\left(\frac{\lambda}{p}\right)^r\right)}{\lambda^2(r-1)\gamma\left(r-1,\frac{\lambda}{p}\right)^2} \\
&- \frac{e^{-\frac{2t}{p}} s(t-s)\left(\left(p^2(r-1)\left(\frac{\lambda}{p}\right)^{2r}\right) e^{\frac{\lambda}{p}} \lambda\left(\Gamma(r)-(r-1)\Gamma\left(r-1,\frac{\lambda}{p}\right)\right)\right)}{\lambda^2(r-1)\gamma\left(r-1,\frac{\lambda}{p}\right)^2}.
\end{aligned}
$$

7.5.3 $w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$

Theorem 7.23. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$ *then*

$$
\begin{aligned}
w_t(n;\phi) &= \frac{\Gamma(n+a+1)_1F_1(n+a+1;n;\lambda(1-t))}{a^2\Gamma(a)\Gamma(n)}. \\
f_{w,t}(n) &= \frac{t^n\lambda^{n-1}\Gamma(n+a+1)_1F_1(n+a+1;n;\lambda(1-t))}{\Gamma(a+2)\Gamma(n+1)\Gamma(n)_1F_1(2+a;2;\lambda)}. \\
g_t(z) &= \frac{(1+t(z-1))_1F_1(2+a;2;\lambda(1+t(z-1)))}{_1F_1(2+a;2;\lambda)}. \\
E(N^w(t)) &= t+\frac{(2+a)\lambda t_1F_1(3+a;3;\lambda)}{2_1F_1(2+a;2;\lambda)}. \\
Var(N^w(t)) &= \frac{t^2\lambda(2+a)(6_1F_1(3+a;3;\lambda)+(3+a)\lambda_1F_1(4+a;4;\lambda))}{6_1F_1(2+a;2;\lambda)} \\
&+ E(N^w(t)) - (E(N^w(t)))^2. \\
g_{t-s}(z) &= \frac{(1+s(1-z)+t(z-1))_1F_1(2+a;2;\lambda(1+s(1-z)+t(z-1)))}{_1F_1(2+a;2;\lambda)}. \\
f_{w,(s,t)}(x,y) &= \begin{pmatrix} y \\ x \end{pmatrix}\frac{(1-\frac{s}{t})^{y-x}(\frac{s}{t})^x t^y\lambda^{y-1}\Gamma(1+a+y)_1F_1(1+a+y,y;\lambda(1-t))}{\Gamma(2+a)\Gamma(1+y)\Gamma(y)_1F_1(2+a;2;\lambda)}. \\
Cov_{s,t} &= \frac{s(2(2+a)\lambda_1F_1(2+a;2;\lambda)(3_1F_1(3+a;3;\lambda)+(3+a)\lambda t_1F_1(4+a;4;\lambda)))}{12_1F_1(2+a;2;\lambda)^2} \\
&- \frac{s\Big(3(2+a)^2t\lambda^2{}_1F_1(3+a;3;\lambda)^2(3_1F_1(3+a;3;\lambda)+(3+a)\lambda t_1F_1(4+a;4;\lambda))\Big)}{12_1F_1(2+a;2;\lambda)^2} \\
&- \frac{s\Big(12(t-1)_1F_1(2+a;2;\lambda)^2(3_1F_1(3+a;3;\lambda)+(3+a)\lambda t_1F_1(4+a;4;\lambda))\Big)}{12_1F_1(2+a;2;\lambda)^2}. \\
Cov_{s,t-s} &= \frac{s(s-t)\Big((4-4a\lambda)_1F_1(2+a;2;\lambda)^2+a^2\lambda^2{}_1F_1(2+a;3;\lambda)^2\Big)}{4_1F_1(2+a;2;\lambda)^2} \\
&+ \frac{s(s-t)2a\lambda(2+\lambda)_1F_1(2+a;2;\lambda)_1F_1(2+a;3;\lambda)}{4_1F_1(2+a;2;\lambda)^2}.
\end{aligned}
$$

7.5.4 $w(n;\phi) = \left(\frac{-1}{\ln(1-p)}\frac{p^n}{n}\right)^{-1}$

Theorem 7.24. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \left(\frac{-1}{\ln(1-p)}\frac{p^n}{n}\right)^{-1}$ *then*

$$
\begin{aligned}
w_t(n;\phi) &= \frac{e^{\frac{\lambda(1-t)}{p}}(\lambda t-\lambda-np)\ln(1-p)}{p^{n+1}}.\\
f_{w,t}(n) &= \frac{e^{-\frac{\lambda t}{p}}t^n\lambda^{n-1}(np+\lambda-\lambda t)}{p^n n!}.\\
g_t(z) &= e^{\frac{\lambda t(z-1)}{p}}\left(1+t(z-1)\right).\\
E(N^w(t)) &= \frac{t(p+\lambda)}{p}.\\
Var(N^w(t)) &= \frac{t(\lambda+p-pt)}{p}.\\
g_{t-s}(z) &= e^{\frac{\lambda(t-s)(z-1)}{p}}\left(1+s(1-z)+t(z-1)\right).\\
f_{w,(s,t)}(x,y) &= \binom{y}{x}\frac{\left(1-\frac{s}{t}\right)^{y-x}\left(\frac{s}{t}\right)^x t^y\lambda^{y-1}e^{-\frac{\lambda t}{p}}(py+\lambda-t\lambda)}{y!}.\\
Cov_{s,t} &= \frac{s(p-pt+\lambda)}{p}.\\
Cov_{s,t-s} &= s(s-t).
\end{aligned}
$$

7.5.5 $w(n;\phi) = \left(\frac{\Gamma(r+n)}{n!\Gamma(r)}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$

Theorem 7.25. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \left(\frac{\Gamma(r+n)}{n!\Gamma(r)}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$ *then*

$$
\begin{aligned}
w_t(n;\phi) &= \frac{Beta(a,b)\Gamma(1+n)\Gamma(r)\,{}_2F_2(1+n,a+b+n+r;b+n,n+r;\lambda(1-t))}{\Gamma(n+r)Beta(a+r,b+n)}.\\
f_{w,t}(n) &= \frac{(\lambda t)^n Beta(a+r,b)\Gamma(r)\,{}_2F_2(1+n,a+b+n+r;b+n,n+r;\lambda(1-t))}{\Gamma(n+r)Beta(a+r,b+n)\,{}_2F_2(1,a+b+r;b,r;\lambda)}.\\
g_t(z) &= \frac{{}_2F_2(1,a+b+r;b,r;\lambda(1+t(z-1)))}{{}_2F_2(1,a+b+r;b,r;\lambda)}.\\
E(N^w(t)) &= \frac{(a+b+r)\lambda t\,{}_2F_2(2,1+a+b+r;1+b,1+r;\lambda)}{br\,{}_2F_2(1,a+b+r;b,r;\lambda)}.\\
Var(N^w(t)) &= \frac{2(a+b+r)(1+a+b+r)(\lambda t)^2\,{}_2F_2(3,2+a+b+r;2+b,2+r;\lambda)}{b(1+b)r(1+r)\,{}_2F_2(1,a+b+r;b,r;\lambda)}\\
&+ E(N^w(t)) - (E(N^w(t)))^2.\\
g_{t-s}(z) &= \frac{{}_2F_2(1,a+b+r;b,r;\lambda(1+s(1-z)+t(z-1)))}{{}_2F_2(1,a+b+r;b,r;\lambda)}.\\
f_{w,(s,t)}(x,y) &= \binom{y}{x}\frac{\left(1-\frac{s}{t}\right)^{y-x}\left(\frac{s}{t}\right)^x(\lambda t)^y Beta(a+r,b)\Gamma(r)}{Beta(a+r,b+y)\Gamma(r+y)}\\
&\times \frac{{}_2F_2(1+y,a+b+r+y;b+y,r+y;\lambda(1-t))}{{}_2F_2(1,a+b+r;b,r;\lambda)}.\\
Cov_{s,t} &= \frac{(a+b+r)st\left(-(a+b+r)t\lambda\,{}_2F_2(2,1+a+b+r;1+b,1+r;\lambda)^2\right)}{(br)^2\,{}_2F_2(1,a+b+r;b,r;\lambda)^2}\\
&+ \frac{{}_2F_2(1,a+b+r;b,r;\lambda)\,{}_2F_2(2,1+a+b+r;1+b,1+r;\lambda)}{br\,{}_2F_2(1,a+b+r;b,r;\lambda)^2}\\
&+ \frac{2(1+a+b+r)t\lambda\,{}_2F_2(1,a+b+r;b,r;\lambda)\,{}_2F_2(3,2+a+b+r;2+b,2+r;\lambda)}{(b+b^2)(r+r^2)\,{}_2F_2(1,a+b+r;b,r;\lambda)^2}.\\
Cov_{s,t-s} &= \frac{(a+b+r)s(s-t)\lambda^2\left(-(a+b+r)\,{}_2F_2(2,1+a+b+r;1+b,1+r;\lambda)^2\right)}{(br)^2\,{}_2F_2(1,a+b+r;b,r;\lambda)^2}\\
&+ \frac{(a+b+r)s(s-t)\lambda^2 2(1+a+b+r)\,{}_2F_2(1,a+b+r;b,r;\lambda)\,{}_2F_2(3,2+a+b+r;2+b,2+r;\lambda)}{(b+b^2)(r+r^2)\,{}_2F_2(1,a+b+r;b,r;\lambda)^2}.
\end{aligned}
$$

7.5.6 $w(n;\phi) = \binom{n}{a}^{-1}$

Theorem 7.26. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \binom{n}{a}^{-1}$ *then*

$$
\begin{aligned}
w_t(n;\phi) &= \frac{{}_1F_1(1-a+n;1+n;\lambda(1-t))}{\binom{n}{a}}. \\
f_{w,t}(n) &= \frac{t^n\lambda^{n-a}a!\,{}_1F_1(1-a+n;1+n;\lambda(1-t))}{\binom{n}{a}n!\,{}_2F_2(1,1;1+a,1+a;\lambda)}. \\
g_t(z) &= \frac{(1+t(z-1))^a\,{}_2F_2(1,1;1+a,1+a;\lambda(1+t(z-1)))}{{}_2F_2(1,1;1+a,1+a;\lambda)}. \\
E(N^w(t)) &= \frac{t\left(a(1+a)^2\,{}_2F_2(1,1;1+a,1+a;\lambda)+\lambda\,{}_2F_2(2,2;2+a,2+a;\lambda)\right)}{(1+a)^2\,{}_2F_2(1,1;1+a,1+a;\lambda)}. \\
Var(N^w(t)) &= \frac{t^2a(a-1)\,{}_2F_2(1,1;1+a,1+a;\lambda)}{{}_2F_2(1,1;1+a,1+a;\lambda)} \\
&+ \frac{2t^2\lambda\left(a(2+a)^2\,{}_2F_2(2,2;2+a,2+a;\lambda)+2\lambda\,{}_2F_2(3,3;3+a,3+a;\lambda)\right)}{(1+a)^2(2+a)^2\,{}_2F_2(1,1;1+a,1+a;\lambda)} \\
&+ E(N^w(t)) - (E(N^w(t)))^2. \\
g_{t-s}(z) &= \frac{(1+s(1-z)+t(z-1))^a\,{}_2F_2(1,1;1+a,1+a;\lambda(1+s(1-z)+t(z-1)))}{{}_2F_2(1,1;1+a,1+a;\lambda)}. \\
f_{w,(s,t)}(x,y) &= \binom{y}{x}\binom{y}{a}^{-1}\frac{(1-\frac{s}{t})^{y-x}(\frac{s}{t})^x t^y\lambda^{y+a}a!\,{}_1F_1(1-a+y;1+y;\lambda(1-t))}{{}_2F_2(1,1;1+a,1+a;\lambda)}. \\
Cov_{s,t} &= sa(1-t) - \frac{st\lambda^2\,{}_2F_2(2,2;2+a,2+a;\lambda)^2}{(1+a)^4\,{}_2F_2(1,1;1+a,1+a;\lambda)^2} \\
&+ \frac{s\lambda\,{}_2F_2(1,1;1+a,1+a;\lambda)\,{}_2F_2(2,2;2+a,2+a;\lambda)}{(1+a)^2\,{}_2F_2(1,1;1+a,1+a;\lambda)^2} \\
&+ \frac{4st\lambda^2\,{}_2F_2(1,1;1+a,1+a;\lambda)\,{}_2F_2(3,3;3+a,3+a;\lambda)}{(1+a)^2(2+a)^2\,{}_2F_2(1,1;1+a,1+a;\lambda)^2}. \\
Cov_{s,t-s} &= \frac{-as(t-s)\Gamma(1+a)^2}{{}_2F_2(1,1;1+a,1+a;\lambda)^2} - \frac{s(t-s)\lambda^2\,{}_2F_2(2,2;2+a,2+a;\lambda)^2}{(1+a)^2\,{}_2F_2(1,1;1+a,1+a;\lambda)^2} \\
&+ \frac{s(t-s)4\,{}_2F_2(3,3;3+a,3+a;\lambda)}{{}_2F_2(1,1;1+a,1+a;\lambda)\Gamma(1+a)^2(2+a)^2}.
\end{aligned}
$$

7.5.7 $w\left(n;\phi\right)=\left(\left(a\right)_{n}\right)^{-1}$

Theorem 7.27. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w\left(n;\phi\right)=\left(\left(a\right)_{n}\right)^{-1}$ *then*

$$
\begin{aligned}
w_t\left(n;\phi\right) &= \tfrac{{}_0F_1(a+n;\lambda(1-t))\Gamma(a)}{\Gamma(a+n)}.\\
f_{w,t}\left(n\right) &= \tfrac{\Gamma(\lambda)(\lambda t)^n{}_0F_1(a+n;\lambda(1-t))}{\Gamma(1+n)\Gamma(a+n){}_0F_1(a;\lambda)}.\\
g_t\left(z\right) &= \tfrac{{}_0F_1(a;\lambda(1+\lambda t(z-1)))}{{}_0F_1(a;\lambda)}.\\
E\left(N^w\left(t\right)\right) &= \tfrac{t\lambda{}_0F_1(1+a;\lambda)}{{}_0F_1(a;\lambda)a}.\\
Var\left(N^w\left(t\right)\right) &= \lambda t^2-\tfrac{(\lambda t)^2{}_0F_1(1+a;\lambda)^2}{a^2{}_0F_1(a;\lambda)^2}+\tfrac{\lambda t(1-at){}_0F_1(1+a;\lambda)}{a{}_0F_1(a;\lambda)}.\\
g_{t-s}\left(z\right) &= \tfrac{{}_0F_1(a;\lambda(1+s(1-z)+t(z-1)))}{{}_0F_1(a;\lambda)}.\\
f_{w,(s,t)}\left(x,y\right) &= \begin{pmatrix} y\\ x\end{pmatrix}\tfrac{\left(1-\frac{s}{t}\right)^{y-x}\left(\frac{s}{t}\right)^{x}(\lambda t)^y{}_0F_1(a+y;\lambda(1-t))\Gamma(a)}{\Gamma(1+y){}_0F_1(a;\lambda)\Gamma(a+y)}.\\
Cov_{s,t} &= \lambda st-\tfrac{st\lambda^2{}_0F_1(1+a;\lambda)^2}{a^2{}_0F_1(a;\lambda)^2}+\tfrac{s\lambda(1-at){}_0F_1(1+a;\lambda)}{a{}_0F_1(a;\lambda)}.\\
Cov_{s,t-s} &= \tfrac{s(s-t)\lambda^2\Gamma(a)^2}{{}_0F_1(a;\lambda)^2}\left(\tfrac{{}_0F_1(1+a;\lambda)^2}{\Gamma(1+a)^2}-\tfrac{{}_0F_1(a;\lambda){}_0F_1(2+a;\lambda)}{\Gamma(a)\Gamma(2+a)}\right).
\end{aligned}
$$

7.5.8 $w(n;\phi) = ((n)_a)^{-1}$

Theorem 7.28. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = ((n)_a)^{-1}$ *then*

$$
\begin{aligned}
w_t(n;\phi) &= \frac{\Gamma(n)\,{}_1F_1(n,n+a;\lambda(1-t))}{\Gamma(n+a)}.\\
f_{w,t}(n) &= \frac{t^n\lambda^{n-1}\,{}_1F_1(n,n+a;\lambda(1-t))}{n\,{}_2F_2(1,1;2,1+a;\lambda)}.\\
g_t(z) &= \frac{(1+t(z-1))\,{}_2F_2(1,1;2,1+a;\lambda(1+t(z-1)))}{{}_2F_2(1,1;2,1+a;\lambda)}.\\
E(N^w(t)) &= \frac{e^\lambda t(\Gamma(1+a)-a\Gamma(a,\lambda))}{\lambda^a\,{}_2F_2(1,1;2,1+a;\lambda)}.\\
Var(N^w(t)) &= \frac{te^{-2a}\left({}_2F_2(1,1;2,1+a;\lambda)\lambda^a\left(at\lambda^a+e^\lambda(1-at+\lambda t)(\Gamma(1+a)-a\Gamma(a,\lambda))\right)\right)}{{}_2F_2(1,1;2,1+a;\lambda)^2}\\
&- \frac{te^{-2a}\left(e^{2\lambda}t(\Gamma(1+a)-a\Gamma(a,\lambda))^2\right)}{{}_2F_2(1,1;2,1+a;\lambda)^2}.\\
g_{t-s}(z) &= \frac{(1+s(1-z)+t(z-1))\,{}_2F_2(1,1;2,1+a;\lambda(1+s(1-z)+t(z-1)))}{{}_2F_2(1,1;2,1+a;\lambda)}.\\
f_{w,(s,t)}(x,y) &= \begin{pmatrix} y \\ x \end{pmatrix} \frac{\left(1-\frac{s}{t}\right)^{y-x}\left(\frac{s}{t}\right)^x t^y\lambda^{y-1}\Gamma(1+a)\,{}_1F_1(y;a+y;\lambda(1-t))}{y\Gamma(a+y)\,{}_2F_2(1,1;2,1+a;\lambda)}.\\
Cov_{s,t} &= \frac{s\lambda^{-2a}\left({}_2F_2(1,1;2,1+a;\lambda)\lambda^a\left(at\lambda^a+e^\lambda(1-at+t\lambda)(\Gamma(1+a)-a\Gamma(a,\lambda))\right)\right)}{{}_2F_2(1,1;2,1+a;\lambda)}\\
&- \frac{s\lambda^{-2a}\left(e^{2\lambda}t(\Gamma(1+a)-a\Gamma(a,\lambda))^2\right)}{{}_2F_2(1,1;2,1+a;\lambda)}.\\
Cov_{s,t-s} &= \frac{s(s-t)\lambda^{-2a}\left({}_2F_2(1,1;2,1+a;\lambda)\lambda^a\left(e^\lambda(\lambda-a)(a\Gamma(a,\lambda)-\Gamma(1+a))-a\lambda^a\right)\right)}{{}_2F_2(1,1;2,1+a;\lambda)}\\
&+ \frac{s(s-t)\lambda^{-2a}\left(e^{2\lambda}t(\Gamma(1+a)-a\Gamma(a,\lambda))^2\right)}{{}_2F_2(1,1;2,1+a;\lambda)}.
\end{aligned}
$$

Chapter 8

Weighted Poisson Process: Applications

In this chapter, the weighted Poisson processes proposed in Chapter 7 will be applied to a range of datasets similar to what was done in Chapter 5.

The number of datasets that will be analysed in this chapter will be smaller than was the case in Chapter 5. The main reason for this is that many of the datasets exhibit cyclical patterns and trends. Typically, when time series are modelled, "seasonal adjustments" need to be made to remove these cycles. To limit this problem in the modelling, the duration of the data analysed has been limited to relatively short time spans.

Theoretically, Poisson and weighted Poisson processes are continuous time stochastic processes, however none of the available datasets have time variables that are fine enough that they can truly be considered "continuous time". For instance, the US gun violence dataset spans 1551 days, but none of the events have an hour, minute or second associated with them. This leads to the practical limitation that the time variable $t \in [0, 1]$ can, at best, be subdivided into 1551 intervals of equal length. As was discussed in Chapter 6, this is a common practical occurrence, and this discrete approximation of the continuous time domain will be present in the upcoming analyses.

The discretisation of time leads to some interesting implications for parameter estimation: A traditional realisation of a process has, at each time, only one value associated with it since two events occurring at exactly the same time is practically impossible. However, since numerous events can happen in the same interval of time, some datasets, like the US gun injuries per incident (shown in Figure 8.1), can have multiple data points at "each time". This has the practical effect that if increments of the processes are modelled, the "sample size" can be larger than one at each time step.

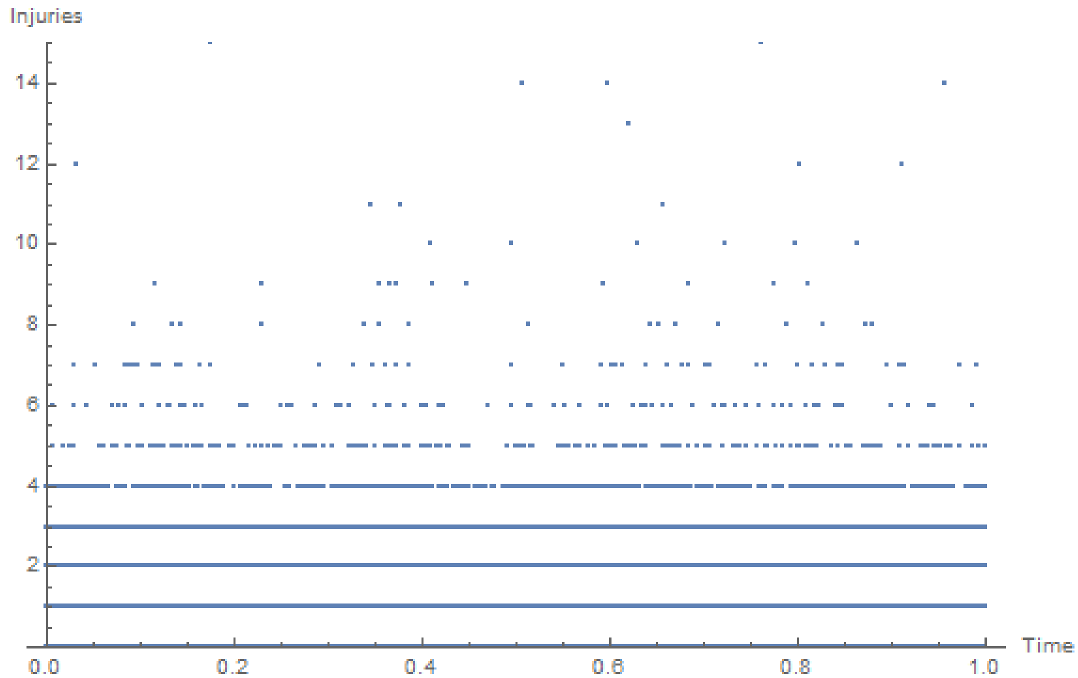


Figure 8.1: US Gun violence - Injuries per incident

For all the weighted Poisson processes discussed in Chapter 7, the maximum likelihood estimates of the parameters are calculated (using the Nelder-Mead algorithm which was described in Chapter 5). Since the time domain is being approximated by many small discrete intervals of equal length it is possible to calculate the maximum likelihood estimates by using the probability mass function of the processes at the various time points. In addition to the maximum log-likelihood, the AIC, AICc and BIC of the fit will also be calculated. (Tables containing these results can be found in Chapter 10.) Of all of the fitted models, the one that has the smallest AIC, AICc and BIC values is selected as the "best model".

While it is possible to plot a "hybrid" graph which combines the discrete variable n and the continuous variable t (see Figure 8.2 as an example), these graphs become extremely difficult to interpret for large ranges of n. As a result, the marginal distribution of the best weighted Poisson process will be plotted (in red) against the Poisson process (in black) at a few t values. For each plot a vertical reference line will be inserted to indicate what the actual value of the process is at that time.

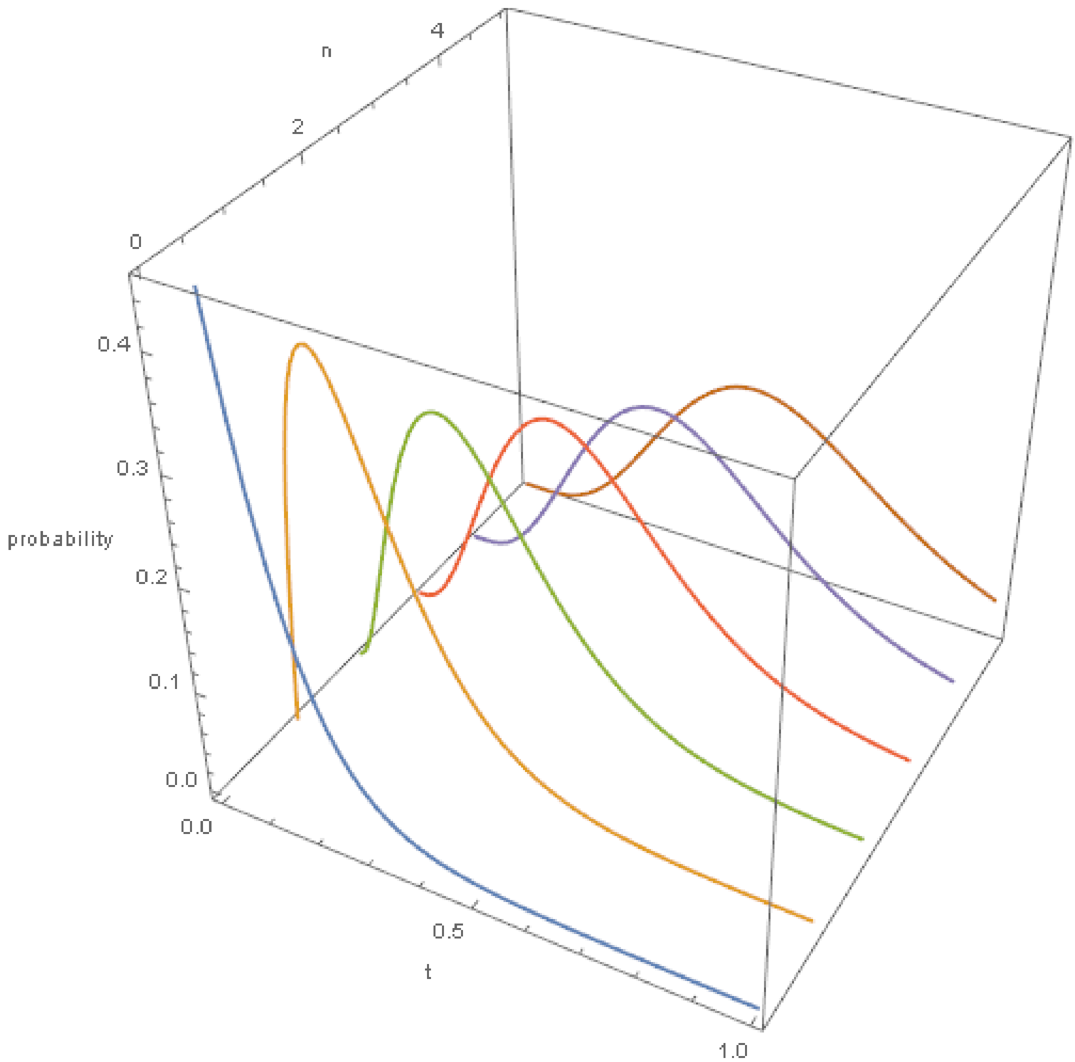


Figure 8.2: Hybrid plot - $Poisson\,(10t)\,; n = 0, 1, 2, ...; t \in [0, 1]$

It should also be noted (as was mentioned earlier in this thesis) that many of the weight functions discussed result in weighted Poisson probability mass functions that are computationally intensive to calculate if n becomes large. When data was modelled from a purely distributional perspective this was a minimal problem since these values were rarely sizable enough to cause major concerns. However, since count processes are non-decreasing as t increases, the number of events in a time period can grow quickly. This is especially true for processes that span long time periods and have high rates of incidents. When certain weighted Poisson processes are too taxing they are excluded from the comparisons.

8.1 Novel data fits

For the Weekly sales - Item 409 data, of the 20 weight functions that were tested, only 1 model performs better than the Poisson process. The best fit is achieved when $w(n;\phi) = (912.948)_n$ and $\lambda = 0.70881$. Figure 8.3 shows the increments of the process and in Figure 8.4 below the probability mass functions of the Poisson and weighted Poisson processes are given for $t = \frac{10}{52}, \frac{20}{52}, \frac{30}{52}, \frac{40}{52}$.

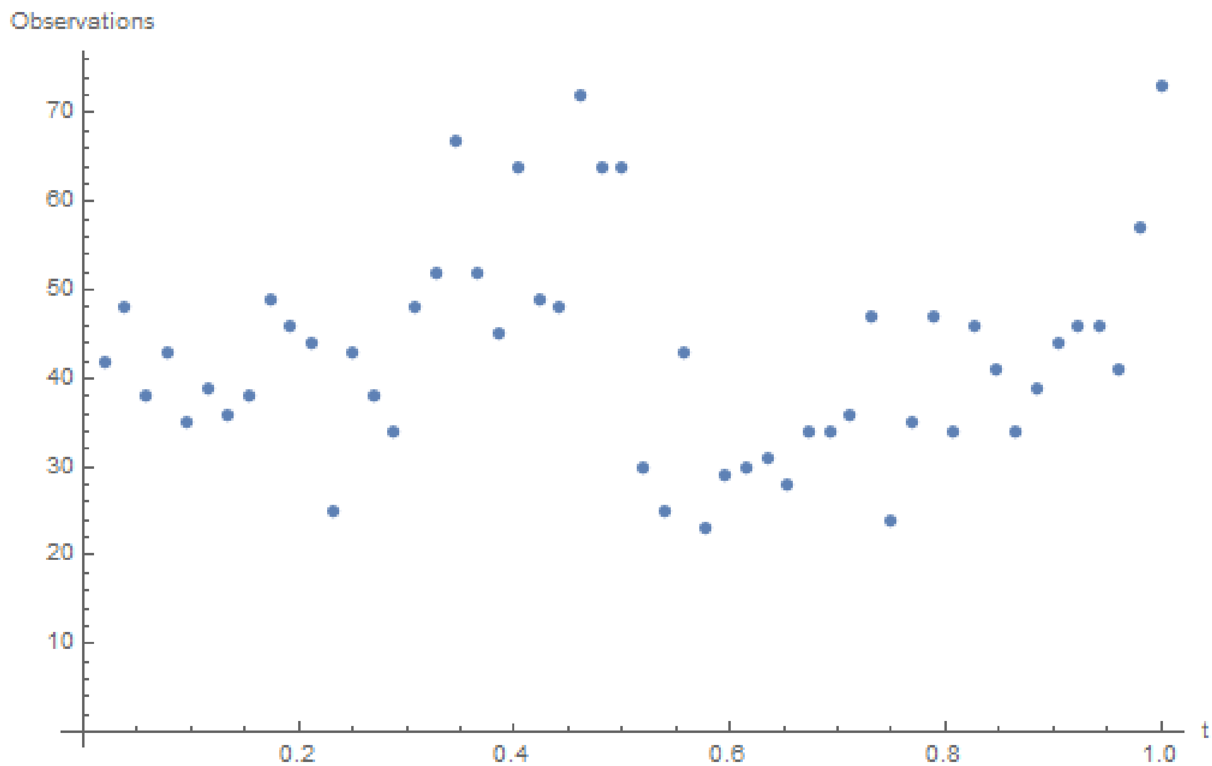


Figure 8.3: Weekly sales - Item 409, increments

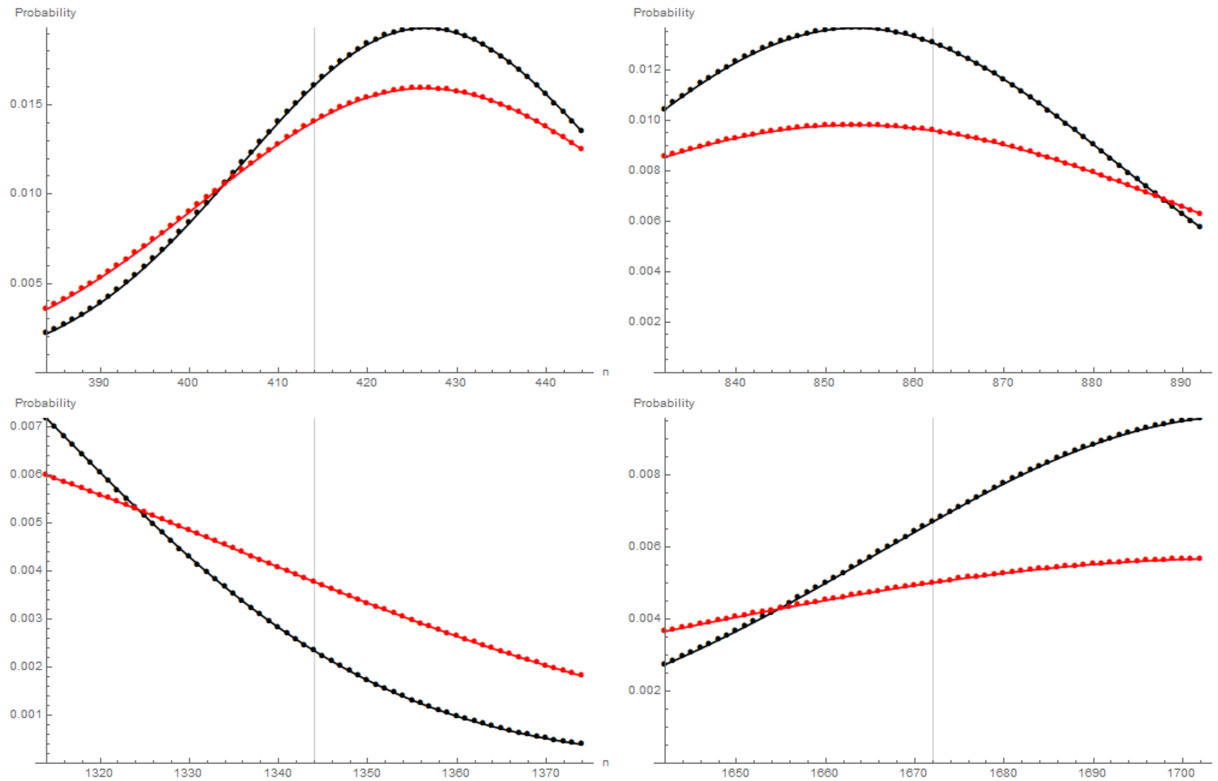


Figure 8.4: Weekly sales - Item 409, fits at various t

To eliminate or minimise the seasonal patterns that are present in nearly all of the datasets discussed thus far, the remaining analyses will be done on small (weekly, monthly or yearly) sections of the previously discussed data. For example, if the number of mass shootings incidents in June 2013 are analysed it is found that of the 25 weighted processes, 5 outperform the Poisson process and 2 additional weighted Poisson processes may. Interestingly, the best fit occurs when $w(n;\phi) = I(n \leq 31)$ and $\lambda = 37.4904$. While there is no practical reason why the number of incidents in a month should be limited to 31, June and July are the months where the most events tend to occur. Thus, this model may actually give insight into how frequently mass shootings occur over monthly time periods. The second best fit is achieved when $w(n;\phi) = \left(1762.3\frac{\Gamma(n)\Gamma(1763.3)}{\Gamma(n+1763.3)}\right)^{-1}$ and $\lambda = 0.422858$. Figure 8.5 shows the increments of the process and in Figure 8.6 below the probability mass functions are given for $t = \frac{5}{30}, \frac{10}{30}, \frac{15}{30}, \frac{20}{30}, \frac{25}{30}, \frac{30}{30}$.

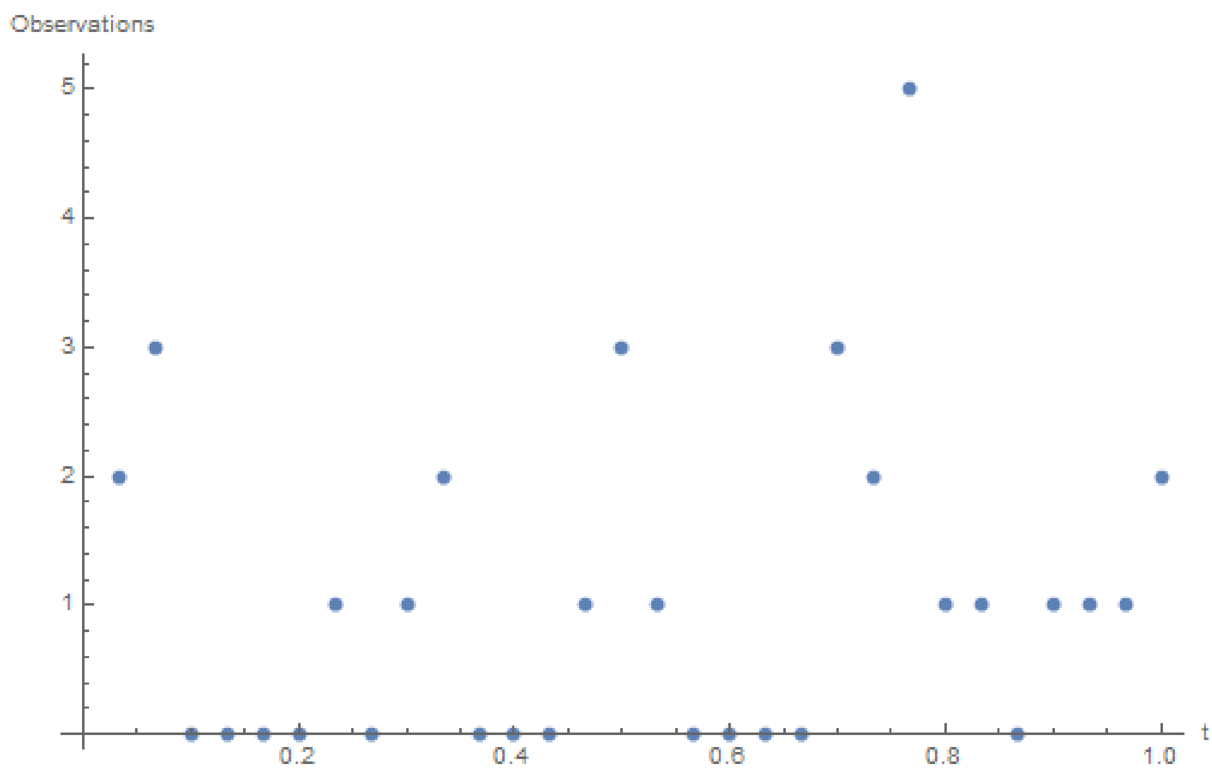


Figure 8.5: Mass shooting incidents - June 2013, increments

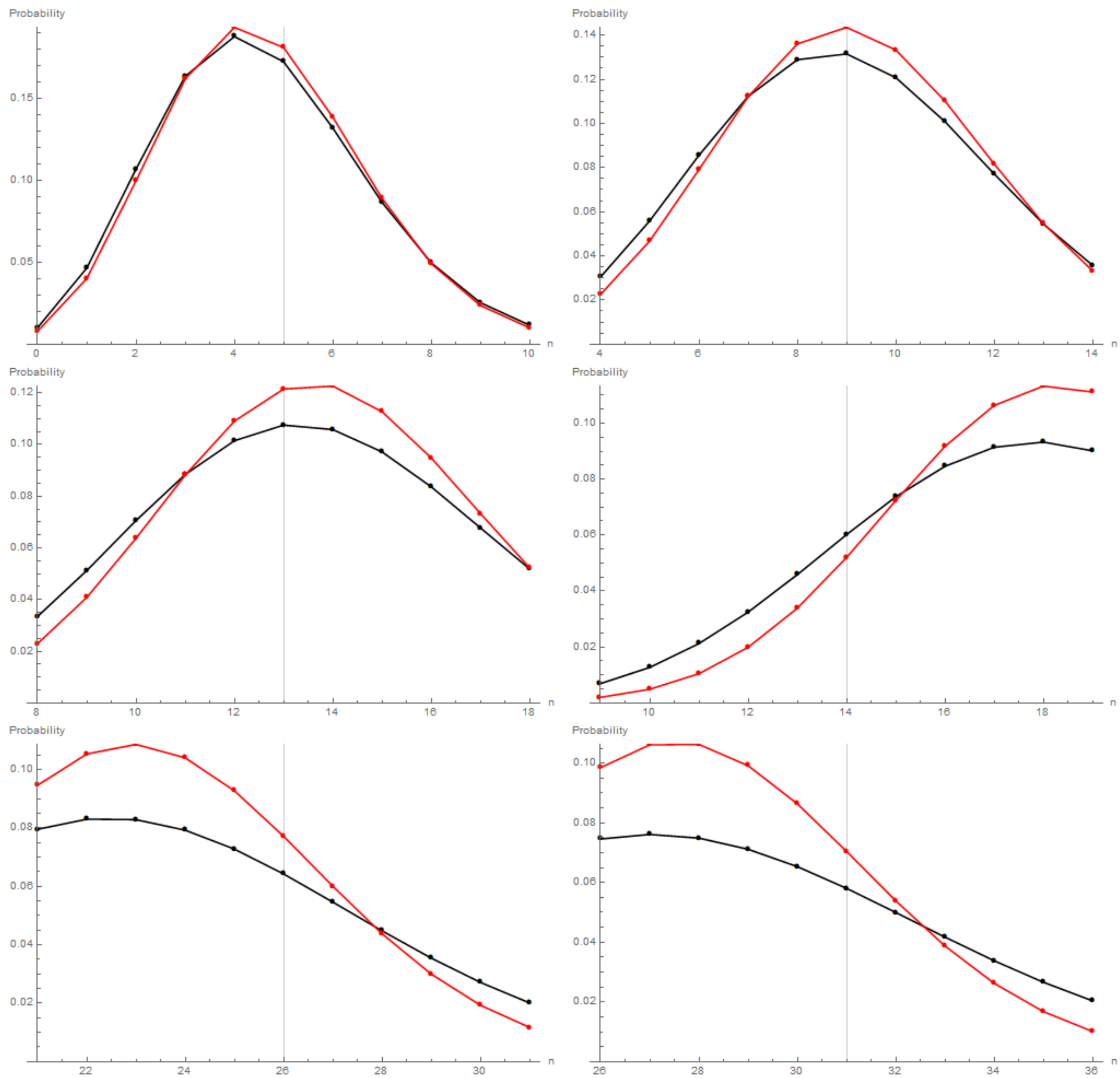


Figure 8.6: Mass shooting incidents - June 2013, fits at various t

When the daily number of gun violence incidents in Rhode Island for 2014 is analysed it is found that out of the 22 weighted Poisson processes, 3 perform better than the Poisson process, and 1 additional weighted Poisson process may be. The best fit is achieved when $w(n;\phi) = \left(\frac{\Gamma(1492.15+n)}{\Gamma(1492.15)n!} \frac{Beta(1492.17,411.724+n)}{Beta(0.017301,411.724)}\right)^{-1}$ and $\lambda = 474.181$. Figure 8.7 shows the increments of the process and in Figure 8.8 below the probability mass functions are given for $t = \frac{60}{365}, \frac{120}{365}, \frac{180}{365}, \frac{240}{365}, \frac{300}{365}, \frac{365}{365}$.

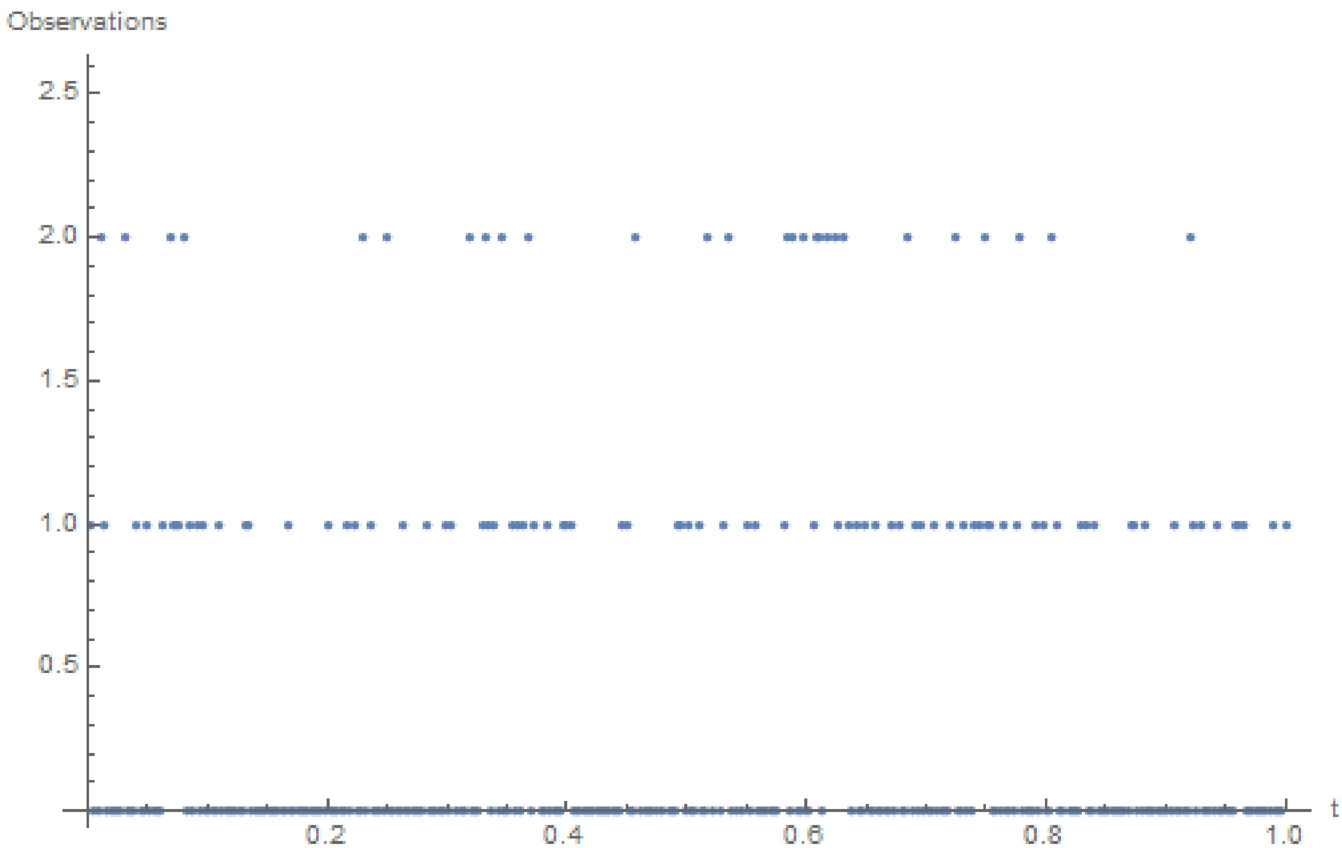


Figure 8.7: US gun violence - Rhode Island, increments

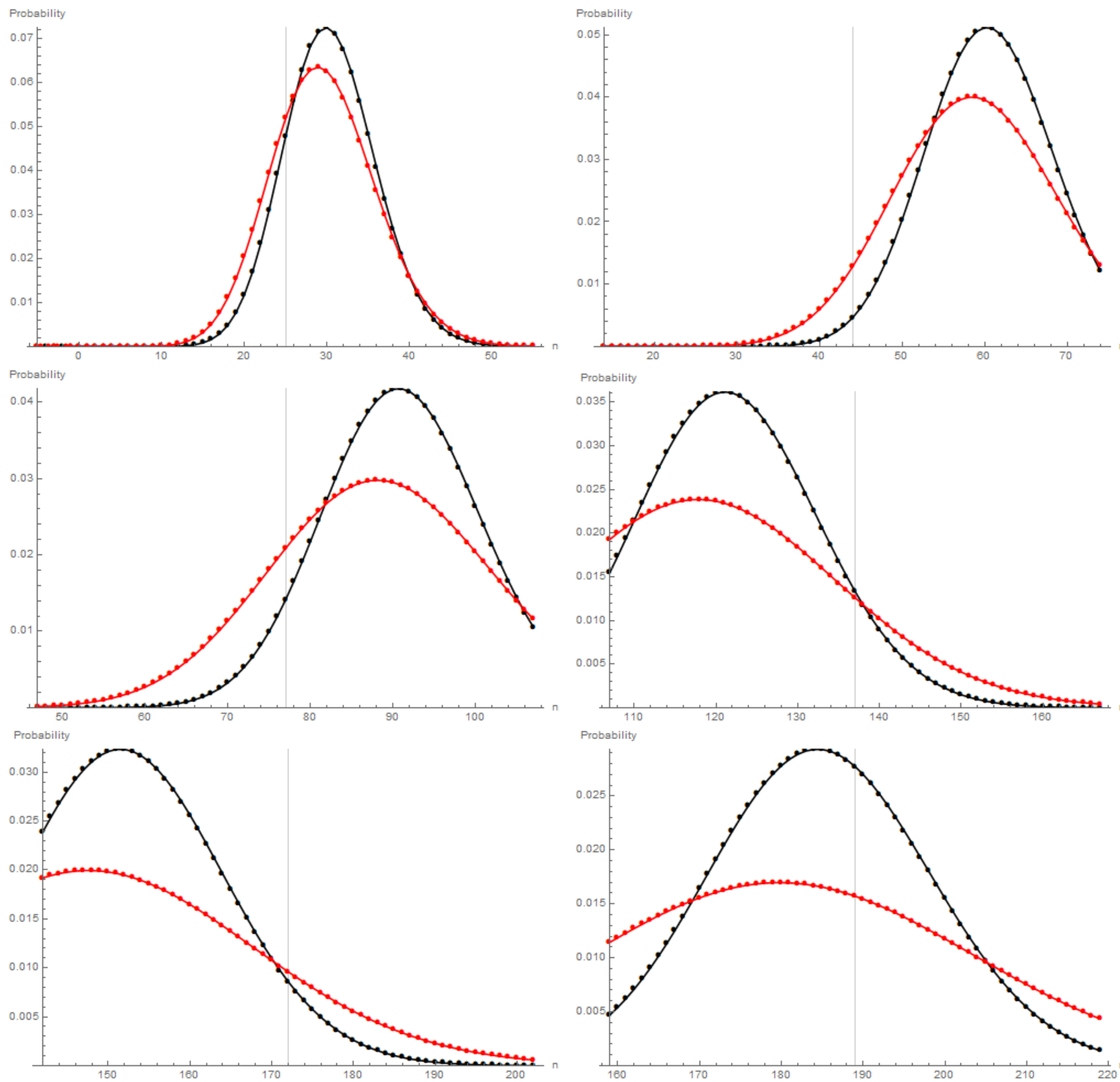


Figure 8.8: Mass shooting incidents, fits at various t

The next dataset contains information about global terrorist attacks between 1970 and 2017. When the number successful bombing attacks in January of 2017 is analysed it is found that out of the 22 weighted Poisson processes, 4 perform better than the Poisson process, and 3 may. The best fit is achieved when $w(n;\phi) = \left(8882.71\frac{\Gamma(n)\Gamma(8883.71)}{\Gamma(n+8883.71)}\right)^{-1}$ and $\lambda = 11.1034$. Figure 8.9 shows the increments of the process and in Figure 8.10 below the probability mass functions are given for $t = \frac{5}{31}, \frac{10}{31}, \frac{15}{31}, \frac{20}{31}, \frac{25}{31}, 1$.

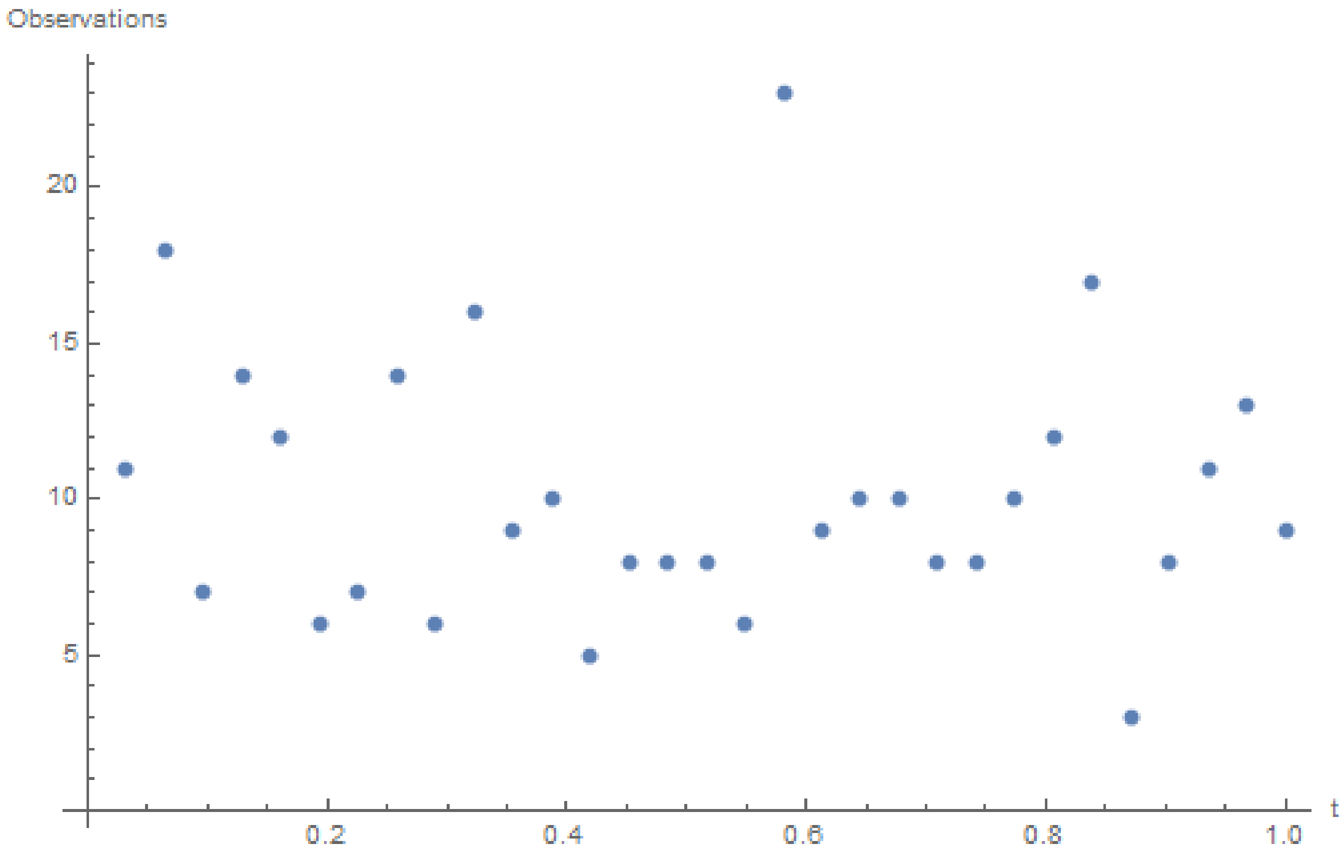


Figure 8.9: Global terrorism - January 2017 successful bombings, increments

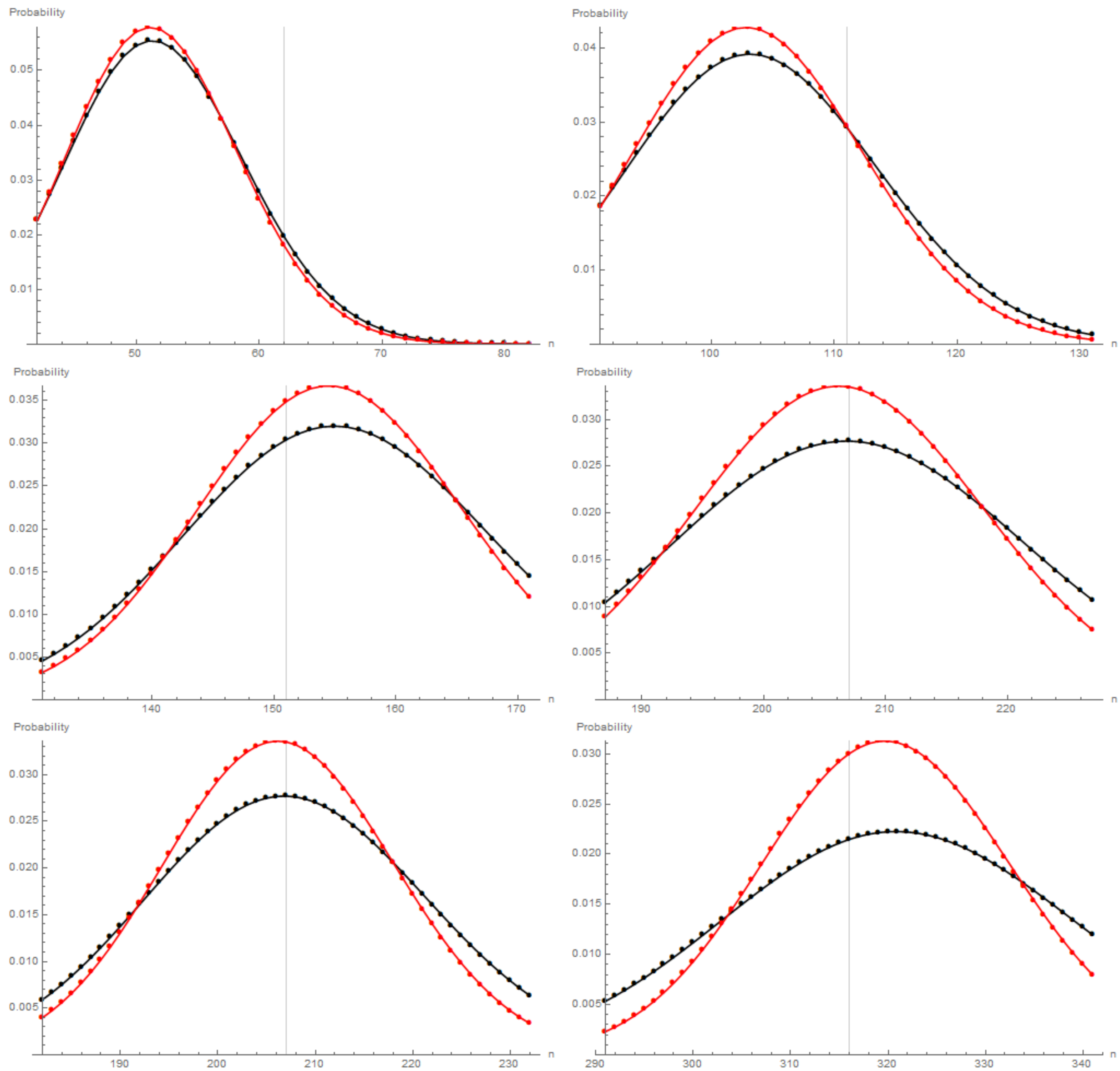


Figure 8.10: Global terrorism - January 2017 successful bombings, fits at various t

The next dataset contains information on accidents reported by the NYPD. When the number injuries to cyclists between January and March of 2016 is analysed it is found that out of the 19 weighted Poisson processes, 2 perform better than the Poisson process. The best fit is achieved when $w(n;\phi) = (196.481)_n$ and $\lambda = 0.733446$. Figure 8.11 shows the increments of the process and in Figure 8.12 below the probability mass functions are given for $t = \frac{15}{91}, \frac{30}{91}, \frac{45}{91}, \frac{60}{91}, \frac{75}{91}, 1$.

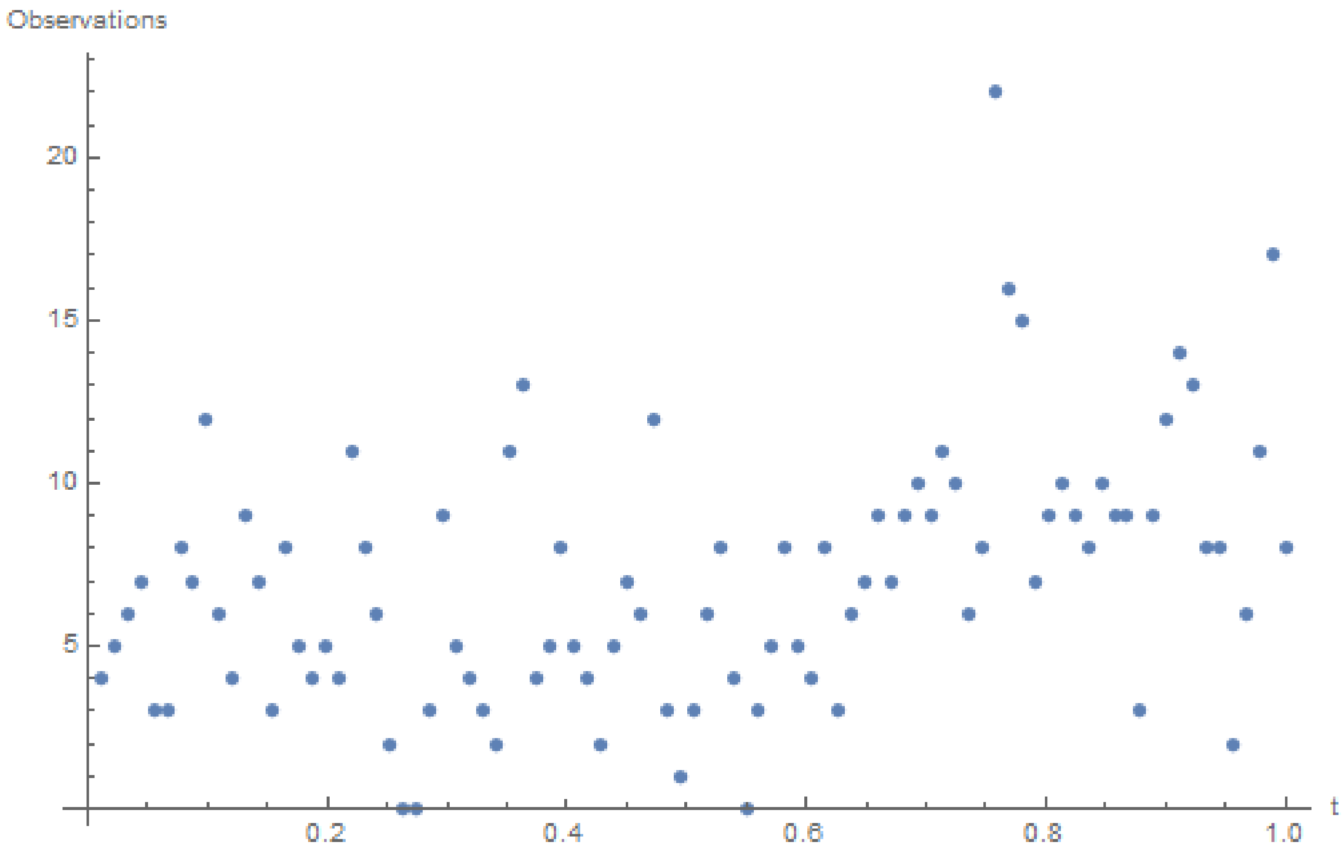


Figure 8.11: NYPD accidents - Cyclist injuries January to March 2016, increments

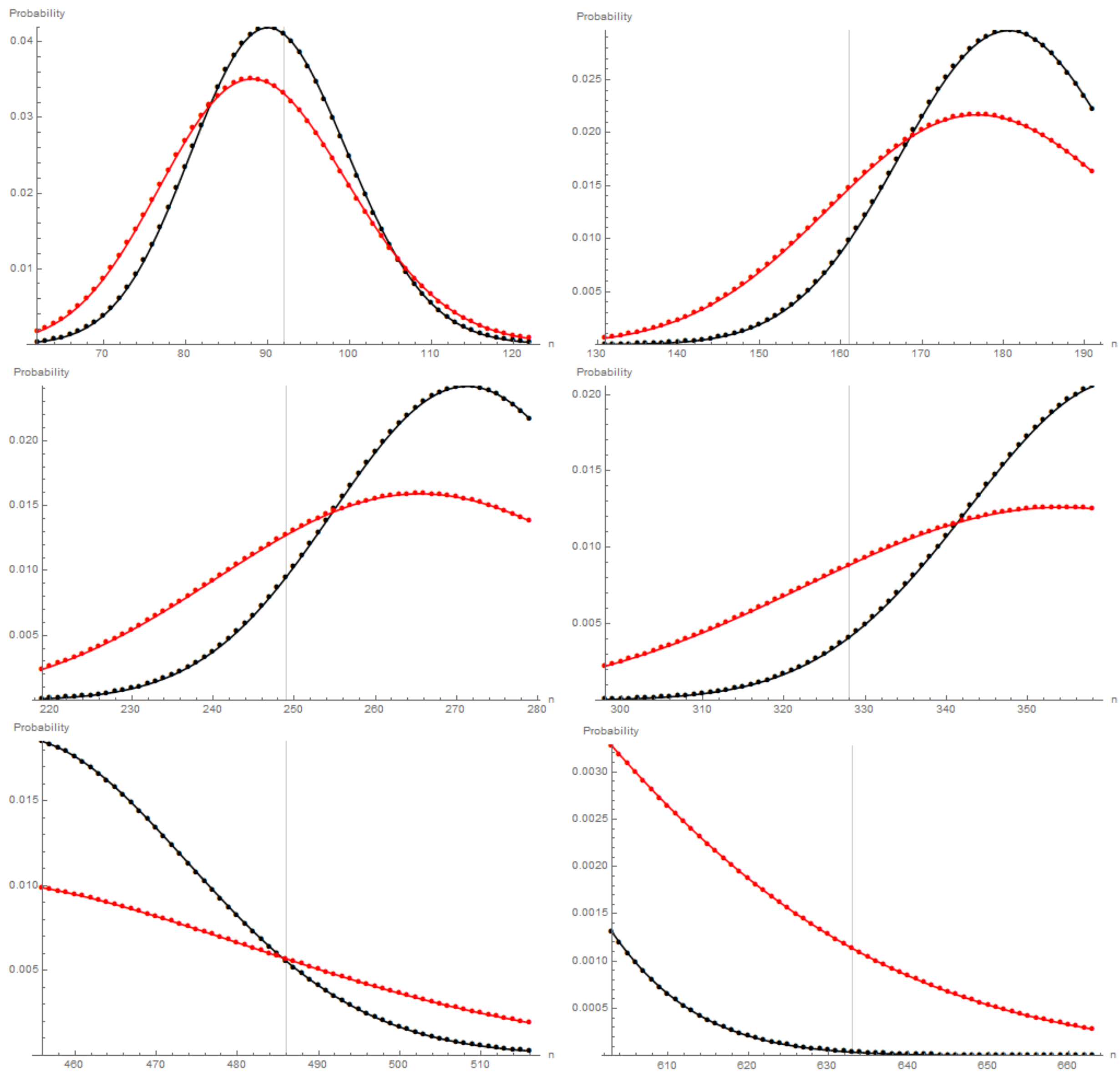


Figure 8.12: NYPD accidents - Cyclist injuries January to March 2016, fits at various t

The next dataset contains information on crimes in Boston between June of 2015 and June of 2018. When the number bomb hoaxes is analysed over the 37 month period it is found that out of the 22 weighted Poisson processes, 5 perform better than the Poisson process. The best fit is achieved when $w(n;\phi) = (278.742)_n$ and $\lambda = 0.790089$. Figure 8.13 shows the increments of the process and in Figure 8.14 below the probability mass functions are given for $t = \frac{6}{37}, \frac{12}{37}, \frac{18}{37}, \frac{24}{37}, \frac{30}{37}, 1$.

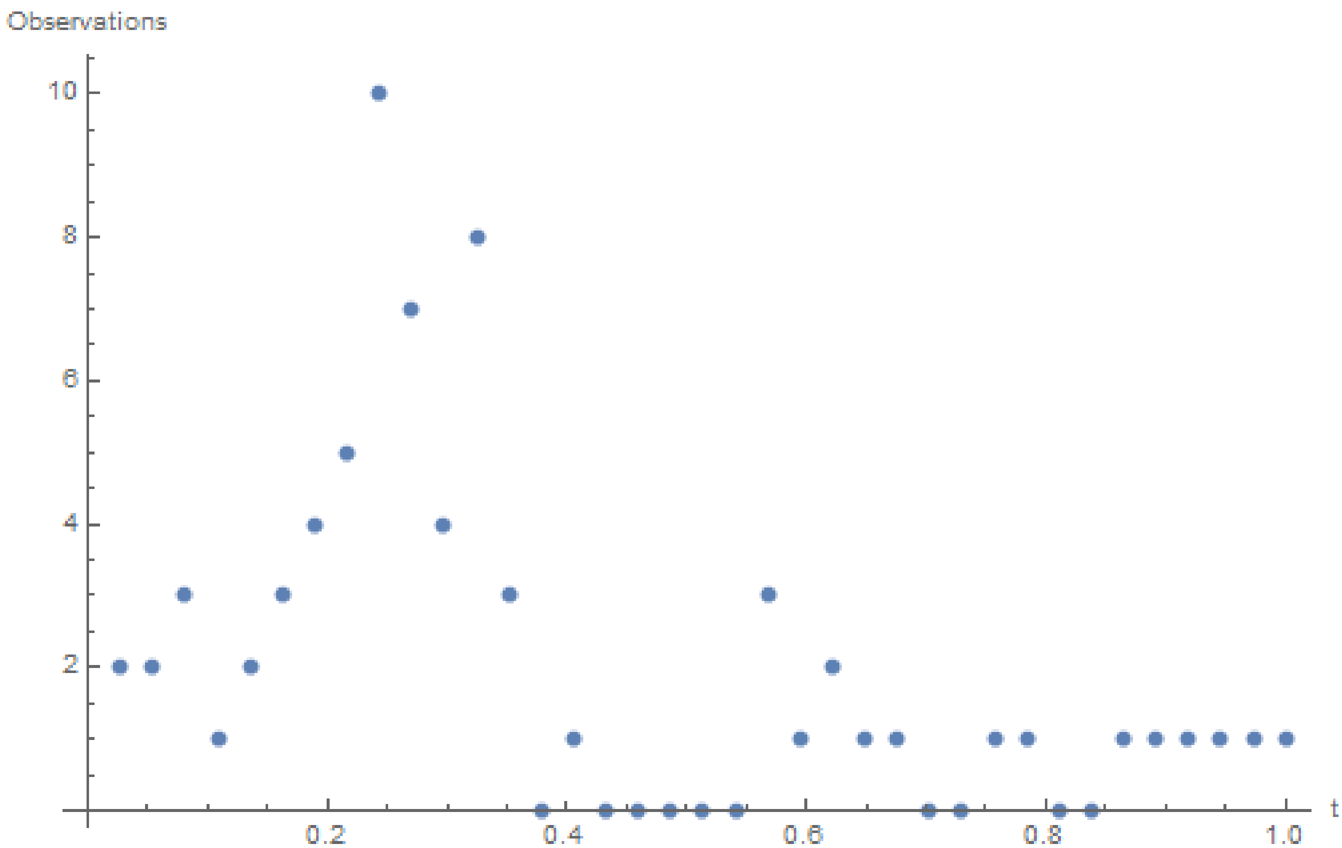


Figure 8.13: Boston crimes - Bomb hoaxes, increments

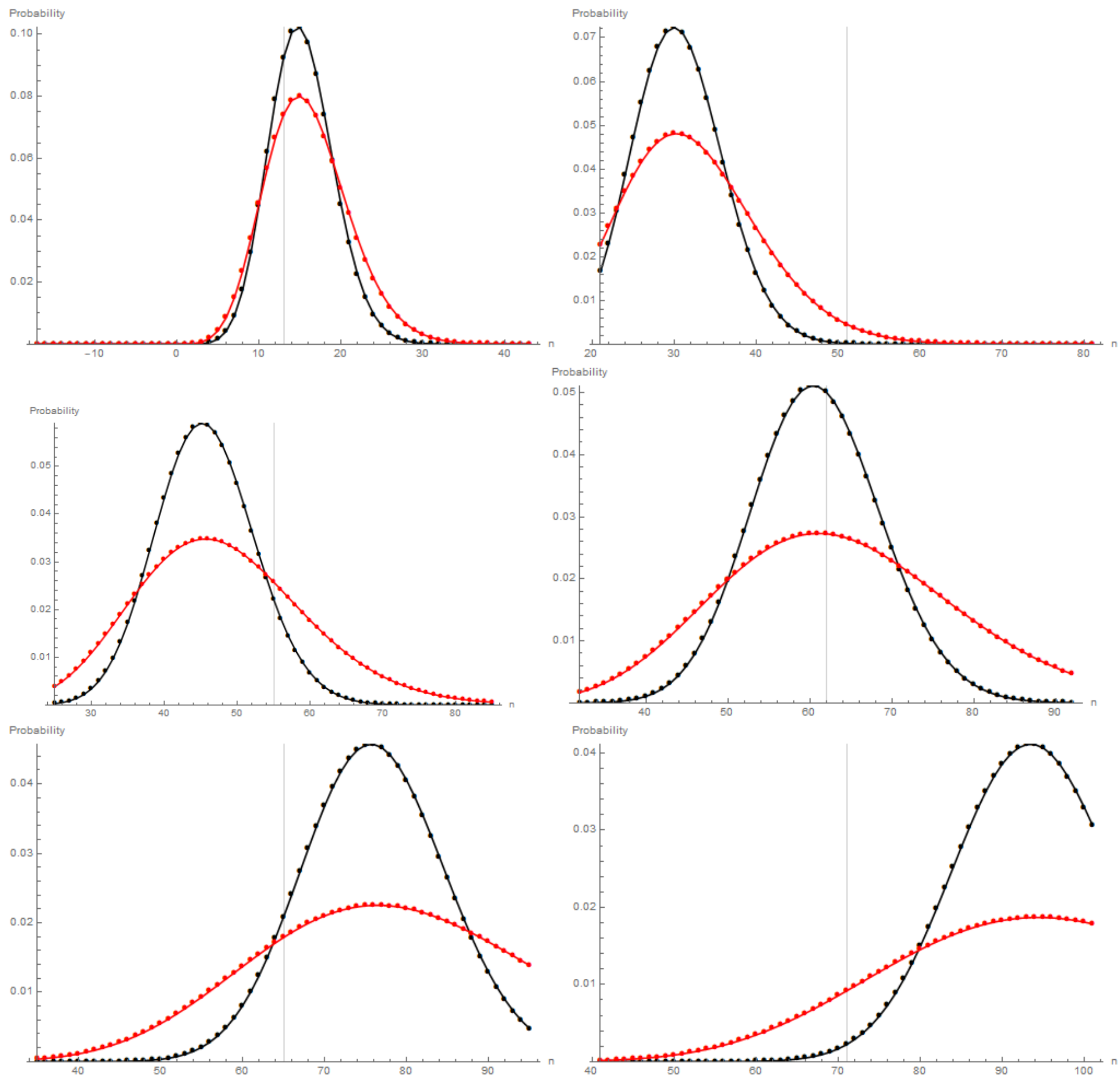


Figure 8.14: Boston crimes - Bomb hoaxes, fits at various t

The final dataset contains weekly unemployment information in Barry County, Missouri for 2019. When the number of unemployment claims is analysed over the 52 week period it is found that out of the 19 weighted Poisson processes, 2 perform better than the Poisson process. The best fit is achieved when $w(n;\phi) = (23.8434)_n$ and $\lambda = 0.0.958732$. Figure 8.15 shows the increments of the process and in Figure 8.16 below the probability mass functions are given for $t = \frac{10}{52}, \frac{20}{52}, \frac{30}{52}, \frac{40}{52}, 1$.

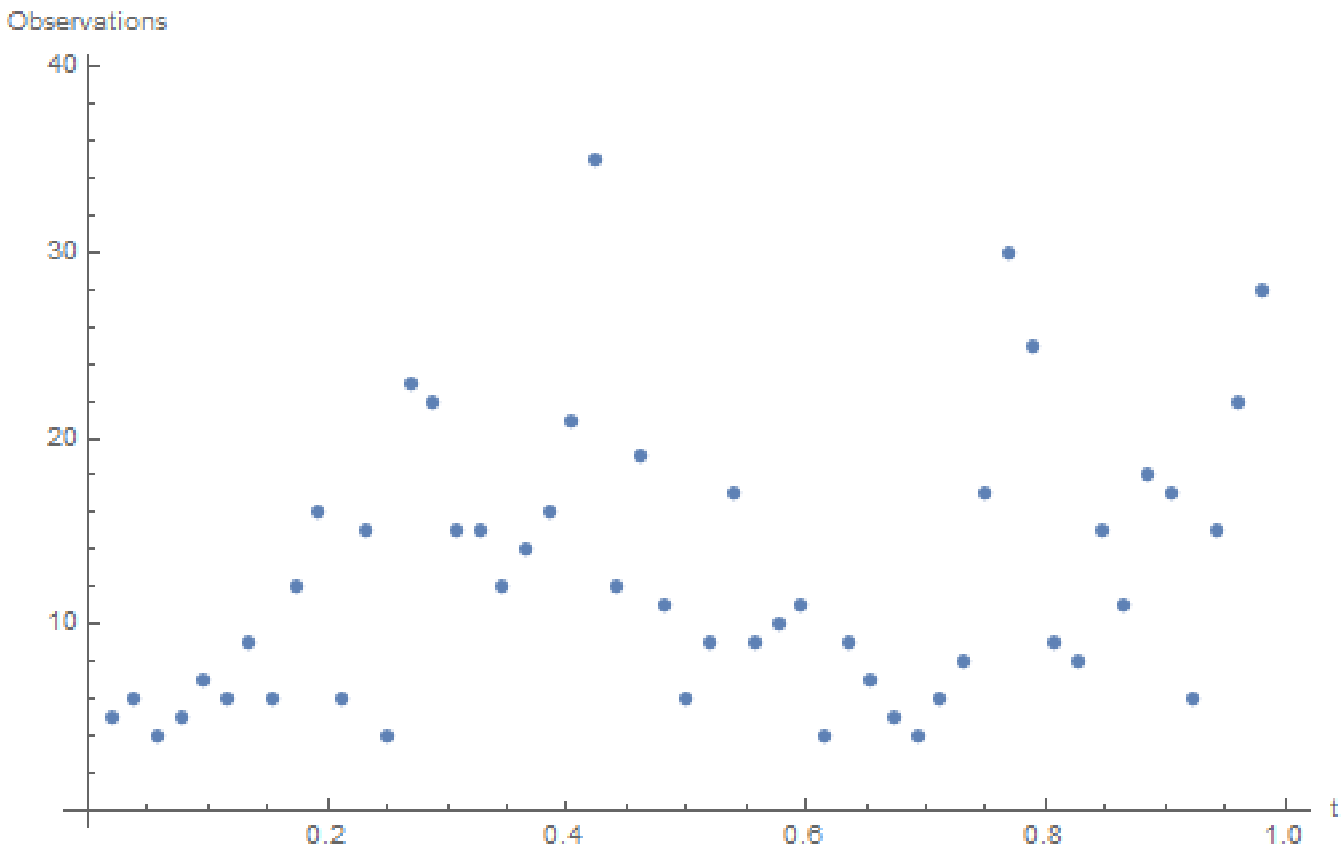


Figure 8.15: Barry County - Unemployment claims, increments

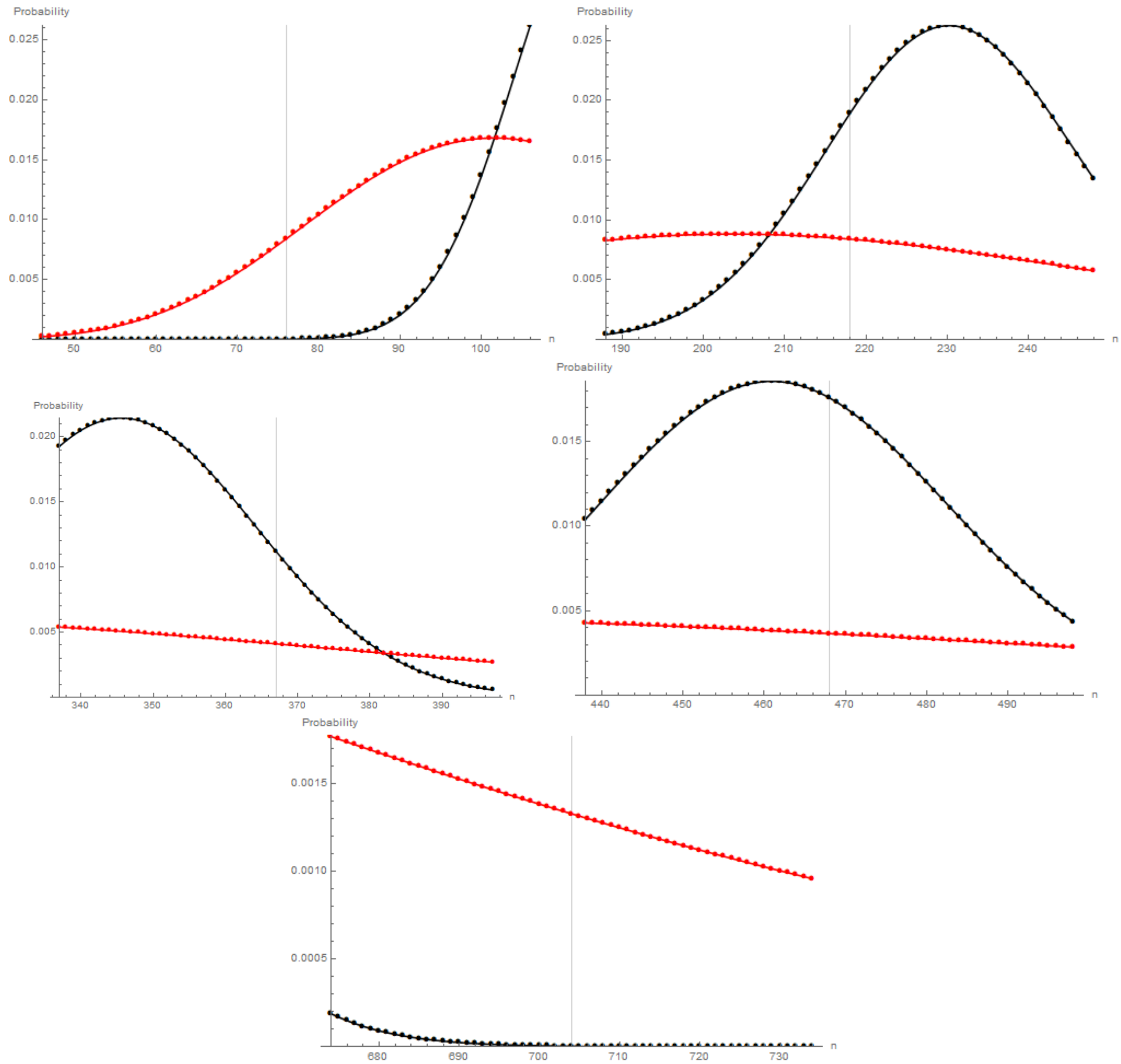


Figure 8.16: Barry County - Unemployment claims, fits at various t

Chapter 9

Conclusion

This thesis aimed to investigate distributions and methods by which non-equidispersed data could be modelled. As a result, weighted Poisson distributions and processes were investigated in-depth.

After a discussion into currently available methods of modelling non-equidispersed data, a power study was performed on a range of classical and more recent tests for Poissonity. This was followed by the introduction of 26 novel weight functions that can be used to augment the probability mass function of the Poisson distribution. The resulting distribution is referred to as a weighted Poisson distribution. Closed form statistical properties were derived for all the proposed weighted Poisson distributions. The newly derived distributions were then fitted to a variety of datasets, where, based on the AIC, AICc and BIC, it was found that the weighted Poisson distribution gave better fits than the Poisson in every case considered. The weighted Poisson distributions were also compared to a series of alternate methods that have been used in the past to model discrete data. In this case the weighted Poisson models outperformed the previous distributions in 13 of 17 cases considered.

Following an investigation into tests for homogeneous Poisson processes, the new weighted Poisson distributions were expanded into stochastic processes, referred to as weighted Poisson processes. Once again, closed form statistical properties of the processes were derived. These weighted Poisson processes were fitted to observed datasets and shown to have favourable fit characteristics relative to the traditional Poisson process.

There is, however, a considerable amount of followup research that can be conducted into the weighted Poisson distributions which would lead to the distribution being more useful in practice:

- The zero-inflated Poisson distribution has received relatively little attention in this thesis, and the conditional weighted Poisson distribution none. It is theoretically possible, however, to model zero-inflated data with conditional weighted Poisson distributions. This concept should be further investigated.

- The potential relationships that may exist between the weighted Poisson distribution and others (like the power distributions and the Lagrangian family of distributions) remains to be studied.

- An exploration into niche "compound" weight functions remains to be conducted. Certain specific practical situations might benefit from combining two or more weights. As an example, imagine a situation where there is an excess amount of zero observations, but where the zero-inflated Poisson distribution does not provide a good fit. It is also possible that a truncating distribution could be needed, but one of the truncating weights mentioned in this thesis might not provide a sufficiently good fit. In this situation it is worth exploring whether a truncating function, in combination with another weight, might be suitable.
- This thesis has solely investigated univariate weighted Poisson distributions. Due to the inherent flexibility of these distributions, expanding them into multivariate forms could have useful practical applications.
- In order for the weighted Poisson distribution to be more widely used in practice, a regression methodology needs to be developed.
- The parameters in this research were estimated with a numerical maximum likelihood method. The development of a formal EM algorithm is yet to be explored.
- An investigation into the weighted Poisson distributions use in a Bayesian framework remains to be done. Specifically, research into suitable priors for the parameters in the weights is worthy of consideration.

Chapter 10

Appendix

10.1 Definitions

Definition 10.1. *Natural numbers*
For notational purposes throughout this thesis $\mathbb{N}_0$ *will represent the set of natural numbers starting at 0. Similarly* $\mathbb{N}_1$ *will represent the set of natural numbers starting at 1.*
$\mathbb{N}_0 = \{0, 1, 2, ...\}$
$\mathbb{N}_1 = \{1, 2, 3, ...\}$.

Definition 10.2. *Empirical probability mass function*
Let $X_1, ..., X_n$ *be* n*, i.i.d. discrete random variables. The empirical probability mass function is then defined as*

$$\hat{f}\left(X_1, ..., X_n\right)_i \quad = \quad \frac{\textit{Number of realizations equal to } i}{n} = \frac{1}{n}\sum_{j=1}^{n} \quad I\left(X_j = i\right), i = 0, 1, 2, ...$$

Definition 10.3. *Poisson probability mass function*
A discrete random variable, X*, is said to follow the Poisson distribution with parameter* $\lambda > 0$ *if it has the following probability mass function*

$$f\left(x; \lambda\right) \quad = \quad \frac{\lambda^x e^{-\lambda}}{x!} \quad , x = 0, 1, 2, ...$$

This is denoted by $X \sim Poi\left(\lambda\right)$.

Definition 10.4. *Generalised Poisson probability mass function*
A discrete random variable, X*, is said to follow the generalised Poisson distribution with parameters* $\lambda_1 > 0$ *and* $|\lambda_2| < 1$ *if it has the following probability mass function*

$$f\left(x; \lambda_1, \lambda_2\right) \quad \frac{\lambda_1(\lambda_1 + x\lambda_2)^{x-1} e^{-(\lambda_1 + x\lambda_2)}}{x!} \quad , \quad x = 0, 1, 2, ...$$

This is denoted by $X \sim GenPoi\left(\lambda_1, \lambda_2\right)$.

Definition 10.5. *Binomial probability mass function*
A discrete random variable, X, is said to follow the binomial distribution with parameters $n \in 0, 1, 2, ...$ and $0 < p < 1$ if it has the following probability mass function

$$f(x; n, p) = \begin{pmatrix} n \\ x \end{pmatrix} p^x (1-p)^{n-x} \quad , x = 0, 1, 2, ...$$

This is denoted by $X \sim bin(n, p)$.

Definition 10.6. *Negative binomial probability mass function*
A discrete random variable, X, is said to follow the negative binomial distribution with parameters $r > 0$ and $0 < p < 1$ if it has the following probability mass function

$$f(x; r, p) = \begin{pmatrix} x + r - 1 \\ x \end{pmatrix} p^x (1-p)^r \quad , x = 0, 1, 2, ...$$

This is denoted by $X \sim negbin(r, p)$. It should be noted that many different forms of the negative binomial probability mass function exist in the literature. These differences occur when the random variable X is counting different things.

Definition 10.7. *Exponential probability density function*
A continuous random variable, X, is said to follow the exponential distribution with parameter $\lambda > 0$ if it has the following probability density function

$$f(x; \lambda) = \lambda e^{-\lambda x} \quad , x \geq 0$$

This is denoted by $X \sim exp(\lambda)$.

Definition 10.8. *Erlang probability density function*
A continuous random variable, X, is said to follow the Erlang distribution with parameters $\lambda > 0$ and $k \in 1, 2, ...$ if it has the following probability density function

$$f(x; \lambda, k) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!} \quad , x \geq 0$$

This is denoted by $X \sim Erl((x; \lambda, k))$.

Definition 10.9. *Gamma function*
The gamma function is defined as

$$\Gamma\left(\alpha\right)=\int_{0}^{\infty}t^{\alpha-1}e^{-t}dt.$$

Definition 10.10. *Upper incomplete gamma function*
The upper incomplete gamma function is defined as

$$\Gamma\left(\alpha,\beta\right)=\int_{\beta}^{\infty}t^{\alpha-1}e^{-t}dt.$$

Definition 10.11. *Lower incomplete gamma function*
The lower incomplete gamma function is defined as

$$\gamma\left(\alpha,\beta\right)=\int_{0}^{\beta}t^{\alpha-1}e^{-t}dt.$$

Definition 10.12. *Beta function*
The beta function is defined as

$$Beta\left(\alpha,\beta\right)=\frac{\Gamma\left(\alpha\right)\Gamma\left(\beta\right)}{\Gamma\left(\alpha+\beta\right)}$$

where $\Gamma\left(\right)$ *is the gamma function (Definition 10.9)*

Definition 10.13. *Locconvexity and concavity*
A function $g\left(\right)$ *is said to be logconvex if* $\forall x_1,x_2$ *in the domain of* $g\left(\right)$ *and* $\forall t\in[0,1]$

$$\begin{aligned} & log\left(g\left(tx_1+\left(1-t\right)x_2\right)\right) && \leq && tlog\left(g\left(x_1\right)\right)+\left(1-t\right)log\left(g\left(x_2\right)\right) \\ = & \; g\left(tx_1+\left(1-t\right)x_2\right) && \leq && g\left(x_1\right)^t g\left(x_2\right)^{1-t}. \end{aligned}$$

If $g\left(\right)$ *is twice differentiable then an equivalent condition exists, that states that* $g\left(\right)$ *is log-convex if* $\forall x$ *in the domain of* $g\left(\right)$

$$g^{''}\left(x\right)g\left(x\right)\geq\left(g^{'}\left(x\right)\right)^2.$$

Similarly a function $h\left(\right)$ *is said to be logconacve if* $\forall x_1,x_2$ *in the domain of* $h\left(\right)$ *and* $\forall t\in[0,1]$

$$\begin{aligned} & log\left(h\left(tx_1+\left(1-t\right)x_2\right)\right) && \geq \quad tlog\left(h\left(x_1\right)\right)+\left(1-t\right)log\left(h\left(x_2\right)\right) \\ = \quad & h\left(tx_1+\left(1-t\right)x_2\right) && \geq \quad h\left(x_1\right)^t h\left(x_2\right)^{1-t}. \end{aligned}$$

If $g()$ is twice differentiable then an equivalent condition exists that states that $g()$ is logconcave if for all x in the domain of $g()$

$$h^{''}\left(x\right)h\left(x\right)\leq\left(h^{'}\left(x\right)\right)^2.$$

Note that strict logconvexity and logconcavity is achieved when the same inequalities hold as above but with strictly larger than or smaller than symbols.

Definition 10.14. *Indicator function*
If A is a set, then the indicator function of A, denoted by I_A is defined as

$$I_A\left(x\right)=\begin{cases}1 & \text{if } x\in A\\ 0 & \text{if } x\notin A.\end{cases}$$

Bain and Engelhardt [4] p341.

Definition 10.15. *Confluent hypergeometric function*
The confluent hypergeometric function, sometime also referred to as the degenerate or Kummer hypergeometric functions, is defined as

$${}_1F_1\left(\alpha;\gamma;z\right)=\sum_{n=0}^{\infty}\frac{\left(\alpha\right)_n}{\left(\gamma\right)_n}\frac{z^n}{n!}$$

where $()_n$ is the Pochhammer symbol defined in Definition 10.18. Gradshteyn and Ryzhik [59] p1023.

Definition 10.16. *Gauss hypergeometric function*
The Gauss hypergeometric function is defined as

$${}_2F_1\left(\alpha,\beta;\gamma;z\right)=\sum_{n=0}^{\infty}\frac{\left(\alpha\right)_n\left(\beta\right)_n}{\left(\gamma\right)_n}\frac{z^n}{n!}$$

where $()_n$ is the Pochhammer symbol defined in Definition 10.18.
Gradshteyn and Ryzhik [59] p1005.

Definition 10.17. *Generalised hypergeometric function*
The generalised hypergeometric function is defined as

$$ {}_pF_q\left(\alpha_1,...,\alpha_p;\beta_1,...,\beta_q;z\right)=\sum_{n=0}^{\infty}\frac{\left(\alpha_1\right)_n...\left(\alpha_p\right)_n}{\left(\beta_1\right)_n...\left(\beta_q\right)_n}\frac{z^n}{n!} $$

where $()_n$ *is the Pochhammer symbol defined in Definition 10.18. Note that* ${}_pF_q\left(\alpha_1,...,\alpha_p;\beta_1,...,\beta_q;z\right)$ *converges absolutely* $\forall z$ *if* $p<q+1$. *If* $p=q+1$ *convergence occurs if* $|z|<1$, *and diverges* $\forall z\neq 0$ *if* $p>q+1$. *While it is possible to use analytic continuation to define that function at* $z=0$ *this extension will not be required in this thesis.*
Gradshteyn and Ryzhik [59] p1010.

Definition 10.18. *Pochhammer symbol*
$(a)_n$ *is known as the Pochhammer symbol. It is defined as*

$$ (a)_n=a\left(a+1\right)...\left(a+n-a\right)=\frac{\Gamma\left(a+n\right)}{\Gamma\left(a\right)} $$

where $(a)_0=1$.
Gradshteyn and Ryzhik [59] pxliii.

Definition 10.19. *Laguerre polynomials*
$L_n\left(x\right)$ *is known as the Laguerre polynomials. They are the solutions to the Laguerre equation*

$$ xy^{''}+\left(1-x\right)y^{'}+ny=0. $$

$L_n\left(x\right)$ *has the following expression:*

$$ L_n\left(x\right)=\sum_{k=0}^{n}\left(\begin{array}{c} n \\ k \end{array}\right)\frac{\left(-1\right)^k}{k!}x^k. $$

$L_n^{\alpha}\left(x\right)$ *is known as the generalised Laguerre polynomials.*
$L_n^{\alpha}\left(x\right)$ *has the following expression:*

$$ L_n^{\alpha}\left(x\right)=\sum_{k=0}^{n}\left(\begin{array}{c} n+\alpha \\ n-k \end{array}\right)\frac{\left(-1\right)^k}{k!}x^k. $$

Note that $L_n^0\left(x\right)=L_n\left(x\right)$.
Gradshteyn and Ryzhik [59] p1000.

Definition 10.20. *Modified Bessel function (1st kind)*
$I_n(x)$ is known as the modified Bessel function of the first kind. It satisfies the following differential equation

$$x^2y'' + xy' - \left(x^2 - n^2\right)y = 0.$$

$I_n(x)$ has the following expression:

$$I_n(x) = \sum_{i=0}^{\infty} \frac{1}{i!\Gamma(i+n+1)} \left(\frac{x}{2}\right)^{2i+n}.$$

Definition 10.21. *Hurwitz zeta function*
The Hurwitz zeta function is defined as

$$\zeta(a,b) = \sum_{n=0}^{\infty} \frac{1}{(b+n)^a}.$$

Definition 10.22. *Exponential Integral function*
$E_n(z)$ is known as the exponential integral function. It is defined as

$$E_n(z) = \int_1^{\infty} \frac{e^{-zt}}{t^n} dt.$$

Definition 10.23. *Stochastic process*
A stochastic process $\{N(t), t \in T\}$ is a collection of random variable. That is, for each t in the set T, $N(t)$ is a random variable. If T is a countable set the stochastic process is said to be "discrete-time", and if T is a continuum it is said to be a "continuous-time" process. If $N(t)$ can only assume integer values the stochastic process is said to be an "integer-valued" stochastic process, whereas if $N(t)$ can assume any real value the process is called a "real-valued" stochastic process.

Definition 10.24. *Counting process*
A stochastic process (see Definition 10.23) $\{N(t), t \geq 0\}$ is said to be a counting process if $N(t)$ represents the total number of events that have occurred up to time t. As such for a stochastic process to qualify as a counting process it must satisfy four fundamental properties:

1. $N(t) > 0$.

2. *$N(t)$ is integer valued.*

3. *If $s < t$, then $N(s) \leq N(t)$.*

4. *For $s < t$, $N(t) - N(s)$ is the number of events that occurred in the interval $(s, t]$.*

Definition 10.25. *Poisson process - version 1*
A counting process (See Definition 10.24) $\{N(t), t \geq 0\}$ is said to be a Poisson process having rate $\lambda > 0$ if:

1. $N(0) = 0$.

2. *The process has independent increments.*

3. *The number of events in any interval of length t is Poisson distributed with mean λt. In other words, for $0 < s < t$ it follows that*

$$P(N(t+s) - N(s) = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}, n = 0, 1, 2, ...$$

Definition 10.26. *Poisson process - version 2*
A counting process (See Definition 10.24) $\{N(t), t \geq 0\}$ is said to be a Poisson process having rate $\lambda > 0$ if:

1. $N(0) = 0$.

2. *The process has stationary and independent increments.*

3. $P(N(h) = 1) = \lambda h + o(h)$.

4. $P(N(h) \geq 2) = o(h)$.

Definition 10.27. *Poisson process - version 3*
A counting process (See Definition 10.24) $\{N(t), t \geq 0\}$ is said to be a Poisson process having rate $\lambda > 0$ if:

$$N(t) = \sum_{n\geq 1} I_{(0,t]}(X_n), t \geq 0,$$

where the sequence (X_n) has i.i.d. increments $T_1, T_2, ...$ each of which follows an $exp(\lambda)$ distribution. The X_n are called arrival epochs and the T_n inter-arrival times.

Definition 10.28. *Compound Poisson process*

A stochastic process $\{X(t), t \geq 0\}$ is said to be a compound Poisson process if it can be represented, for $t \geq 0$, by

$$X(t) = \sum_{i=1}^{N(t)} X_i$$

where $\{N(t), t \geq 0\}$ is a Poisson process and $X_1, X_2, ...$ is a family of independent identically distributed random variables that is independent of the process $\{N(t), t \geq 0\}$.

Definition 10.29. *Sample mean*
Let $X_1, ..., X_n$ be a sample of size n, The sample mean is then defined to be

$$\bar{X} = \tfrac{1}{n} \textstyle\sum_{j=0}^{\infty} X_j.$$

Definition 10.30. *Sample variance*
Let $X_1, ..., X_n$ be a sample of size n, The sample variance is then defined to be

$$S^2 = \tfrac{1}{n-1} \textstyle\sum_{j=0}^{\infty} \left(X_j - \bar{X}\right).$$

Definition 10.31. *Waiting and inter-arrival times*

Consider a counting process $N(t)$, and let X_1 denote the time of the first event. Further, for $n \geq 1$, let X_n denote the time between the $(n-1)^{st}$ and the n^{th} event. The sequence $\{X_n, n \geq 1\}$ is called the sequence of inter-arrival times. If S_n is defined as the sum of inter-arrival times up to an including the n^{th} event (i.e. $S_n = \sum_{i=1}^{n} X_i, n \geq 1$), then S_n is said to be the waiting time until the n^{th} event occurs.

10.2 Results and proofs

Theorem 10.1. *The Poisson properties*
Suppose that X is a Poisson random variable with with parameter $\lambda > 0$ (Definition 10.3). Then $E(X) = \lambda$ and $Var(X) = \lambda$.

See Bain and Engelhardt [4] p104.

Theorem 10.2. *The binomial properties*
Suppose that X is a binomial random variable with with parameters $n = 0, 1, 2, ...$ and $0 < p < 1$ (Definition 10.5). Then $E(X) = np$ and $Var(X) = np(1-p)$.

See Bain and Engelhardt [4] p95.

Theorem 10.3. *The negative binomial properties*
Suppose that X is a negative binomial random variable with with parameters $r = 0, 1, 2, ...$ and $0 < p| < 1$ (Definition 10.6). Then $E(X) = \frac{rp}{1-p}$ and $Var(X) = \frac{rp}{(1-p)^2}$.

See Bain and Engelhardt [4] p102.

Theorem 10.4. *The generalised Poisson*
Suppose that X is a generalised Poisson random variable with with parameters $\lambda_1 > 0$ and $|\lambda_2| < 1$ (Definition 10.4). Then $E(X) = \frac{\lambda_1}{1-\lambda_2}$ and $Var(X) = \frac{\lambda_1}{(1-\lambda_2)^3}$.

See Consul and Jain [29].

Theorem 10.5. *Derivative of the confluent hypergeometric function*
Suppose that ${}_1F_1(\alpha;\gamma;z)$ confluent hypergeometric function (Definition 10.15). Then its first derivative with respect to z is given by

$$\frac{\partial}{\partial z}{}_1F_1(\alpha;\gamma;z) = \frac{\alpha}{\gamma}{}_1F_1(\alpha+1;\gamma+1;z).$$

See Erdélyi [46].

Theorem 10.6. *Derivative of the generalised hypergeometric function*
Suppose that ${}_pF_q(\alpha_1,...,\alpha_p;\beta_1,...,\beta_q;z)$ is the hypergeometric function (Definition 10.17). Then its first derivative with respect to z is given by

$$\frac{\partial}{\partial z}{}_pF_q(\alpha_1,...,\alpha_p;\beta_1,...,\beta_q;z) = \frac{\alpha_1\alpha_2...\alpha_p}{\beta_1\beta_2...\beta_q}{}_pF_q(\alpha_1+1,...,\alpha_p+1;\beta_1+1,...,\beta_q+1;z).$$

See Slater [130].

Theorem 10.7. *Generalised hypergeometric function derivative identity*
Suppose that ${}_pF_q(\alpha_1,...,\alpha_p;\beta_1,...,\beta_q;z)$ *is the hypergeometric function (Definition 10.17) and that* $\alpha_p+1=\beta_q$*. Then its first derivative of the hypergeometric function with respect to* z *is given by*

$$\begin{aligned}\frac{\partial}{\partial z}{}_pF_q(\alpha_1,...,\alpha_p;\beta_1,...,\beta_q;z) &= \frac{\alpha_p}{z}{}_{p-1}F_{q-1}(\alpha_1,...,\alpha_{p-1};\beta_1,...,\beta_{q-1};z)\\ &- \frac{\alpha_p}{z}{}_pF_q(\alpha_1,...,\alpha_p;\beta_1,...,\beta_q;z).\end{aligned}$$

Proof. From Theorem 10.6 we know that □

$$\begin{aligned}\frac{\partial}{\partial z}{}_pF_q(\alpha_1,...,\alpha_p;\beta_1,...,\beta_q;z) &= \frac{\alpha_1\alpha_2...\alpha_p}{\beta_1\beta_2...\beta_q}{}_pF_q(\alpha_1+1,...,\alpha_p+1;\beta_1+1,...,\beta_q+1;z)\\ &= {}_pF_q(\alpha_1,...,\alpha_p;\beta_1,...,\beta_q;z)\\ &+ z\frac{\alpha_1\alpha_2...\alpha_{p-1}}{\beta_1\beta_2...\beta_q}{}_pF_q(\alpha_1+1,...,\alpha_p+1;\beta_1+1,...,\beta_q+1;z)\\ &= \left(1+\frac{z}{\alpha_p}\frac{d}{dz}\right){}_pF_q(\alpha_1,...,\alpha_p;\beta_1,...,\beta_q;z).\end{aligned}$$

By using Euler's integral transform (Slater [130]) it follows that

$$\begin{aligned}{}_pF_q(\alpha_1,...,\alpha_p;\beta_1,...,\beta_q;z) &= \frac{\Gamma(\beta_q)}{\Gamma(\alpha_p)\Gamma(\beta_q-\beta_p)}\left(\int_0^1 t^{\alpha_p-1}(1-t)^{\beta_q-\alpha_p-1}\right.\\ &\times\ {}_{p-1}F_{q-1}(\alpha_1,...,\alpha_{p-1};\beta_1,...,\beta_{q-1};t)\,dt).\end{aligned}$$

By using the assumption that $\alpha_p+1=\beta_q$, and by reparameterising $zt=u$ it follows that

$$\begin{aligned}{}_pF_q(\alpha_1,...,\alpha_p;\beta_1,...,\beta_q;z) &= \alpha_p\int_0^1 t^{\alpha_p-1}{}_{p-1}F_{q-1}(\alpha_1,...,\alpha_{p-1};\beta_1,...,\beta_{q-1};zt)\,dt.\\ &= \frac{\alpha_p}{z^{\alpha_p}}\int_0^z u^{\alpha_p-1}{}_{p-1}F_{q-1}(\alpha_1,...,\alpha_{p-1};\beta_1,...,\beta_{q-1};u)\,du.\end{aligned}$$

Consequently it holds that

$$\begin{aligned}\frac{\partial}{\partial z}{}_pF_q(\alpha_1,...,\alpha_p;\beta_1,...,\beta_q;z) &= \frac{\alpha_p}{z}{}_{p-1}F_{q-1}(\alpha_1,...,\alpha_{p-1};\beta_1,...,\beta_{q-1};z)\\ &- \frac{\alpha_p}{z}{}_pF_q(\alpha_1,...,\alpha_p;\beta_1,...,\beta_q;z).\end{aligned}$$

Theorem 10.8. *Confluent hypergeometric relations*
Suppose that ${}_1F_1(\alpha;\gamma;z)$ *is the confluent hypergeometric function (Definition 10.15). Then the following relationships hold:*

$$\begin{aligned}{}_1F_1(\alpha;\gamma;z) &= e^z{}_1F_1(\gamma-\alpha;\gamma;-z).\\ \frac{z}{\gamma}{}_1F_1(\alpha+1;\gamma+1;z) &= {}_1F_1(\alpha+1;\gamma;z)-{}_1F_1(\alpha;\gamma;z).\\ \alpha{}_1F_1(\alpha+1;\gamma+1;z) &= (\alpha-\gamma){}_1F_1(\alpha;\gamma+1;z)+\gamma{}_1F_1(\alpha;\gamma;z).\\ \alpha{}_1F_1(\alpha+1;\gamma;z) &= (z+2\alpha-\gamma){}_1F_1(\alpha;\gamma;z)+(\gamma-\alpha){}_1F_1(\alpha-1;\gamma;z).\end{aligned}$$

See Gradshteyn and Ryzhik [59] p1023.

Theorem 10.9. *Gauss hypergeometric relations*
Suppose that ${}_2F_1(\alpha, \beta; \gamma; z)$ *is the Gauss hypergeometric function (Definition 10.16). Then the following relationships hold:*

$$\begin{aligned} {}_2F_1(-n, \beta; \beta; -z) &= (1+z)^n. \\ {}_2F_1(1,1;2;-z) &= \frac{ln(1+z)}{z}. \end{aligned}$$

See Gradshteyn and Ryzhik [59] p1023.

Theorem 10.10. *Gamma function relationships*
Suppose that $\Gamma(\alpha), \Gamma(\alpha, \beta)$ *and* $\gamma(\alpha, \beta)$ *are gamma, upper incomplete gamma and lower incomplete gamma functions as defined in Definitions 10.9, 10.10 and 10.11 respectively. Then the following relationships hold:*

$$\Gamma(\alpha) = \gamma(\alpha, \beta) + \Gamma(\alpha, \beta).$$

$$\begin{aligned} \Gamma(\alpha, \beta) &= (\alpha-1)! e^{-\beta} \sum_{i=0}^{\alpha-1} \frac{\beta^i}{i!}. \\ \gamma(\alpha, \beta) &= (\alpha-1)! \left(1 - e^{-\beta} \sum_{i=0}^{\alpha-1} \frac{\beta^i}{i!}\right). \\ \Gamma(\alpha, \beta) &= \Gamma(\alpha) - \sum_{i=0}^{\infty} \frac{(-1)^i \beta^{a+i}}{i!(a+i)}. \\ \gamma(\alpha, \beta) &= \sum_{i=0}^{\infty} \frac{(-1)^i \beta^{a+i}}{i!(a+i)}. \\ \gamma(\alpha, \beta) &= \beta^a \sum_{i=0}^{\infty} \frac{(-\beta)^i}{i!(a+i)}. \end{aligned}$$

$$\begin{aligned} \Gamma(\alpha, \beta) - \Gamma(\alpha, \beta+\delta) &= \gamma(\alpha, \beta+\delta) - \gamma(\alpha, \beta). \\ \Gamma(\alpha)\Gamma(\alpha+\delta, \beta) - \Gamma(\alpha+\delta)\Gamma(\alpha, \beta) &= \Gamma(\alpha+\delta)\gamma(\alpha, \beta) - \Gamma(\alpha)\gamma(\alpha+\delta, \beta). \end{aligned}$$

$$\begin{aligned} \gamma(\alpha+1, \beta) &= \alpha\gamma(\alpha, \beta) - \beta^\alpha e^{-\beta}. \\ \Gamma(\alpha+1, \beta) &= \alpha\Gamma(\alpha, \beta) + \beta^\alpha e^{-\beta}. \end{aligned}$$

$$\begin{aligned} \frac{d\gamma(\alpha, \beta)}{d\beta} &= \beta^{\alpha-1} e^{-\beta}. \\ \frac{d\Gamma(\alpha, \beta)}{d\beta} &= -\beta^{\alpha-1} e^{-\beta}. \end{aligned}$$

See Gradshteyn and Ryzhik [59] p899-901.

Theorem 10.11. *Laguerre polynomial derivatives*
If $L_n(x)$ *are the Laguerre polynomials and* $L_n^\alpha(x)$ *are the generalised Laguerre polynomials as defined in Definition 10.19 then*

$$\frac{d}{dx} L_n^\alpha(x) = -L_{n-1}^{\alpha+1}(x)$$

See Gradshteyn and Ryzhik [59] p 1001.

Theorem 10.12. *Hurwitz zeta relations*
Suppose that $\zeta(a,b)$ is the Hurwitz zeta function (Definition 10.21). Then the following relationships hold:

$$\zeta(a,b) = \zeta(a,b+1) + b^{-a}.$$

$$\frac{\partial \zeta(a,b)}{\partial b} = -a\zeta(a+1,b).$$

See Coffey [22].

Theorem 10.13. *Properties of weighted Poisson processes*
Suppose that $N^w(t)$ is a weighted Poisson process as given in Definition 10.25 and N^w is a weighted Poisson random variable as given in Definition 2.2. Using the properties of compound distributions it then follows that:
The expected value of the process is given by

$$\begin{aligned} E(N^w(t)) &= E(N^w) E\left(I_{[0,t]}[U_j]\right) \\ &= E(N^w) t. \end{aligned}$$

The variance of the process is given by

$$\begin{aligned} Var(N^w(t)) &= E(N^w) Var\left(I_{[0,t]}[U_j]\right) + Var(N^w) E\left(I_{[0,t]}[U_j]\right)^2 \\ &= E(N^w) t(1-t) + Var(N^w) t^2. \end{aligned}$$

The probability generating function of the process is given by

$$\begin{aligned} G_{N^w(t)}(z) &= G_{N^w}\left(G_{I_{[0,t]}[U_j]}(z)\right) \\ &= G_{N^w}(1-t+tz). \end{aligned}$$

The probability generating function of an increment of the process is given by

$$\begin{aligned} G_{N^w(t)-N^w(s)}(z) &= G_{N^w}\left(G_{I_{[s,t]}[U_j]}(z)\right) \\ &= G_{N^w}(1-(t-s)+(t-s)z). \end{aligned}$$

The joint probability mass function of the process is given by

$$P(N^w(s)=k, N^w(t)=n) = \binom{n}{k}\left(\frac{s}{t}\right)^k\left(1-\frac{s}{t}\right)^{n-k} P(N^w=n).$$

The covariance of the process is given by

$$\begin{aligned} Cov\left(N^{w}\left(s\right),N^{w}\left(t\right)\right) &= stVar\left(N^{w}\right)+s\left(1-t\right)E\left(N^{w}\right) \\ &= s\left(tVar\left(N^{w}\right)+\left(1-t\right)E\left(N^{w}\right)\right). \end{aligned}$$

The covariance between two independent increments of the process is given by

$$\begin{aligned} Cov\left(N^{w}\left(s\right),N^{w}\left(t\right)-N^{w}\left(s\right)\right) &= Cov\left(N^{w}\left(s\right),N^{w}\left(t\right)\right)-Var\left(N^{w}\left(s\right)\right) \\ &= s\left(t-s\right)\left(Var\left(N^{w}\right)-E\left(N^{w}\right)\right). \end{aligned}$$

See Balakrishnan and Kozubowski [6].

Theorem 10.14. *If the weight function in Equation 2.7 is $w\left(n\right)=n$ then*

$$\begin{aligned} E\left(w\left(N;\phi\right)\right) &= \lambda. \\ f_{w}\left(n\right) &= \tfrac{e^{-\lambda}\lambda^{n-1}}{(n-1)!}. \\ g\left(z\right) &= e^{\lambda(z-1)}z. \\ E\left(N^{w}\right) &= \lambda+1. \\ Var\left(N^{w}\right) &= \lambda. \end{aligned}$$

Proof. From the definition of the normalising constant, it follows that

$$\begin{aligned} E\left(w\left(N\right)\right) &= \textstyle\sum_{k=0}^{\infty}w\left(k\right)f\left(k\right) \\ &= \textstyle\sum_{k=0}^{\infty}k\frac{e^{-\lambda}\lambda^{k}}{k!}. \end{aligned}$$

Since this is the first moment of a Poisson distribution with parameter λ it follows that $E\left(w\left(N\right)\right)=\lambda$.

From the definition of the weighted Poisson probability mass function, it follows that

$$\begin{aligned} f_{w}\left(n\right) &= \tfrac{w(n)f(n)}{\sum_{k=0}^{\infty}w(k)f(k)} \\ &= \tfrac{n\frac{e^{-\lambda}\lambda^{n}}{n!}}{\lambda} \\ &= \tfrac{ne^{-\lambda}\lambda^{n-1}}{n!} \\ &= \tfrac{e^{-\lambda}\lambda^{n-1}}{(n-1)!}. \end{aligned}$$

From the definition of the probability generating function, it follows that

$$\begin{aligned} g(z) &= \sum_{k=0}^{\infty} f_w(k) z^k \\ &= \sum_{k=0}^{\infty} \frac{ke^{-\lambda}\lambda^{k-1}}{k!} z^k \\ &= \frac{ze^{-\lambda}}{e^{-\lambda z}} \sum_{k=1}^{\infty} \frac{e^{-\lambda z}(\lambda z)^{k-1}}{(k-1)!}. \end{aligned}$$

Since this is the probability mass function of a Poisson distribution with parameter λz it follows that $g(z) = e^{(z-1)\lambda} z$.

From the definition of the expected value it follows that

$$E(N^w) = \lim_{z \to -1} \frac{\partial}{\partial z} g(z).$$

However,

$$\begin{aligned} \frac{\partial}{\partial z} g(z) &= \frac{\partial}{\partial z} \left(e^{(z-1)\lambda} z \right) \\ &= \left(\frac{\partial}{\partial z} e^{\lambda(z-1)} \right) (z) + \left(e^{\lambda(z-1)} \right) \left(\frac{\partial}{\partial z} z \right) \\ &= e^{\lambda(z-1)} \lambda z + e^{\lambda(z-1)} \\ &= e^{\lambda(z-1)} (\lambda z + 1). \end{aligned}$$

Consequently, it follows that

$$\begin{aligned} E(N^w) &= \lim_{z \to -1} e^{\lambda(z-1)} (\lambda z + 1) \\ &= \lambda + 1. \end{aligned}$$

From the definition of the variance it follows that

$$Var(N^w) = \lim_{z \to -1} \frac{\partial^2}{\partial z^2} g(z) + E(N^w) - (E(N^w))^2.$$

However,

$$\begin{aligned} \frac{\partial^2}{\partial z^2} g(z) &= \frac{\partial}{\partial z} \left(e^{\lambda(z-1)} (\lambda z + 1) \right) \\ &= \left(\frac{\partial}{\partial z} e^{\lambda(z-1)} \right) (\lambda z + 1) + \left(e^{\lambda(z-1)} \right) \left(\frac{\partial}{\partial z} \lambda z + 1 \right) \\ &= e^{\lambda(z-1)} \lambda (\lambda z + 1) + e^{\lambda(z-1)} \lambda \\ &= e^{\lambda(z-1)} (\lambda (\lambda z + 1) + \lambda). \end{aligned}$$

Consequently, it follows that

$$\begin{aligned} Var\left(N^{w}\right) &= \lim_{z\to-1}\frac{\partial^{2}}{\partial z^{2}}g\left(z\right)+E\left(N^{w}\right)-\left(E\left(N^{w}\right)\right)^{2} \\ &= \lim_{z\to-1}e^{\lambda(z-1)}\left(\lambda\left(\lambda z+1\right)+\lambda\right)+\left(1+\lambda\right)-\left(1+\lambda\right)^{2} \\ &= \left(\lambda\left(\lambda+1\right)+\lambda\right)+\left(1+\lambda\right)-\left(1+\lambda\right)^{2} \\ &= \lambda^{2}+3\lambda+1-\left(\lambda^{2}+2\lambda+1\right) \\ &= \lambda. \end{aligned}$$

□

Theorem 10.15. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w\left(n;\phi\right)=n+\varepsilon$ *then*

$$\begin{aligned} E\left(w\left(N;\phi\right)\right) &= \lambda+\varepsilon. \\ f_{w}\left(n\right) &= \frac{(n+\varepsilon)e^{-\lambda}\lambda^{n}}{(\lambda+\varepsilon)n!}. \\ g\left(z\right) &= \frac{e^{(z-1)\lambda}(\varepsilon+\lambda z)}{\lambda+\varepsilon}. \\ E\left(N^{w}\right) &= \frac{\lambda(1+\lambda+\varepsilon)}{\lambda+\varepsilon}. \\ Var\left(N^{w}\right) &= \frac{\lambda\left((\lambda+\varepsilon)^{2}+\varepsilon\right)}{(\lambda+\varepsilon)^{2}}. \end{aligned}$$

Proof. From the definition of the normalising constant, it follows that

$$\begin{aligned} E\left(w\left(N\right)\right) &= \sum_{k=0}^{\infty}w\left(k\right)f\left(k\right) \\ &= \sum_{k=0}^{\infty}\left(k+\varepsilon\right)\frac{e^{-\lambda}\lambda^{k}}{k!} \\ &= \sum_{k=0}^{\infty}k\frac{e^{-\lambda}\lambda^{k}}{k!}+\varepsilon\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^{k}}{k!}. \end{aligned}$$

Since $\sum_{k=0}^{\infty}k\frac{e^{-\lambda}\lambda^{k}}{k!}$ is the first moment of a Poisson distribution with parameter λ, and $\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^{k}}{k!}=1$, it follows that $E\left(w\left(N\right)\right)=\lambda+\varepsilon$.

From the definition of the weighted Poisson probability mass function, it follows that

$$\begin{aligned} f_{w}\left(n\right) &= \frac{w(n)f(n)}{\sum_{k=0}^{\infty}w(k)f(k)} \\ &= \frac{(n+\varepsilon)\frac{e^{-\lambda}\lambda^{n}}{n!}}{\lambda+\varepsilon} \\ &= \frac{(n+\varepsilon)e^{-\lambda}\lambda^{n}}{(\lambda+\varepsilon)n!}. \end{aligned}$$

From the definition of the probability generating function, it follows that

$$\begin{aligned} g(z) &= \sum_{k=0}^{\infty} f_w(k)\, z^k \\ &= \sum_{k=0}^{\infty} \frac{(k+\varepsilon)e^{-\lambda}\lambda^k}{(\lambda+\varepsilon)k!} z^k \\ &= \frac{e^{-\lambda}}{(\lambda+\varepsilon)} \sum_{k=0}^{\infty} \frac{(k+\varepsilon)(\lambda z)^k}{k!} \\ &= \frac{e^{-\lambda}}{(\lambda+\varepsilon)} \left(\sum_{k=0}^{\infty} \frac{k(\lambda z)^k}{k!} + \sum_{k=0}^{\infty} \frac{\varepsilon(\lambda z)^k}{k!} \right) \\ &= \frac{e^{-\lambda}}{(\lambda+\varepsilon)} \left(\sum_{k=1}^{\infty} \frac{k(\lambda z)^k}{k!} + \sum_{k=0}^{\infty} \frac{\varepsilon(\lambda z)^k}{k!} \right) \\ &= \frac{e^{-\lambda}}{(\lambda+\varepsilon)} \left(\lambda z e^{\lambda z} + \varepsilon e^{\lambda z} \right) \\ &= \frac{e^{(z-1)\lambda}(\varepsilon+\lambda z)}{\lambda+\varepsilon}. \end{aligned}$$

From the definition of the expected value, it follows that

$$E(N^w) = \lim_{z\to -1} \frac{\partial}{\partial z} g(z).$$

However,

$$\begin{aligned} \frac{\partial}{\partial z} g(z) &= \frac{\partial}{\partial z} \left(\frac{e^{(z-1)\lambda}(\varepsilon+\lambda z)}{\lambda+\varepsilon} \right) \\ &= \frac{e^{(z-1)\lambda}\lambda}{\lambda+\varepsilon} + \frac{e^{(z-1)\lambda}\lambda(\varepsilon+\lambda z)}{\lambda+\varepsilon} \\ &= \frac{e^{(z-1)\lambda}\lambda(\varepsilon+\lambda z+1)}{\lambda+\varepsilon}. \end{aligned}$$

Consequently, it follows that

$$\begin{aligned} E(N^w) &= \lim_{z\to -1} \frac{e^{(z-1)\lambda}\lambda(\varepsilon+\lambda z+1)}{\lambda+\varepsilon} \\ &= \frac{\lambda(1+\lambda+\varepsilon)}{\lambda+\varepsilon}. \end{aligned}$$

From the definition of the variance, it follows that

$$Var(N^w) = \lim_{z\to -1} \frac{\partial^2}{\partial z^2} g(z) + E(N^w) - (E(N^w))^2.$$

However,

$$\begin{aligned}
\frac{\partial^2}{\partial z^2} g(z) &= \frac{\partial}{\partial z}\left(\frac{e^{(z-1)\lambda}\lambda(\varepsilon+\lambda z+1)}{\lambda+\varepsilon}\right) \\
&= \frac{2e^{(z-1)\lambda}\lambda^2}{\lambda+\varepsilon} + \frac{e^{(z-1)\lambda}\lambda^2(\varepsilon+\lambda z)}{\lambda+\varepsilon} \\
&= \frac{e^{(z-1)\lambda}\lambda^2(\varepsilon+\lambda z+2)}{\lambda+\varepsilon}.
\end{aligned}$$

Consequently, it follows that

$$\begin{aligned}
Var(N^w) &= \lim_{z\to -1}\frac{\partial^2}{\partial z^2} g(z) + E(N^w) - (E(N^w))^2 \\
&= \lim_{z\to -1}\frac{e^{(z-1)\lambda}\lambda^2(\varepsilon+\lambda z+2)}{\lambda+\varepsilon} + \frac{\lambda(1+\lambda+\varepsilon)}{\lambda+\varepsilon} - \left(\frac{\lambda(1+\lambda+\varepsilon)}{\lambda+\varepsilon}\right)^2 \\
&= \frac{\lambda^2(\varepsilon+\lambda+2)}{\lambda+\varepsilon} + \frac{\lambda(1+\lambda+\varepsilon)}{\lambda+\varepsilon} - \left(\frac{\lambda(1+\lambda+\varepsilon)}{\lambda+\varepsilon}\right)^2 \\
&= \frac{(\lambda+\varepsilon)\lambda^2(\varepsilon+\lambda+2)+(\lambda+\varepsilon)\lambda(1+\lambda+\varepsilon)-\lambda^2(1+\lambda+\varepsilon)^2}{(\lambda+\varepsilon)^2} \\
&= \frac{\lambda\left(\varepsilon+\varepsilon^2+2\varepsilon\lambda+\lambda^2\right)}{(\lambda+\varepsilon)^2} \\
&= \frac{\lambda\left((\lambda+\varepsilon)^2+\varepsilon\right)}{(\lambda+\varepsilon)^2}.
\end{aligned}$$

□

Theorem 10.16. *If the weight function used in the weighted Poisson probability mass function is chosen as $w(n;\phi) = an^3 + bn^2 + cn$ then*

$$\begin{aligned}
E(w(N;\phi)) &= \lambda(a+b+c) + \lambda^2(3a+b+\lambda a). \\
f_w(n) &= \frac{(an^3+bn^2+cn)}{(\lambda(3a+b+\lambda a)+(a+b+c))}\frac{e^{-\lambda}\lambda^n}{n!}. \\
g(z) &= \frac{e^{\lambda(z-1)}z\left(a\left((\lambda z)^2+3\lambda z+1\right)+b(\lambda z+1)+c\right)}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c}. \\
E(N^w) &= \frac{a\left(\lambda^3+6\lambda^2+7\lambda+1\right)+b\left(\lambda^2+3\lambda+1\right)+c(\lambda+1)}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c}. \\
Var(N^w) &= \frac{a\left(\lambda^4+9\lambda^3+19\lambda^2+8\lambda\right)+b\left(\lambda^3+5\lambda^2+4\lambda\right)+c\left(\lambda^2+2\lambda\right)}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c} + E(N^w) - (E(N^w))^2.
\end{aligned}$$

Proof. From the definition of the normalising constant, it follows that

$$
\begin{aligned}
E\left(w\left(N;\phi\right)\right) &= \textstyle\sum_{k=0}^{\infty} w\left(k\right) f\left(k\right) \\
&= \textstyle\sum_{k=0}^{\infty} \left(ak^3+bk^2+ck\right) \frac{e^{-\lambda}\lambda^k}{k!} \\
&= \textstyle\sum_{k=0}^{\infty} ak^3 \frac{e^{-\lambda}\lambda^k}{k!} + \sum_{k=0}^{\infty} bk^2 \frac{e^{-\lambda}\lambda^k}{k!} + \sum_{k=0}^{\infty} ck \frac{e^{-\lambda}\lambda^k}{k!} \\
&= a\textstyle\sum_{k=0}^{\infty} k^3 \frac{e^{-\lambda}\lambda^k}{k!} + b\sum_{k=0}^{\infty} k^2 \frac{e^{-\lambda}\lambda^k}{k!} + c\sum_{k=0}^{\infty} k \frac{e^{-\lambda}\lambda^k}{k!}.
\end{aligned}
$$

Since the three sums in the above equations are the third, second and first moments respectively of a Poisson distribution with parameter λ, it follows that

$$
\begin{aligned}
E\left(w\left(N;\phi\right)\right) &= a\left(\lambda^3+3\lambda^2+\lambda\right)+b\left(\lambda^2+\lambda\right)+c\lambda \\
&= \lambda\left(a+b+c\right)+\lambda^2\left(3a+b+\lambda a\right).
\end{aligned}
$$

From the definition of the weighted Poisson probability mass function, it follows that

$$
\begin{aligned}
f_w\left(n\right) &= \frac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\
&= \frac{\left(an^3+bn^2+cn\right)\frac{e^{-\lambda}\lambda^n}{n!}}{\lambda(a+b+c)+\lambda^2(3a+b+\lambda a)} \\
&= \frac{\left(an^3+bn^2+cn\right)}{\lambda(a+b+c)+\lambda^2(3a+b+\lambda a)} \frac{e^{-\lambda}\lambda^n}{n!}.
\end{aligned}
$$

From the definition of the probability generating function, it follows that

$$
\begin{aligned}
g\left(z\right) &= \textstyle\sum_{k=0}^{\infty} f_w\left(k\right) z^k \\
&= \textstyle\sum_{k=0}^{\infty} \frac{\left(ak^3+bk^2+ck\right)}{\lambda(a+b+c)+\lambda^2(3a+b+\lambda a)} \frac{e^{-\lambda}\lambda^k}{k!} z^k \\
&= \frac{1}{\lambda(a+b+c)+\lambda^2(3a+b+\lambda a)} \left(\textstyle\sum_{k=0}^{\infty} ak^3 \frac{\lambda^k e^{-\lambda}}{k!} z^k\right. \\
&+ \left.\textstyle\sum_{k=0}^{\infty} bk^2 \frac{\lambda^k e^{-\lambda}}{k!} z^k + \sum_{k=0}^{\infty} ck \frac{\lambda^k e^{-\lambda}}{k!} z^k\right) \\
&= \frac{e^{-\lambda}}{e^{-\lambda z}\left(\lambda(a+b+c)+\lambda^2(3a+b+\lambda a)\right)} \left(a\textstyle\sum_{k=0}^{\infty} k^3 \frac{(\lambda z)^k e^{-\lambda z}}{k!}\right. \\
&+ \left. b\textstyle\sum_{k=0}^{\infty} k^2 \frac{(\lambda z)^k e^{-\lambda z}}{k!} + c\sum_{k=0}^{\infty} k \frac{(\lambda z)^k e^{-\lambda z}}{k!}\right).
\end{aligned}
$$

Since the three sums in the above equations are the third, second and first moments respectively of a Poisson distribution with parameter λz, it follows that

$$
\begin{aligned}
g\left(z\right) &= \frac{e^{-\lambda}}{e^{-\lambda z}\left(\lambda(a+b+c)+\lambda^2(3a+b+\lambda a)\right)} \left(a\left(\left(\lambda z\right)^3+3\left(\lambda z\right)^2+\lambda z\right)+b\left(\left(\lambda z\right)^2+\lambda\right)+c\lambda z\right) \\
&= \frac{e^{-\lambda+\lambda z}\lambda z\left(a\left((\lambda z)^2+3\lambda z+1\right)+b(\lambda z+1)+c\right)}{\lambda(a+b+c)+\lambda^2(3a+b+\lambda a)} \\
&= \frac{e^{\lambda(z-1)}z\left(a\left((\lambda z)^2+3\lambda z+1\right)+b(\lambda z+1)+c\right)}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c}.
\end{aligned}
$$

From the definition of the expected value, it follows that

$$E\left(N^{w}\right) = \lim_{z\to -1}\frac{\partial}{\partial z}g\left(z\right).$$

However,

$$\begin{aligned}
\frac{\partial}{\partial z}g\left(z\right) &= \frac{\partial}{\partial z}\frac{e^{\lambda(z-1)}z\left(a\left((\lambda z)^{2}+3\lambda z+1\right)+b(\lambda z+1)+c\right)}{a(\lambda^{2}+3\lambda+1)+b(\lambda+1)+c}\\
&= \left(\frac{\partial}{\partial z}e^{\lambda(z-1)}\right)\frac{z\left(a\left((\lambda z)^{2}+3\lambda z+1\right)+b(\lambda z+1)+c\right)}{a(\lambda^{2}+3\lambda+1)+b(\lambda+1)+c}\\
&+ \frac{e^{\lambda(z-1)}}{a(\lambda^{2}+3\lambda+1)+b(\lambda+1)+c}\left(\frac{\partial}{\partial z}z\left(a\left((\lambda z)^{2}+3\lambda z+1\right)+b\left(\lambda z+1\right)+c\right)\right)\\
&= \frac{e^{\lambda(z-1)}\lambda z\left(a\left((\lambda z)^{2}+3\lambda z+1\right)+b(\lambda z+1)+c\right)}{a(\lambda^{2}+3\lambda+1)+b(\lambda+1)+c}\\
&+ \frac{e^{\lambda(z-1)}\left(a\left((\lambda z)^{2}+3\lambda z+1\right)+b(\lambda z+1)+c+z\left(a\left(2\lambda^{2}z+3\lambda\right)+b\lambda\right)\right)}{a(\lambda^{2}+3\lambda+1)+b(\lambda+1)+c}\\
&= \frac{e^{\lambda(z-1)}\left((\lambda z)^{3}a+(\lambda z)^{2}(6a+b)+\lambda z(7a+3b+c)+a+b+c\right)}{a(\lambda^{2}+3\lambda+1)+b(\lambda+1)+c}.
\end{aligned}$$

Consequently, it follows that

$$\begin{aligned}
E\left(N^{w}\right) &= \lim_{z\to -1}\frac{e^{\lambda(z-1)}\left((\lambda z)^{3}a+(\lambda z)^{2}(6a+b)+\lambda z(7a+3b+c)+a+b+c\right)}{a(\lambda^{2}+3\lambda+1)+b(\lambda+1)+c}\\
&= \frac{e^{\lambda(1-1)}\left(\lambda^{3}a+\lambda^{2}(6a+b)+\lambda(7a+3b+c)+a+b+c\right)}{a(\lambda^{2}+3\lambda+1)+b(\lambda+1)+c}\\
&= \frac{a\left(\lambda^{3}+6\lambda^{2}+7\lambda+1\right)+b\left(\lambda^{2}+3\lambda+1\right)+c(\lambda+1)}{a(\lambda^{2}+3\lambda+1)+b(\lambda+1)+c}.
\end{aligned}$$

From the definition of the variance, it follows that

$$Var\left(N^{w}\right) = \lim_{z\to -1}\frac{\partial^{2}}{\partial z^{2}}g\left(z\right)+E\left(N^{w}\right)-\left(E\left(N^{w}\right)\right)^{2}\ .$$

However,

$$
\begin{aligned}
\frac{\partial^2}{\partial z^2} g(z) &= \frac{\partial}{\partial z} \frac{e^{\lambda(z-1)}\left((\lambda z)^3 a+(\lambda z)^2(6a+b)+\lambda z(7a+3b+c)+a+b+c\right)}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c} \\
&= \left(\frac{\partial}{\partial z} e^{\lambda(z-1)}\right) \frac{(\lambda z)^3 a+(\lambda z)^2(6a+b)+\lambda z(7a+3b+c)+a+b+c}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c} \\
&+ \frac{e^{\lambda(z-1)}}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c}\left(\frac{\partial}{\partial z}(\lambda z)^3 a+(\lambda z)^2(6a+b)\right) \\
&+ \frac{e^{\lambda(z-1)}}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c}\left(\frac{\partial}{\partial z}\lambda z(7a+3b+c)+a+b+c\right) \\
&= \frac{e^{\lambda(z-1)}\lambda\left((\lambda z)^3 a+(\lambda z)^2(6a+b)+\lambda z(7a+3b+c)+a+b+c\right)}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c} \\
&+ \frac{e^{\lambda(z-1)}\left(3\lambda^3 z^2 a+2\lambda^2 z(6a+b)+\lambda(7a+3b+c)\right)}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c} \\
&= \frac{e^{\lambda(z-1)}\lambda\left((\lambda z)^3 a+(\lambda z)^2(6a+b)+\lambda z(7a+3b+c)+a+b+c+3\lambda^2 z^2 a+2\lambda z(6a+b)+(7a+3b+c)\right)}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c} \\
&= \frac{e^{\lambda(z-1)}\lambda\left((\lambda z)^3 a+(\lambda z)^2(6a+b)+\lambda z(19a+5b+c)+8a+4b+2c\right)}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c}.
\end{aligned}
$$

Consequently, it follows that

$$
\begin{aligned}
Var(N^w) &= \lim_{z\to -1} \frac{\partial^2}{\partial z^2} g(z) + E(N^w) - (E(N^w))^2 \\
&= \lim_{z\to -1} \frac{e^{\lambda(z-1)}\lambda\left((\lambda z)^3 a+(\lambda z)^2(9a+b)+\lambda z(19a+5b+c)+8a+4b+2c\right)}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c} \\
&+ E(N^w) - (E(N^w))^2 \\
&= \frac{e^{\lambda(1-1)}\lambda\left(\lambda^3 a+\lambda^2(9a+b)+\lambda(19a+5b+c)+8a+4b+2c\right)}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c} + E(N^w) - (E(N^w))^2 \\
&= \frac{\lambda\left(\lambda^3 a+\lambda^2(9a+b)+\lambda(19a+5b+c)+8a+4b+2c\right)}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c} + E(N^w) - (E(N^w))^2 \\
&= \frac{a\left(\lambda^4+9\lambda^3+19\lambda^2+8\lambda\right)+b\left(\lambda^3+5\lambda^2+4\lambda\right)+c\left(\lambda^2+2\lambda\right)}{a(\lambda^2+3\lambda+1)+b(\lambda+1)+c} + E(N^w) - (E(N^w))^2 .
\end{aligned}
$$

□

Theorem 10.17. *If the weight function used in the weighted Poisson probability mass func-*

tion is chosen as $w(n;\phi) = (n+a)(n-b)^2$ *then*

$$
\begin{aligned}
E(w(N;\phi)) &= \lambda(b^2 - 2ab - 2b + a + 1) + \lambda^2(\lambda + a - 2b + 3) + ab^2. \\
f_w(n) &= \frac{(n+a)(n-b)^2}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2}\frac{e^{-\lambda}\lambda^n}{n!}. \\
g(z) &= \frac{e^{\lambda(z-1)}\left((\lambda z)(b^2-2ab-2b+a+1)+(\lambda z)^2(\lambda z+a-2b+3)+ab^2\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2}. \\
E(N^w) &= \frac{\lambda\left(1+7\lambda+6\lambda^2+\lambda^3+b^2(1+\lambda)-2b\left(1+3\lambda+\lambda^2\right)+a\left(1+b^2+3\lambda+\lambda^2-2b(1+\lambda)\right)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2}. \\
Var(N^w) &= \frac{\lambda^2\left(8+19\lambda+9\lambda^2+\lambda^3+b^2(2+\lambda)-2b\left(4+5\lambda+\lambda^2\right)+a\left(4+b^2+5\lambda+\lambda^2-2b(2+\lambda)\right)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\
&+ E(N^w) - (E(N^w))^2.
\end{aligned}
$$

Proof. From the definition of the normalising constant, it follows that

$$
\begin{aligned}
E(w(N;\phi)) &= \sum_{k=0}^{\infty} w(k)\, f(k) \\
&= \sum_{k=0}^{\infty} (k+a)(k-b)^2 \frac{e^{-\lambda}\lambda^k}{k!} \\
&= \sum_{k=0}^{\infty} \left(k^3 + k^2(a-2b) + k(b^2-2ab) + ab^2\right)\frac{e^{-\lambda}\lambda^k}{k!} \\
&= \sum_{k=0}^{\infty} k^3 \frac{e^{-\lambda}\lambda^k}{k!} + (a-2b)\sum_{k=0}^{\infty} k^2 \frac{e^{-\lambda}\lambda^k}{k!} \\
&+ (b^2-2ab)\sum_{k=0}^{\infty} k \frac{e^{-\lambda}\lambda^k}{k!} + ab^2 \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!}.
\end{aligned}
$$

Since the first three sums in the above equations are the third, second and first moments respectively of a Poisson distribution with parameter λ, and the fourth sum equals 1, it follows that

$$
\begin{aligned}
E(w(N;\phi)) &= (\lambda^3 + 3\lambda^2 + \lambda) + (a-2b)(\lambda^2+\lambda) + (b^2-2ab)\lambda + ab^2 \\
&= \lambda(b^2 - 2ab - 2b + a + 1) + \lambda^2(\lambda + a - 2b + 3) + ab^2.
\end{aligned}
$$

From the definition of the weighted Poisson probability mass function, it follows that

$$
\begin{aligned}
f_w(n) &= \frac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\
&= \frac{(n+a)(n-b)^2}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2}\frac{e^{-\lambda}\lambda^n}{n!}.
\end{aligned}
$$

From the definition of the probability generating function, it follows that

$$\begin{aligned}
g(z) &= \textstyle\sum_{k=0}^{\infty} f_w(k)\, z^k \\
&= \textstyle\sum_{k=0}^{\infty} \frac{(k+a)(k-b)^2}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \frac{e^{-\lambda}\lambda^k}{k!} z^k \\
&= \frac{\sum_{k=0}^{\infty}\left(k^3+k^2(a-2b)+k\left(b^2-2ab\right)+ab^2\right)\frac{e^{-\lambda}\lambda^k}{k!} z^k}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\
&= \frac{e^{-\lambda}\left(\sum_{k=0}^{\infty}\left(k^3\frac{\lambda^k}{k!}z^k+k^2(a-2b)\frac{\lambda^k}{k!}z^k+k\left(b^2-2ab\right)\frac{\lambda^k}{k!}z^k+ab^2\frac{\lambda^k}{k!}z^k\right)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\
&= \frac{e^{-\lambda}\left(\sum_{k=0}^{\infty}k^3\frac{\lambda^k}{k!}z^k+\sum_{k=0}^{\infty}k^2(a-2b)\frac{\lambda^k}{k!}z^k+\sum_{k=0}^{\infty}k\left(b^2-2ab\right)\frac{\lambda^k}{k!}z^k+\sum_{k=0}^{\infty}ab^2\frac{\lambda^k}{k!}z^k\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\
&= \frac{\frac{e^{-\lambda}}{e^{-\lambda z}}\left(\sum_{k=0}^{\infty}k^3a_{k,\lambda,z}+(a-2b)\sum_{k=0}^{\infty}k^2a_{k,\lambda,z}+\left(b^2-2ab\right)\sum_{k=0}^{\infty}ka_{k,\lambda,z}+ab^2\sum_{k=0}^{\infty}a_{k,\lambda,z}\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2}
\end{aligned}$$

with $a_{k,\lambda,z} = \frac{(\lambda z)^k e^{-\lambda z}}{k!}$.
Since the first three sums in the above equations are the third, second and first moments respectively of a Poisson distribution with parameter λz, and the fourth sum equals 1, it follows that

$$\begin{aligned}
g(z) &= \frac{\frac{e^{-\lambda}}{e^{-\lambda z}}\left(\left((\lambda z)^3+3(\lambda z)^2+\lambda z\right)+(a-2b)\left((\lambda z)^2+\lambda\right)+\left(b^2-2ab\right)\lambda z+ab^2\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\
&= \frac{e^{\lambda(z-1)}\left((\lambda z)\left(b^2-2ab-2b+a+1\right)+(\lambda z)^2(\lambda z+a-2b+3)+ab^2\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2}.
\end{aligned}$$

From the definition of the expected value, it follows that

$$E(N^w) = \lim_{z\to^-1}\frac{\partial}{\partial z}g(z).$$

However,

$$\begin{aligned}
\frac{\partial}{\partial z}g(z) &= \frac{\partial}{\partial z}\frac{e^{\lambda(z-1)}(\lambda z)\left(\left(b^2-2ab-2b+a+1\right)+(\lambda z)(\lambda z+a-2b+3)+ab^2\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\
&= \left(\frac{\partial}{\partial z}e^{\lambda(z-1)}\right)\frac{\left((\lambda z)\left(b^2-2ab-2b+a+1\right)+(\lambda z)^2(\lambda z+a-2b+3)+ab^2\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\
&+ \frac{e^{\lambda(z-1)}}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\
&\times \frac{\partial}{\partial z}\left((\lambda z)\left(b^2-2ab-2b+a+1\right)+(\lambda z)^2(\lambda z+a-2b+3)+ab^2\right) \\
&= \frac{e^{\lambda(z-1)}\lambda\left((\lambda z)\left(b^2-2ab-2b+a+1\right)+(\lambda z)^2(\lambda z+a-2b+3)+ab^2\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\
&+ \frac{e^{\lambda(z-1)}\left(\lambda\left(b^2-2ab-2b+a+1\right)+\lambda^2 z(3\lambda z+2a-4b+6)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\
&= \frac{e^{\lambda(z-1)}\lambda\left(1+7z\lambda+6z^2\lambda^2+z^3\lambda^3+b^2(1+z\lambda)-2b\left(1+3z\lambda+z^2\lambda^2\right)+a\left(1+b^2+3z\lambda+z^2\lambda^2-2b(1+z\lambda)\right)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2}.
\end{aligned}$$

Consequently, it follows that

$$\begin{aligned} E\left(N^w\right) &= \lim_{z\to-1}\frac{e^{\lambda(z-1)}\lambda\left(1+7z\lambda+6z^2\lambda^2+z^3\lambda^3+b^2(1+z\lambda)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\ &+ \lim_{z\to-1}\frac{e^{\lambda(z-1)}\lambda\left(a\left(1+b^2+3z\lambda+z^2\lambda^2-2b(1+z\lambda)\right)-2b\left(1+3z\lambda+z^2\lambda^2\right)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\ &= \frac{\lambda\left(1+7\lambda+6\lambda^2+\lambda^3+b^2(1+\lambda)-2b\left(1+3\lambda+\lambda^2\right)+a\left(1+b^2+3\lambda+\lambda^2-2b(1+\lambda)\right)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2}. \end{aligned}$$

From the definition of the variance, it follows that

$$Var\left(N^w\right) = \lim_{z\to-1}\frac{\partial^2}{\partial z^2}g\left(z\right)+E\left(N^w\right)-\left(E\left(N^w\right)\right)^2.$$

However,

$$\begin{aligned} \frac{\partial^2}{\partial z^2}g\left(z\right) &= \frac{\partial}{\partial z}\frac{e^{\lambda(z-1)}\lambda\left(1+7z\lambda+6z^2\lambda^2+z^3\lambda^3+b^2(1+z\lambda)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\ &+ \frac{\partial}{\partial z}\frac{e^{\lambda(z-1)}\lambda\left(a\left(1+b^2+3z\lambda+z^2\lambda^2-2b(1+z\lambda)\right)-2b\left(1+3z\lambda+z^2\lambda^2\right)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\ &= \left(\frac{\partial}{\partial z}e^{\lambda(z-1)}\right)\left(\frac{\lambda\left(1+7z\lambda+6z^2\lambda^2+z^3\lambda^3+b^2(1+z\lambda)-2b\left(1+3z\lambda+z^2\lambda^2\right)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2}\right) \\ &+ \left(\frac{e^{\lambda(z-1)}+\lambda a\left(1+b^2+3z\lambda+z^2\lambda^2-2b(1+z\lambda)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2}\right) \\ &\times \frac{\partial}{\partial z}\left(\lambda\left(1+7z\lambda+6z^2\lambda^2+z^3\lambda^3+b^2\left(1+z\lambda\right)-2b\left(1+3z\lambda+z^2\lambda^2\right)\right)\right. \\ &\quad \left.+\lambda a\left(1+b^2+3z\lambda+z^2\lambda^2-2b\left(1+z\lambda\right)\right)\right) \\ &= \frac{e^{\lambda(z-1)}\lambda^2\left(z\lambda\left(1+b^2+3z\lambda+z^2\lambda^2-2b(1+z\lambda)\right)+a\left(b^2-2bz\lambda+z\lambda(1+z\lambda)\right)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\ &+ \frac{2e^{\lambda(z-1)}\lambda^2\left(1+a-2b-2ab+b^2+6z\lambda+2az\lambda-4bz\lambda+3z^2\lambda^2\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\ &+ \frac{e^{\lambda(z-1)}\lambda^2(6+2a-4b+6z\lambda)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\ &= \frac{e^{\lambda(z-1)}\lambda^2\left(8+19z\lambda+9z^2\lambda^2+z^3\lambda^3+b^2(2+z\lambda)-2b\left(4+5z\lambda+z^2\lambda^2\right)+a\left(4+b^2+5z\lambda+z^2\lambda^2-2b(2+z\lambda)\right)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2}. \end{aligned}$$

Consequently, it follows that

$$\begin{aligned} Var\left(N^w\right) &= \lim_{z\to-1}\frac{\partial^2}{\partial z^2}g\left(z\right)+E\left(N^w\right)-\left(E\left(N^w\right)\right)^2 \\ &= \lim_{z\to-1}\frac{e^{\lambda(z-1)}\lambda^2\left(8+19z\lambda+9z^2\lambda^2+z^3\lambda^3+b^2(2+z\lambda)-2b\left(4+5z\lambda+z^2\lambda^2\right)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\ &+ \lim_{z\to-1}\frac{e^{\lambda(z-1)}\lambda^2\left(a\left(4+b^2+5z\lambda+z^2\lambda^2-2b(2+z\lambda)\right)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2}+E\left(N^w\right)-\left(E\left(N^w\right)\right)^2 \\ &= \frac{\lambda^2\left(8+19\lambda+9\lambda^2+\lambda^3+b^2(2+\lambda)-2b\left(4+5\lambda+\lambda^2\right)+a\left(4+b^2+5\lambda+\lambda^2-2b(2+\lambda)\right)\right)}{\lambda(b^2-2ab-2b+a+1)+\lambda^2(\lambda+a-2b+3)+ab^2} \\ &+ E\left(N^w\right)-\left(E\left(N^w\right)\right)^2. \end{aligned}$$

□

Theorem 10.18. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n) = a + \frac{b-ac}{n+c}$ *then*

$$E(w(N;\phi)) = e^{-\lambda}(-\lambda)^{-c}\left(ae^{\lambda}(-\lambda)^{c} + (b-ac)\gamma(c,-\lambda)\right).$$

$$f_w(n) = \frac{(b+an)(-\lambda)^{c}\lambda^{n}}{(c+n)n!\left(ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)\right)}.$$

$$g(z) = \frac{(-\lambda)^{c}\left(ae^{z\lambda}(-z\lambda)^{c}+(b-ac)\gamma(c,-\lambda z)\right)}{\left(ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)\right)(-z\lambda)^{c}}.$$

$$E(N^w) = \frac{\left(e^{\lambda}(-\lambda)^{c}(b-ac+a\lambda)+c(ac-b)\gamma(c,-\lambda)\right)}{\left(ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)\right)}.$$

$$\begin{aligned} Var(N^w) &= \frac{(-\lambda)^{c}\left(e^{\lambda}(-\lambda)^{c}\left((1+c)(ac-b)+(b-ac)\lambda+a\lambda^{2}\right)+c(1+c)(b-ac)\gamma(c,-\lambda)\right)}{(-\lambda)^{c}\left(ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)\right)} \\ &+ E(N^w) - \left(E(N^w)\right)^{2}. \end{aligned}$$

Proof. From the definition of the normalising constant, it follows that

$$\begin{aligned} E(w(N;\phi)) &= \sum_{k=0}^{\infty} w(k) f(k) \\ &= \sum_{k=0}^{\infty}\left(a+\frac{b-ac}{k+c}\right)\frac{e^{-\lambda}\lambda^{k}}{k!} \\ &= \sum_{k=0}^{\infty} a\frac{e^{-\lambda}\lambda^{k}}{k!} + \sum_{k=0}^{\infty}\frac{b-ac}{k+c}\frac{e^{-\lambda}\lambda^{k}}{k!} \\ &= a\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^{k}}{k!} + e^{-\lambda}(b-ac)\sum_{k=0}^{\infty}\frac{\lambda^{k}}{(k+c)k!} \\ &= a\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^{k}}{k!} + e^{-\lambda}(b-ac)(-\lambda)^{-c}(-\lambda)^{c}\sum_{k=0}^{\infty}\frac{\lambda^{k}}{(k+c)k!}. \end{aligned}$$

Since the first infinite sum above equals 1 and $(-\lambda)^{c}\sum_{k=0}^{\infty}\frac{\lambda^{k}}{(k+c)k!} = \gamma(c,-\lambda)$ (See Theorem 10.10), it follows that

$$E(w(N;\phi)) = a + e^{-\lambda}(b-ac)(-\lambda)^{-c}\gamma(c,-\lambda).$$

From the definition of the weighted Poisson probability mass function, it follows that

$$\begin{aligned} f_w(n) &= \frac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\ &= \frac{\left(a+\frac{b-ac}{n+c}\right)\frac{e^{-\lambda}\lambda^{n}}{n!}}{a+e^{-\lambda}(b-ac)(-\lambda)^{-c}\gamma(c,-\lambda)} \\ &= \frac{\frac{an!(n+c)+(b-ac)e^{-\lambda}\lambda^{n}}{n!(n+c)}}{a+e^{-\lambda}(b-ac)(-\lambda)^{-c}\gamma(c,-\lambda)} \\ &= \frac{an!(n+c)+(b-ac)e^{-\lambda}\lambda^{n}}{(n+c)n!\left(a+e^{-\lambda}(b-ac)(-\lambda)^{-c}\gamma(c,-\lambda)\right)} \\ &= \frac{(b+an)(-\lambda)^{c}\lambda^{n}}{(c+n)n!\left(ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)\right)}. \end{aligned}$$

From the definition of the probability generating function, it follows that

$$\begin{aligned}
g(z) &= \sum_{k=0}^{\infty} f_w(k) z^k \\
&= \sum_{k=0}^{\infty} \frac{(b+ak)(-\lambda)^c \lambda^k}{(c+k)k!\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)} z^k \\
&= \frac{(-\lambda)^c}{\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)} \sum_{k=0}^{\infty} \frac{(b+ak)(\lambda z)^k}{(c+k)k!} \\
&= \frac{(-\lambda)^c}{\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)} \left(b\sum_{k=0}^{\infty} \frac{(\lambda z)^k}{(c+k)k!} + a\sum_{k=0}^{\infty} \frac{k(\lambda z)^k}{(c+k)k!}\right) \\
&= \frac{(-\lambda)^c}{\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)} \left(b\sum_{k=0}^{\infty} \frac{(\lambda z)^k}{(c+k)k!} + a\lambda z\sum_{k=1}^{\infty} \frac{(\lambda z)^{k-1}}{(c+k)(k-1)!}\right).
\end{aligned}$$

By reparameterising $m = k-1$ and using the fact that $\sum_{k=0}^{\infty} \frac{(\lambda z)^k}{(c+k)k!} = \frac{\gamma(c,-\lambda z)}{(-\lambda z)^c}$ and $\sum_{m=0}^{\infty} \frac{(\lambda z)^m}{(c+1+m)m!} = -\frac{\gamma(c+1,-\lambda z)}{(\lambda z)(-\lambda z)^c}$ (See Theorem 10.10), it follows that

$$\begin{aligned}
g(z) &= \frac{(-\lambda)^c}{\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)} \left(\frac{b}{(-\lambda z)^c}\gamma(c,-\lambda z) - a\frac{\gamma(c+1,-\lambda z)}{(-\lambda z)^c}\right) \\
&= \frac{(-\lambda)^c}{\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)} \left(\frac{b}{(-\lambda z)^c}\gamma(c,-\lambda z) - a\frac{\gamma(c+1,-\lambda z)}{(-\lambda z)^c}\right).
\end{aligned}$$

By using the relation $\gamma(\alpha+1,\beta) = \alpha\gamma(\alpha,\beta) - \beta^{\alpha}e^{-\beta}$ (See Theorem 10.10), it follows that

$$\begin{aligned}
g(z) &= \frac{(-\lambda)^c}{\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)} \left(\frac{b}{(-\lambda z)^c}\gamma(c,-\lambda z) - \frac{a}{(-\lambda z)^c}\left(c\gamma(c,-\lambda z) - (-\lambda z)^c e^{\lambda z}\right)\right) \\
&= \frac{(-\lambda)^c\left(ae^{z\lambda}(-\lambda z)^c+(b-ac)\gamma(c,-\lambda z)\right)}{\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)(-\lambda z)^c}.
\end{aligned}$$

From the definition of the expected value, it follows that

$$E(N^w) = \lim_{z\to -1} \frac{\partial}{\partial z} g(z).$$

However,

$$\begin{aligned}
\frac{\partial}{\partial z} g(z) &= \frac{\partial}{\partial z} \frac{(-\lambda)^c\left(ae^{z\lambda}(-\lambda z)^c+(b-ac)\gamma(c,-\lambda z)\right)}{\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)(-\lambda z)^c} \\
&= \frac{(-\lambda)^c}{\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)} \frac{\partial}{\partial z} \frac{\left(ae^{z\lambda}(-\lambda z)^c+(b-ac)\gamma(c,-\lambda z)\right)}{(-\lambda z)^c} \\
&= \frac{(-\lambda)^c}{ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)} \\
&\times \left(\frac{\left(ae^{z\lambda}(-\lambda z)^c+(b-ac)\gamma(c,-\lambda z)\right)}{c\lambda(-\lambda z)^{c+1}} + \frac{\left(-ace^{z\lambda}\lambda(-\lambda z)^{c-1}+ae^{z\lambda}\lambda(-\lambda z)^c-(b-ac)e^{z\lambda}\lambda(-\lambda z)^{c-1}\right)}{(-\lambda z)^c}\right) \\
&= \frac{(-\lambda)^c\left(e^{z\lambda}(-\lambda z)^c(b-ac+a\lambda z)+c(ac-b)\gamma(c,-\lambda z)\right)}{(-\lambda z)^c z\left(ae^{\lambda}(-\lambda)^c+(b-ac)\gamma(c,-\lambda)\right)}.
\end{aligned}$$

Note that the expression for the derivative of the lower incomplete gamma function can be found in Theorem 10.10.
Consequently, it follows that

$$\begin{aligned} E\left(N^{w}\right) &= \lim_{z\to-1}\frac{(-\lambda)^{c}\left(e^{z\lambda}(-\lambda z)^{c}(b-ac+a\lambda z)+c(ac-b)\gamma(c,-\lambda z)\right)}{(-\lambda z)^{c}z\left(ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)\right)} \\ &= \frac{\left(e^{\lambda}(-\lambda)^{c}(b-ac+a\lambda)+c(ac-b)\gamma(c,-\lambda)\right)}{\left(ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)\right)}. \end{aligned}$$

From the definition of the variance, it follows that

$$Var\left(N^{w}\right) = \lim_{z\to-1}\frac{\partial^{2}}{\partial z^{2}}g\left(z\right)+E\left(N^{w}\right)-\left(E\left(N^{w}\right)\right)^{2}.$$

However,

$$\begin{aligned} \frac{\partial^{2}}{\partial z^{2}}g\left(z\right) &= \frac{\partial}{\partial z}\frac{(-\lambda)^{c}\left(e^{z\lambda}(-\lambda z)^{c}(b-ac+a\lambda z)+c(ac-b)\gamma(c,-\lambda z)\right)}{(-\lambda z)^{c}z\left(ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)\right)} \\ &= \frac{(-\lambda)^{c}e^{\lambda z}(-\lambda z)^{c}(b(\lambda z+c-1)+a\lambda z(c+\lambda z))}{(-\lambda z)^{c}ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)} \\ &+ \frac{2c\lambda^{2}(-\lambda)^{c}e^{\lambda z}\left(a(-\lambda z)^{c}-b(-\lambda z)^{c-1}\right)}{(-\lambda z)^{c+1}ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)} \\ &- \frac{c(-c-1)(-\lambda)^{c}\lambda^{2}\left(ae^{\lambda z}(-\lambda z)^{c}+(b-ac)\gamma(c,-\lambda z)\right)}{(-\lambda z)^{c+2}ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)} \\ &= \frac{(-\lambda)^{c}\left(e^{\lambda z}(-\lambda z)^{c}\left((1+c)(ac-b)+(b-ac)\lambda z+az^{2}\lambda^{2}\right)+c(1+c)(b-ac)\gamma(c,-\lambda z)\right)}{z^{2}(-\lambda z)^{c}\left(ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)\right)}. \end{aligned}$$

Consequently, it follows that

$$\begin{aligned} Var\left(N^{w}\right) &= \lim_{z\to-1}\frac{(-\lambda)^{c}\left(e^{\lambda z}(-\lambda z)^{c}\left((1+c)(ac-b)+(b-ac)\lambda z+az^{2}\lambda^{2}\right)+c(1+c)(b-ac)\gamma(c,-\lambda z)\right)}{z^{2}(-\lambda z)^{c}\left(ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)\right)} \\ &+ E\left(N^{w}\right)-\left(E\left(N^{w}\right)\right)^{2} \\ &= \frac{(-\lambda)^{c}\left(e^{\lambda}(-\lambda)^{c}\left((1+c)(ac-b)+(b-ac)\lambda+a\lambda^{2}\right)+c(1+c)(b-ac)\gamma(c,-\lambda)\right)}{(-\lambda)^{c}\left(ae^{\lambda}(-\lambda)^{c}+(b-ac)\gamma(c,-\lambda)\right)} \\ &+ E\left(N^{w}\right)-\left(E\left(N^{w}\right)\right)^{2}. \end{aligned}$$

□

Theorem 10.19. *If the weight function used in the weighted Poisson probability mass func-*

tion is chosen as $w(n;\phi) = \begin{pmatrix} n+r-1 \\ n \end{pmatrix} p^n (1-p)^r$ *then*

$$
\begin{aligned}
E(w(N;\phi)) &= e^{-\lambda}(1-p)^r {}_1F_1(r;1;p\lambda). \\
f_w(n) &= \frac{\begin{pmatrix} n+r-1 \\ n \end{pmatrix}}{{}_1F_1(r;1;p\lambda)} \frac{(p\lambda)^n}{n!}. \\
g(z) &= \frac{{}_1F_1(r;1;pz\lambda)}{{}_1F_1(r;1;p\lambda)}. \\
E(N^w) &= \frac{\lambda p r {}_1F_1(r+1;2;p\lambda)}{{}_1F_1(r;1;p\lambda)}. \\
Var(N^w) &= \frac{\lambda^2 p^2 r(r+1) {}_1F_1(r+2;3;p\lambda)}{2 {}_1F_1(r;1;p\lambda)} + E(N^w) - (E(N^w))^2.
\end{aligned}
$$

where ${}_1F_1(;;)$ *is the confluent hypergeometric function in Definition 10.15.*

Proof. From the definition of the normalising constant, it follows that

$$
\begin{aligned}
E(w(N;\phi)) &= \sum_{k=0}^{\infty} w(k) f(k) \\
&= \sum_{k=0}^{\infty} \begin{pmatrix} k+r-1 \\ k \end{pmatrix} p^k (1-p)^r \frac{e^{-\lambda}\lambda^k}{k!} \\
&= \sum_{k=0}^{\infty} \frac{(k+r-1)!}{k!(r-1)!} p^k (1-p)^r \frac{e^{-\lambda}\lambda^k}{k!} \\
&= e^{-\lambda}(1-p)^r \sum_{k=0}^{\infty} \frac{\Gamma(k+r)}{\Gamma(k+1)(r-1)!} \frac{(p\lambda)^k}{k!} \\
&= e^{-\lambda}(1-p)^r \sum_{k=0}^{\infty} \frac{\Gamma(k+r)}{\Gamma(k+1)\Gamma(r)} \frac{(p\lambda)^k}{k!} \\
&= e^{-\lambda}(1-p)^r \sum_{k=0}^{\infty} \frac{\Gamma(k+r)}{\Gamma(r)} \frac{\Gamma(1)}{\Gamma(k+1)} \frac{(p\lambda)^k}{k!}.
\end{aligned}
$$

The sum in the above equation is ${}_1F_1(r;1;p\lambda)$, and thus it follows that

$$
E(w(N;\phi)) = e^{-\lambda}(1-p)^r {}_1F_1(r;1;p\lambda).
$$

From the definition of the weighted Poisson probability mass function, it follows that

$$
\begin{aligned}
f_w(n) &= \frac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\
&= \frac{\begin{pmatrix} n+r-1 \\ n \end{pmatrix} p^n(1-p)^r}{e^{-\lambda}(1-p)^r {}_1F_1(r;1;p\lambda)} \frac{e^{-\lambda}\lambda^n}{n!} \\
&= \frac{\begin{pmatrix} n+r-1 \\ n \end{pmatrix}}{{}_1F_1(r;1;p\lambda)} \frac{(p\lambda)^n}{n!}.
\end{aligned}
$$

From the definition of the probability generating function, it follows that

$$\begin{aligned} g(z) &= \textstyle\sum_{k=0}^{\infty} f_w(k)\, z^k \\ &= \textstyle\sum_{k=0}^{\infty} \frac{\binom{k+r-1}{k}}{{}_1F_1(r;1;p\lambda)} \frac{(p\lambda)^k}{k!} z^k \\ &= \frac{1}{{}_1F_1(r;1;p\lambda)} \textstyle\sum_{k=0}^{\infty} \binom{k+r-1}{k} \frac{(p\lambda z)^k}{k!} \\ &= \frac{1}{{}_1F_1(r;1;p\lambda)} \textstyle\sum_{k=0}^{\infty} \frac{\Gamma(k+r)}{\Gamma(k+1)\Gamma(r)} \frac{(p\lambda z)^k}{k!} \\ &= \frac{1}{{}_1F_1(r;1;p\lambda)} \textstyle\sum_{k=0}^{\infty} \frac{\Gamma(k+r)}{\Gamma(r)} \frac{\Gamma(1)}{\Gamma(k+1)} \frac{(p\lambda z)^k}{k!}. \end{aligned}$$

The sum in the above equation is ${}_1F_1(r;1;pz\lambda)$, and thus it follows that

$$g(z) = \frac{{}_1F_1(r;1;pz\lambda)}{{}_1F_1(r;1;p\lambda)}.$$

From the definition of the expected value, it follows that

$$E(N^w) = \lim_{z\to -1} \frac{\partial}{\partial z} g(z).$$

However,

$$\frac{\partial}{\partial z} g(z) = \frac{\partial}{\partial z} \frac{{}_1F_1(r;1;pz\lambda)}{{}_1F_1(r;1;p\lambda)}.$$

The derivative of confluent hypergeometric functions is well known with respect to its argument z. This result is given in Theorem 10.5. Consequently,

$$\begin{aligned} \frac{\partial}{\partial z} g(z) &= \frac{\frac{\partial}{\partial z} {}_1F_1(r;1;pz\lambda)}{{}_1F_1(r;1;p\lambda)} \\ &= \frac{r\,{}_1F_1(r+1;2;pz\lambda) \frac{\partial}{\partial z}(pz\lambda)}{{}_1F_1(r;1;p\lambda)} \\ &= \frac{\lambda p r\,{}_1F_1(r+1;2;pz\lambda)}{{}_1F_1(r;1;p\lambda)}. \end{aligned}$$

It then follows that

$$\begin{aligned} E(N^w) &= \lim_{z\to -1} \frac{\lambda p r\,{}_1F_1(r+1;2;pz\lambda)}{{}_1F_1(r;1;p\lambda)} \\ &= \frac{\lambda p r\,{}_1F_1(r+1;2;p\lambda)}{{}_1F_1(r;1;p\lambda)}. \end{aligned}$$

From the definition of the variance, it follows that

$$Var(N^w) = \lim_{z\to -1} \frac{\partial^2}{\partial z^2} g(z) + E(N^w) - (E(N^w))^2.$$

However,

$$\frac{\partial^2}{\partial z^2} g(z) = \frac{\partial}{\partial z} \frac{\lambda p r\,{}_1F_1(r+1;2;pz\lambda)}{{}_1F_1(r;1;p\lambda)}.$$

By using the result in Theorem 10.5, if follows that

$$\begin{aligned}\frac{\partial^2}{\partial z^2} g(z) &= \frac{\frac{r+1}{2}\lambda p r\, {}_1F_1(r+2;3;pz\lambda)\frac{\partial}{\partial z}(pz\lambda)}{{}_1F_1(r;1;p\lambda)} \\ &= \frac{\lambda^2 p^2 r(r+1)\,{}_1F_1(r+2;3;pz\lambda)}{2\,{}_1F_1(r;1;p\lambda)}.\end{aligned}$$

Consequently, the variance can be expressed as

$$\begin{aligned}Var(N^w) &= \lim_{z\to -1}\frac{\partial^2}{\partial z^2} g(z) + E(N^w) - (E(N^w))^2 \\ &= \lim_{z\to -1}\frac{\lambda^2 p^2 r(r+1)\,{}_1F_1(r+2;3;pz\lambda)}{2\,{}_1F_1(r;1;p\lambda)} + E(N^w) - (E(N^w))^2 \\ &= \frac{\lambda^2 p^2 r(r+1)\,{}_1F_1(r+2;3;p\lambda)}{2\,{}_1F_1(r;1;p\lambda)} + E(N^w) - (E(N^w))^2.\end{aligned}$$

□

Theorem 10.20. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \binom{m}{n} p^n (1-p)^{m-n}$ *then*

$$\begin{aligned}E(w(N;\phi)) &= e^{-\lambda}(1-p)^m L_m\left(\frac{p\lambda}{p-1}\right). \\ f_w(n) &= \binom{m}{n}\frac{(p\lambda)^n}{(1-p)^n n! L_m\left(\frac{p\lambda}{p-1}\right)}. \\ g(z) &= \frac{L_m\left(\frac{p\lambda z}{p-1}\right)}{L_m\left(\frac{p\lambda}{p-1}\right)}. \\ E(N^w) &= \frac{p\lambda L^1_{m-1}\left(\frac{p\lambda}{p-1}\right)}{(1-p)L_m\left(\frac{p\lambda}{p-1}\right)}. \\ Var(N^w) &= \frac{(p\lambda)^2 L^2_{m-2}\left(\frac{p\lambda}{p-1}\right)}{(p-1)^2 L_m\left(\frac{p\lambda}{p-1}\right)} \\ &+ E(N^w) - (E(N^w))^2.\end{aligned}$$

Where $L_m()$ *is the Laguerre polynomial and* $L^\alpha_m()$ *is the generalised Laguerre polynomial (Definition 10.19).*

Proof. Note that $\binom{m}{n} = 0$ when $n > m$. As a result $\sum_{n=0}^{\infty}\binom{m}{n} = \sum_{n=0}^{m}\binom{m}{n}$. This fact will be used often in the proofs for this specific weight function, as well as many other truncating weight functions.
From the definition of the normalising constant, it follows that

$$\begin{aligned} E\left(w\left(N;\phi\right)\right) &= \textstyle\sum_{k=0}^{\infty} w\left(k\right) f\left(k\right) \\ &= \textstyle\sum_{k=0}^{\infty} \left(\begin{array}{c} m \\ k \end{array} \right) p^k \left(1-p\right)^{m-k} \frac{e^{-\lambda}\lambda^k}{k!} \\ &= \textstyle\sum_{k=0}^{m} \left(\begin{array}{c} m \\ k \end{array} \right) p^k \left(1-p\right)^{m-k} \frac{e^{-\lambda}\lambda^k}{k!} \\ &= e^{-\lambda}\left(1-p\right)^m \textstyle\sum_{k=0}^{m} \left(\begin{array}{c} m \\ k \end{array} \right) \left(\frac{p}{1-p}\right)^k \frac{\lambda^k}{k!} \\ &= e^{-\lambda}\left(1-p\right)^m \textstyle\sum_{k=0}^{m} \left(\begin{array}{c} m \\ k \end{array} \right) \frac{\left(-1\right)^k \left(\frac{p\lambda}{p-1}\right)^k}{k!}. \end{aligned}$$

Since the sum is the Laguerre polynomial, $L_m\left(\frac{p\lambda}{p-1}\right)$, it follows that

$$E\left(w\left(N;\phi\right)\right) = e^{-\lambda}\left(1-p\right)^m L_m\left(\frac{p\lambda}{p-1}\right).$$

From the definition of the weighted Poisson probability mass function, it follows that

$$\begin{aligned} f_w\left(n\right) &= \frac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\ &= \frac{\left(\begin{array}{c} m \\ n \end{array} \right) p^n (1-p)^{m-n}}{e^{-\lambda}(1-p)^m L_m\left(\frac{p\lambda}{p-1}\right)} \frac{e^{-\lambda}\lambda^n}{n!} \\ &= \left(\begin{array}{c} m \\ n \end{array} \right) \frac{(p\lambda)^n}{(1-p)^n n! L_m\left(\frac{p\lambda}{p-1}\right)}. \end{aligned}$$

From the definition of the probability generating function, it follows that

$$\begin{aligned} g\left(z\right) &= \textstyle\sum_{k=0}^{\infty} f_w\left(k\right) z^k \\ &= \textstyle\sum_{k=0}^{\infty} \left(\begin{array}{c} m \\ k \end{array} \right) \frac{(p\lambda)^k}{(1-p)^k k! L_m\left(\frac{p\lambda}{p-1}\right)} z^k \\ &= \frac{1}{L_m\left(\frac{p\lambda}{p-1}\right)} \textstyle\sum_{k=0}^{m} \left(\begin{array}{c} m \\ k \end{array} \right) \frac{\left(\frac{p\lambda z}{1-p}\right)^k}{k!} \\ &= \frac{1}{L_m\left(\frac{p\lambda}{p-1}\right)} \textstyle\sum_{k=0}^{m} \left(\begin{array}{c} m \\ k \end{array} \right) \frac{(-1)^k \left(\frac{p\lambda z}{p-1}\right)^k}{k!}. \end{aligned}$$

Since the sum is the Laguerre polynomial it follows that

$$g\left(z\right) = \frac{L_m\left(\frac{p\lambda z}{p-1}\right)}{L_m\left(\frac{p\lambda}{p-1}\right)}.$$

From the definition of the expected value, it follows that

$$E\left(N^w\right) = \lim_{z\to -1}\frac{\partial}{\partial z}g\left(z\right).$$

However,

$$\frac{\partial}{\partial z}g\left(z\right) = \frac{\partial}{\partial z}\frac{L_m\left(\frac{p\lambda z}{p-1}\right)}{L_m\left(\frac{p\lambda}{p-1}\right)}$$

$$= \frac{1}{L_m\left(\frac{p\lambda}{p-1}\right)}\frac{\partial}{\partial z}L_m\left(\frac{p\lambda z}{p-1}\right).$$

From Theorem 10.11, it follows that

$$\frac{\partial}{\partial z}g\left(z\right) = -\frac{L^1_{m-1}\left(\frac{p\lambda z}{p-1}\right)}{L_m\left(\frac{p\lambda}{p-1}\right)}\frac{\partial}{\partial z}\frac{p\lambda z}{p-1}$$

$$= \frac{p\lambda L^1_{m-1}\left(\frac{p\lambda z}{p-1}\right)}{(1-p)L_m\left(\frac{p\lambda}{p-1}\right)}.$$

It then follows that

$$E\left(N^w\right) = \lim_{z\to -1}\frac{p\lambda L^1_{m-1}\left(\frac{p\lambda z}{p-1}\right)}{(1-p)L_m\left(\frac{p\lambda}{p-1}\right)}$$

$$= \frac{p\lambda L^1_{m-1}\left(\frac{p\lambda}{p-1}\right)}{(1-p)L_m\left(\frac{p\lambda}{p-1}\right)}.$$

From the definition of the variance, it follows that

$$Var\left(N^w\right) = \lim_{z\to -1}\frac{\partial^2}{\partial z^2}g\left(z\right) + E\left(N^w\right) - \left(E\left(N^w\right)\right)^2.$$

However,

$$\frac{\partial^2}{\partial z^2}g\left(z\right) = \frac{\partial}{\partial z}\frac{p\lambda L^1_{m-1}\left(\frac{p\lambda z}{p-1}\right)}{(1-p)L_m\left(\frac{p\lambda}{p-1}\right)}.$$

By using the result in Theorem 10.11, it follows that

$$\frac{\partial^2}{\partial z^2}g\left(z\right) = \frac{p^2\lambda^2 L^2_{m-2}\left(\frac{p\lambda z}{p-1}\right)}{(p-1)^2 L_m\left(\frac{p\lambda}{p-1}\right)}.$$

Consequently, the variance can be expressed as

$$Var\left(N^w\right) = \lim_{z\to -1}\frac{\partial^2}{\partial z^2}g\left(z\right) + E\left(N^w\right) - \left(E\left(N^w\right)\right)^2$$

$$= \lim_{z\to -1}\frac{p^2\lambda^2 L^2_{m-2}\left(\frac{p\lambda z}{p-1}\right)}{(p-1)^2 L_m\left(\frac{p\lambda}{p-1}\right)} + E\left(N^w\right) - \left(E\left(N^w\right)\right)^2$$

$$= \frac{p^2\lambda^2 L^2_{m-2}\left(\frac{p\lambda}{p-1}\right)}{(p-1)^2 L_m\left(\frac{p\lambda}{p-1}\right)} + E\left(N^w\right) - \left(E\left(N^w\right)\right)^2.$$

□

Theorem 10.21. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \frac{ab^a}{n^{a+1}}$ *then the following results follow*

$$\begin{aligned}
E(w(N;\phi)) &= \frac{ae^{-\lambda}\lambda^b({}_{a+2}F_{a+2}(1,b,...,b;1+b,1+b,...,1+b;\lambda))}{b!b}. \\
f_w(n) &= \frac{\lambda^{n-b}b^{a+1}b!}{n^{a+1}n!{}_{a+2}F_{a+2}(1,b,...,b;1+b,1+b,...,1+b;\lambda)}. \\
g(z) &= \frac{z^b{}_{a+2}F_{a+2}(1,b,...,b;1+b,1+b,...,1+b;\lambda z)}{{}_{a+2}F_{a+2}(1,b,...,b;1+b,1+b,...,1+b;\lambda)}. \\
E(N^w) &= \frac{b{}_{a+1}F_{a+1}(1,b,...,b;1+b,1+b,...,1+b;\lambda)}{{}_{a+2}F_{a+2}(1,b,...,b;1+b,1+b,...,1+b;\lambda)}. \\
Var(N^w) &= \frac{b^2{}_aF_a(1,b,...,b;1+b,1+b,...,1+b;\lambda)-b_{a+1}F_{a+1}(1,b,...,b;1+b,1+b,...,1+b;\lambda)}{{}_{a+2}F_{a+2}(1,b,...,b;1+b,1+b,...,1+b;\lambda)} \\
&+ E(N^w) - (E(N^w))^2.
\end{aligned}$$

Proof. From the definition of the normalising constant, it follows that

$$\begin{aligned}
E(w(N;\phi)) &= \sum_{k=0}^{\infty} w(k) f(k) \\
&= \sum_{k=b}^{\infty} \frac{ab^a}{k^{a+1}} \frac{e^{-\lambda}\lambda^k}{k!} \\
&= ab^a e^{-\lambda} \sum_{k=b}^{\infty} \frac{\lambda^k}{k^{a+1}k!} \\
&= ab^a e^{-\lambda} \sum_{k=b}^{\infty} \frac{\Gamma(k)^{a+1}\lambda^k}{\Gamma(k+1)^{a+1}k!}.
\end{aligned}$$

By reparameterising $m = k - b$, it follows that

$$\begin{aligned}
E(w(N;\phi)) &= ab^a e^{-\lambda} \sum_{m=0}^{\infty} \frac{\Gamma(m+b)^{a+1}\lambda^{(m+b)}}{\Gamma(m+b+1)^{a+1}(m+b)!} \\
&= ab^a e^{-\lambda}\lambda^b \sum_{m=0}^{\infty} \frac{\Gamma(m+b)^{a+1}\lambda^m}{\Gamma(m+b+1)^{a+2}} \\
&= ab^a e^{-\lambda}\lambda^b \sum_{m=0}^{\infty} \frac{\Gamma(m+1)\Gamma(m+b)^{a+1}}{\Gamma(m+b+1)^{a+2}} \frac{\lambda^m}{m!} \\
&= ab^a e^{-\lambda}\lambda^b \sum_{m=0}^{\infty} \frac{\Gamma(m+1)\Gamma(m+b)^{a+1}}{\Gamma(1)\Gamma(b+1)^{a+2}\Gamma(m+b+1)^{a+2}} \frac{\lambda^m}{m!} \\
&= \frac{ae^{-\lambda}\lambda^b}{b!b} \sum_{m=0}^{\infty} \frac{\Gamma(m+1)\Gamma(m+b)^{a+1}\Gamma(b+1)^{a+2}}{\Gamma(1)\Gamma(b)^{a+1}\Gamma(m+b+1)^{a+2}} \frac{\lambda^m}{m!} \\
&= \frac{ae^{-\lambda}\lambda^b({}_{a+2}F_{a+2}(1,b,...,b;1+b,1+b,...,1+b;\lambda))}{b!b}.
\end{aligned}$$

From the definition of the weighted Poisson probability mass function, it follows that

$$
\begin{aligned}
f_w(n) &= \frac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\
&= \frac{\frac{ab^a}{n^{a+1}}\frac{e^{-\lambda}\lambda^n}{n!}}{\frac{ae^{-\lambda}\lambda^b\left({}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)\right)}{b!b}} \\
&= \frac{ab^a}{n^{a+1}}\frac{e^{-\lambda}\lambda^n}{n!}\frac{b!b}{ae^{-\lambda}\lambda^b({}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda))} \\
&= \frac{\lambda^{n-b}b^{a+1}b!}{n^{a+1}n!({}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda))}.
\end{aligned}
$$

From the definition of the probability generating function, it follows that

$$
\begin{aligned}
g(z) &= \sum_{k=b}^{\infty} f_w(k)\, z^k \\
&= \sum_{k=b}^{\infty} \frac{\lambda^{k-b}b^{a+1}b!}{k^{a+1}k!({}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda))} z^k.
\end{aligned}
$$

By reparameterising $m = k - b$, it follows that

$$
\begin{aligned}
g(z) &= \frac{1}{({}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda))} \sum_{m=0}^{\infty} \frac{\lambda^m b^{a+1}b!}{(m+b)^{a+1}(m+b)!} z^{m+b} \\
&= \frac{z^b}{({}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda))} \sum_{m=0}^{\infty} \frac{b^{a+1}b!(\lambda z)^m}{(m+b)^{a+1}(m+b)!} \\
&= \frac{z^b}{({}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda))} \sum_{m=0}^{\infty} \frac{\Gamma(m+b)^{a+1}b^{a+1}b!(\lambda z)^m}{\Gamma(m+b+1)^{a+1}(m+b)!} \\
&= \frac{z^b}{({}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda))} \sum_{m=0}^{\infty} \frac{\Gamma(m+b)^{a+1}b^{a+1}b!(\lambda z)^m}{\Gamma(m+b+1)^{a+2}} \\
&= \frac{z^b}{({}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda))} \sum_{m=0}^{\infty} \frac{\Gamma(m+1)\Gamma(m+b)^{a+1}b^{a+1}b!}{\Gamma(1)\Gamma(m+b+1)^{a+2}}\frac{(\lambda z)^m}{m!} \\
&= \frac{z^b}{({}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda))} \sum_{m=0}^{\infty} \frac{\Gamma(m+1)\Gamma(m+b)^{a+1}\Gamma(m)^{a+2}}{\Gamma(1)\Gamma(m)^{a+1}\Gamma(m+b+1)^{a+2}}\frac{(\lambda z)^m}{m!} \\
&= \frac{z^b\,{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda z)}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)}.
\end{aligned}
$$

From the definition of the expected value, it follows that

$$
E(N^w) = \lim_{z\to -1}\frac{\partial}{\partial z} g(z).
$$

However,

$$
\frac{\partial}{\partial z} g(z) = \frac{\partial}{\partial z}\left(\frac{z^b\,{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda z)}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)}\right).
$$

By using Theorem 10.7, it follows that

$$
\begin{aligned}
\frac{\partial}{\partial z} g(z) &= \frac{bz^{b-1}\,{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda z)}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)} \\
&\quad - \frac{\frac{b}{z}({}_{a+1}F_{a+1}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda z) - {}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda z))}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)} \\
&= \frac{bz^{b-1}\,{}_{a+1}F_{a+1}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda z)}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)}.
\end{aligned}
$$

It then follows that

$$\begin{aligned} E(N^w) &= \lim_{z\to-1} \frac{bz^{b-1}{}_{a+1}F_{a+1}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda z)}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)} \\ &= \frac{b_{a+1}F_{a+1}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)}. \end{aligned}$$

From the definition of the variance, it follows that

$$Var(N^w) = \lim_{z\to-1} \frac{\partial^2}{\partial z^2} g(z) + E(N^w) - (E(N^w))^2.$$

However,

$$\begin{aligned} \frac{\partial^2}{\partial z^2} g(z) &= \frac{\partial}{\partial z}\left(\frac{bz^{b-1}{}_{a+1}F_{a+1}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda z)}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)}\right) \\ &= \frac{b(b-1)z^{b-2}{}_{a+1}F_{a+1}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda z)}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)} \\ &+ \frac{\frac{b}{z}({}_aF_a(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda z) - {}_{a+1}F_{a+1}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda z))}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)} \\ &= \frac{bz^{b-2}(b_aF_a(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda z) - {}_{a+1}F_{a+1}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda z))}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)}. \end{aligned}$$

Consequently, the variance can be expressed as

$$\begin{aligned} Var(N^w) &= \lim_{z\to-1} \frac{bz^{b-2}(b_aF_a(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda z) - {}_{a+1}F_{a+1}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda z))}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)} \\ &+ E(N^w) - (E(N^w))^2 \\ &= \frac{b^2{}_aF_a(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda) - b_{a+1}F_{a+1}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)}{{}_{a+2}F_{a+2}(1,b,\ldots,b;1+b,1+b,\ldots,1+b;\lambda)} \\ &+ E(N^w) - (E(N^w))^2. \end{aligned}$$

□

Theorem 10.22. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \frac{-1}{\ln(1-p)}\frac{p^n}{n}$ *then the following results follow*

$$\begin{aligned} E(w(N;\phi)) &= \frac{-p}{\ln(1-p)} e^{-\lambda} \lambda_2F_2(1,1;2,2;\lambda p). \\ f_w(n) &= \frac{(p\lambda)^{n-1}}{n.n!_2F_2(1,1;2,2;\lambda p)}. \\ g(z) &= \frac{z_2F_2(1,1;2,2;\lambda pz)}{{}_2F_2(1,1;2,2;\lambda p)}. \\ E(N^w) &= \frac{{}_1F_1(1;2;\lambda p)}{{}_2F_2(1,1;2,2;\lambda p)}. \\ Var(N^w) &= \frac{\lambda p_1F_1(2;3;\lambda p)}{2_2F_2(1,1;2,2;\lambda p)} \\ &+ E(N^w) - (E(N^w))^2. \end{aligned}$$

Proof. From the definition of the normalising constant, it follows that

$$\begin{aligned}
E\left(w\left(N;\phi\right)\right) &= \textstyle\sum_{k=0}^{\infty} w\left(k\right) f\left(k\right) \\
&= \textstyle\sum_{k=1}^{\infty} \frac{-1}{\ln(1-p)} \frac{p^k}{k} \frac{e^{-\lambda}\lambda^k}{k!} \\
&= \frac{-1}{\ln(1-p)} e^{-\lambda} \textstyle\sum_{k=1}^{\infty} \frac{(p\lambda)^k}{k!k}.
\end{aligned}$$

By reparameterising $m = k-1$, it follows that

$$\begin{aligned}
E\left(w\left(N;\phi\right)\right) &= \frac{-1}{\ln(1-p)} e^{-\lambda} \textstyle\sum_{m=0}^{\infty} \frac{(p\lambda)^{m+1}}{(m+1)!(m+1)} \\
&= \frac{-p\lambda}{\ln(1-p)} e^{-\lambda} \textstyle\sum_{m=0}^{\infty} \frac{(p\lambda)^{m}}{(m+1)!(m+1)} \\
&= \frac{-p\lambda}{\ln(1-p)} e^{-\lambda} \textstyle\sum_{m=0}^{\infty} \frac{(p\lambda)^{m}}{(m+1)(m+1)m!} \\
&= \frac{-p\lambda}{\ln(1-p)} e^{-\lambda} \textstyle\sum_{m=0}^{\infty} \frac{\Gamma(m+1)\Gamma(m+1)}{\Gamma(m+2)\Gamma(m+2)} \frac{(p\lambda)^{m}}{m!} \\
&= \frac{-p\lambda}{\ln(1-p)} e^{-\lambda} \textstyle\sum_{m=0}^{\infty} \frac{\Gamma(m+1)\Gamma(m+1)\Gamma(2)\Gamma(2)}{\Gamma(1)\Gamma(1)\Gamma(m+2)\Gamma(m+2)} \frac{(p\lambda)^{m}}{m!} \\
&= \frac{-p\lambda}{\ln(1-p)} e^{-\lambda} {}_2F_2\left(1,1;2,2;\lambda p\right).
\end{aligned}$$

From the definition of the weighted Poisson probability mass function, it follows that

$$\begin{aligned}
f_w\left(n\right) &= \frac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\
&= \frac{\frac{-1}{\ln(1-p)}\frac{p^n}{n}\frac{e^{-\lambda}\lambda^n}{n!}}{\frac{-p\lambda}{\ln(1-p)}e^{-\lambda}{}_2F_2(1,1;2,2;\lambda p)} \\
&= \frac{-1}{\ln(1-p)}\frac{p^n}{n}\frac{e^{-\lambda}\lambda^n}{n!}\frac{\ln(1-p)}{-p\lambda e^{-\lambda}{}_2F_2(1,1;2,2;\lambda p)} \\
&= \frac{(p\lambda)^{n-1}}{n.n!{}_2F_2(1,1;2,2;\lambda p)}.
\end{aligned}$$

From the definition of the probability generating function, it follows that

$$\begin{aligned}
g\left(z\right) &= \textstyle\sum_{k=0}^{\infty} f_w\left(k\right) z^k \\
&= \textstyle\sum_{k=1}^{\infty} \frac{(p\lambda)^{k-1}}{k.k!{}_2F_2(1,1;2,2;\lambda p)} z^k \\
&= \frac{1}{{}_2F_2(1,1;2,2;\lambda p)} \textstyle\sum_{k=1}^{\infty} \frac{(p\lambda)^{k-1}}{k.k!} z^k.
\end{aligned}$$

By reparameterising $m = k-1$, it follows that

$$\begin{aligned} g(z) &= \tfrac{1}{{}_2F_2(1,1;2,2;\lambda p)} \sum_{m=0}^{\infty} \tfrac{(p\lambda)^m}{(m+1).(m+1)!} z^{m+1} \\ &= \tfrac{z}{{}_2F_2(1,1;2,2;\lambda p)} \sum_{m=0}^{\infty} \tfrac{(p\lambda z)^m}{(m+1).(m+1)!} \\ &= \tfrac{z}{{}_2F_2(1,1;2,2;\lambda p)} \sum_{m=0}^{\infty} \tfrac{(p\lambda z)^m}{(m+1)(m+1)m!} \\ &= \tfrac{z}{{}_2F_2(1,1;2,2;\lambda p)} \sum_{m=0}^{\infty} \tfrac{\Gamma(m+1)\Gamma(m+1)}{\Gamma(m+2)\Gamma(m+2)} \tfrac{(p\lambda z)^m}{m!} \\ &= \tfrac{z}{{}_2F_2(1,1;2,2;\lambda p)} \sum_{m=0}^{\infty} \tfrac{\Gamma(m+1)\Gamma(m+1)\Gamma(2)\Gamma(2)}{\Gamma(1)\Gamma(1)\Gamma(m+2)\Gamma(m+2)} \tfrac{(p\lambda z)^m}{m!} \\ &= \tfrac{z\,{}_2F_2(1,1;2,2;\lambda pz)}{{}_2F_2(1,1;2,2;\lambda p)}. \end{aligned}$$

From the definition of the expected value, it follows that

$$E(N^w) = \lim_{z\to -1} \tfrac{\partial}{\partial z} g(z).$$

However, by using Theorem 10.7, it follows that

$$\begin{aligned} \tfrac{\partial}{\partial z} g(z) &= \tfrac{\partial}{\partial z}\left(\tfrac{z\,{}_2F_2(1,1;2,2;\lambda pz)}{{}_2F_2(1,1;2,2;\lambda p)}\right) \\ &= \tfrac{{}_2F_2(1,1;2,2;\lambda pz)}{{}_2F_2(1,1;2,2;\lambda p)} + \tfrac{{}_1F_1(1;2;\lambda pz) - {}_2F_2(1,1;2,2;\lambda pz)}{{}_2F_2(1,1;2,2;\lambda p)} \\ &= \tfrac{{}_1F_1(1;2;\lambda pz)}{{}_2F_2(1,1;2,2;\lambda p)}. \end{aligned}$$

It then follows that

$$\begin{aligned} E(N^w) &= \lim_{z\to -1} \tfrac{{}_1F_1(1;2;\lambda pz)}{{}_2F_2(1,1;2,2;\lambda p)} \\ &= \tfrac{{}_1F_1(1;2;\lambda p)}{{}_2F_2(1,1;2,2;\lambda p)}. \end{aligned}$$

From the definition of the variance, it follows that

$$Var(N^w) = \lim_{z\to -1} \tfrac{\partial^2}{\partial z^2} g(z) + E(N^w) - (E(N^w))^2.$$

However,

$$\begin{aligned} \tfrac{\partial^2}{\partial z^2} g(z) &= \tfrac{\partial}{\partial z}\left(\tfrac{{}_1F_1(1;2;\lambda pz)}{{}_2F_2(1,1;2,2;\lambda p)}\right) \\ &= \tfrac{\lambda p\,{}_1F_1(2;3;\lambda pz)}{2\,{}_2F_2(1,1;2,2;\lambda p)}. \end{aligned}$$

Consequently, the variance can be expressed as

$$\begin{aligned} Var(N^w) &= \lim_{z\to -1} \tfrac{\lambda p\,{}_1F_1(2;3;\lambda pz)}{2\,{}_2F_2(1,1;2,2;\lambda p)} + E(N^w) - (E(N^w))^2 \\ &= \tfrac{\lambda p\,{}_1F_1(2;3;\lambda p)}{2\,{}_2F_2(1,1;2,2;\lambda p)} + E(N^w) - (E(N^w))^2. \end{aligned}$$

□

Theorem 10.23. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \frac{\Gamma(r+n)}{n!\Gamma(r)} \frac{Beta(a+r,b+n)}{Beta(a,b)}$ *then the following results follow*

$$E(w(N;\phi)) = \frac{e^{-\lambda} Beta(a+r,b) {}_2F_2(b,r;1,a+b+r;\lambda)}{Beta(a,b)}.$$

$$f_w(n) = \frac{\lambda^n Beta(a+r,b+n)\Gamma(r+n)}{Beta(a+r,b)(n!)^2\Gamma(r) {}_2F_2(b,r;1,a+b+r;\lambda)}.$$

$$g(z) = \frac{{}_2F_2(b,r;1,a+b+r;\lambda z)}{{}_2F_2(b,r;1,a+b+r;\lambda)}.$$

$$E(N^w) = \frac{br\lambda {}_2F_2(1+b,1+r;2,1+a+b+r;\lambda)}{(a+b+r) {}_2F_2(b,r;1,a+b+r;\lambda)}.$$

$$Var(N^w) = \frac{b(1+b)r(1+r)\lambda^2 {}_2F_2(2+b,2+r;3,2+a+b+r;\lambda)}{2(a+b+r)(1+a+b+r) {}_2F_2(b,r;1,a+b+r;\lambda)} + E(N^w) - (E(N^w))^2.$$

Proof. From the definition of the normalising constant, it follows that

$$\begin{aligned}
E(w(N;\phi)) &= \sum_{k=0}^{\infty} w(k) f(k) \\
&= \sum_{k=0}^{\infty} \frac{\Gamma(r+k)}{k!\Gamma(r)} \frac{Beta(a+r,b+k)}{Beta(a,b)} \frac{e^{-\lambda}\lambda^k}{k!} \\
&= \frac{e^{-\lambda}}{Beta(a,b)} \sum_{k=0}^{\infty} \frac{\Gamma(r+k)}{k!\Gamma(r)} \frac{\Gamma(a+r)\Gamma(b+k)}{\Gamma(a+r+b+k)} \frac{\lambda^k}{k!} \\
&= \frac{e^{-\lambda}}{Beta(a,b)} \frac{\Gamma(a+r)\Gamma(b)}{\Gamma(a+r+b)} \sum_{k=0}^{\infty} \frac{\Gamma(a+r+b)}{\Gamma(a+r)\Gamma(b)} \frac{\Gamma(r+k)}{k!\Gamma(r)} \frac{\Gamma(a+r)\Gamma(b+k)}{\Gamma(a+r+b+k)} \frac{\lambda^k}{k!} \\
&= \frac{e^{-\lambda} Beta(a+r,b)}{Beta(a,b)} \sum_{k=0}^{\infty} \frac{\Gamma(b+k)\Gamma(r+k)\Gamma(1)\Gamma(a+b+r)}{\Gamma(b)\Gamma(r)\Gamma(1+k)\Gamma(a+b+r+k)} \frac{\lambda^k}{k!} \\
&= \frac{e^{-\lambda} Beta(a+r,b) {}_2F_2(b,r;1,a+b+r;\lambda)}{Beta(a,b)} \\
&= \frac{e^{-\lambda} Beta(a+r,b) {}_2F_2(b,r;1,a+b+r;\lambda)}{Beta(a,b)}.
\end{aligned}$$

From the definition of the weighted Poisson probability mass function, it follows that

$$\begin{aligned}
f_w(n) &= \frac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\
&= \frac{\frac{\Gamma(r+n)}{n!\Gamma(r)} \frac{Beta(a+r,b+n)}{Beta(a,b)} \frac{e^{-\lambda}\lambda^n}{n!}}{\frac{e^{-\lambda} Beta(a+r,b) {}_2F_2(b,r;1,a+b+r;\lambda)}{Beta(a,b)}} \\
&= \frac{\Gamma(r+n)}{n!\Gamma(r)} \frac{Beta(a+r,b+n)}{Beta(a,b)} \frac{e^{-\lambda}\lambda^n}{n!} \frac{Beta(a,b)}{e^{-\lambda} Beta(a+r,b) {}_2F_2(b,r;1,a+b+r;\lambda)} \\
&= \frac{\lambda^n Beta(a+r,b+n)\Gamma(r+n)}{Beta(a+r,b)(n!)^2\Gamma(r) {}_2F_2(b,r;1,a+b+r;\lambda)}.
\end{aligned}$$

From the definition of the probability generating, function it follows that

$$\begin{aligned} g(z) &= \sum_{k=0}^{\infty} f_w(k) z^k \\ &= \sum_{k=0}^{\infty} \frac{\lambda^k Beta(a+r,b+k)\Gamma(r+k)}{Beta(a+r,b)(k!)^2\Gamma(r)\,{}_2F_2(b,r;1,a+b+r;\lambda)} z^k \\ &= \frac{1}{{}_2F_2(b,r;1,a+b+r;\lambda)} \sum_{k=0}^{\infty} \frac{\Gamma(a+r)\Gamma(b+k)\Gamma(a+r+b)\Gamma(r+k)}{\Gamma(a+r+b+k)\Gamma(a+r)\Gamma(b)(k!)^2\Gamma(r)} (\lambda z)^k \\ &= \frac{1}{{}_2F_2(b,r;1,a+b+r;\lambda)} \sum_{k=0}^{\infty} \frac{\Gamma(b+k)\Gamma(r+k)\Gamma(k)\Gamma(a+r+b)}{\Gamma(b)\Gamma(r)\Gamma(k+1)\Gamma(a+r+b+k)} \frac{(\lambda z)^k}{k!} \\ &= \frac{{}_2F_2(b,r;1,a+b+r;\lambda z)}{{}_2F_2(b,r;1,a+b+r;\lambda)}. \end{aligned}$$

From the definition of the expected value, it follows that

$$E(N^w) = \lim_{z\to -1} \frac{\partial}{\partial z} g(z).$$

However, by using Theorem 10.6

$$\begin{aligned} \frac{\partial}{\partial z} g(z) &= \frac{\partial}{\partial z} \frac{{}_2F_2(b,r;1,a+b+r;\lambda z)}{{}_2F_2(b,r;1,a+b+r;\lambda)} \\ &= \frac{br\lambda\,{}_2F_2(1+b,1+r;2,1+a+b+r;\lambda z)}{(a+b+r)\,{}_2F_2(b,r;1,a+b+r;\lambda)}. \end{aligned}$$

It then follows that

$$\begin{aligned} E(N^w) &= \lim_{z\to -1} \frac{br\lambda\,{}_2F_2(1+b,1+r;2,1+a+b+r;\lambda z)}{(a+b+r)\,{}_2F_2(b,r;1,a+b+r;\lambda)} \\ &= \frac{br\lambda\,{}_2F_2(1+b,1+r;2,1+a+b+r;\lambda)}{(a+b+r)\,{}_2F_2(b,r;1,a+b+r;\lambda)}. \end{aligned}$$

From the definition of the variance, it follows that

$$Var(N^w) = \lim_{z\to -1} \frac{\partial^2}{\partial z^2} g(z) + E(N^w) - (E(N^w))^2 .$$

However,

$$\begin{aligned} \frac{\partial^2}{\partial z^2} g(z) &= \frac{\partial}{\partial z} \frac{br\lambda\,{}_2F_2(1+b,1+r;2,1+a+b+r;\lambda z)}{(a+b+r)\,{}_2F_2(b,r;1,a+b+r;\lambda)} \\ &= \frac{b(b+1)r(r+1)\lambda^2\,{}_2F_2(1+b,1+r;2,1+a+b+r;\lambda z)}{2(a+b+r)(a+b+r+1)\,{}_2F_2(b,r;1,a+b+r;\lambda)}. \end{aligned}$$

Consequently, the variance can be expressed as

$$\begin{aligned} Var(N^w) &= \lim_{z\to -1} \frac{b(b+1)r(r+1)\lambda^2\,{}_2F_2(1+b,1+r;2,1+a+b+r;\lambda z)}{2(a+b+r)(a+b+r+1)\,{}_2F_2(b,r;1,a+b+r;\lambda)} + E(N^w) - (E(N^w))^2 \\ &= \frac{b(b+1)r(r+1)\lambda^2\,{}_2F_2(1+b,1+r;2,1+a+b+r;\lambda)}{2(a+b+r)(a+b+r+1)\,{}_2F_2(b,r;1,a+b+r;\lambda)} + E(N^w) - (E(N^w))^2 \end{aligned}$$

□

Theorem 10.24. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = I(n \geq a)$ *then*

$$\begin{aligned} E(w(N;\phi)) &= \tfrac{\gamma(a,\lambda)}{\Gamma(a)}. \\ f_w(n) &= \tfrac{I(n\geq a)e^{-\lambda}\lambda^n\Gamma(a)}{n!\gamma(a,\lambda)}. \\ g(z) &= \tfrac{e^{\lambda(z-1)}\gamma(a,\lambda z)}{\gamma(a,\lambda)}. \\ E(N^w) &= \tfrac{e^{-\lambda}\lambda^a+\lambda\gamma(a,\lambda)}{\gamma(a,\lambda)}. \\ Var(N^w) &= \tfrac{e^{-\lambda}\left(\lambda^{1+a}+\lambda^a(a-1)+e^{\lambda}\lambda^2\gamma(a,\lambda)\right)}{\gamma(a,\lambda)} \\ &+ E(N^w) - \left(E(N^w)\right)^2. \end{aligned}$$

Proof. From the definition of the normalising constant, it follows that

$$\begin{aligned} E(w(N;\phi)) &= \textstyle\sum_{k=0}^{\infty} w(k) f(k) \\ &= \textstyle\sum_{k=0}^{\infty} I(k\geq a)\tfrac{e^{-\lambda}\lambda^k}{k!} \\ &= \textstyle\sum_{k=a}^{\infty} \tfrac{e^{-\lambda}\lambda^k}{k!}. \end{aligned}$$

Since $\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!} = 1$ (the Poisson probability mass function sums to one) it follows that

$$\sum_{k=a}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!} = 1 - \sum_{k=0}^{a-1}\frac{e^{-\lambda}\lambda^k}{k!}.$$

However, $\sum_{k=0}^{a-1}\frac{e^{-\lambda}\lambda^k}{k!}$ is the the cumulative distribution function of a Poisson distribution with parameter λ at the point $a-1$. The fact that this cumulative distribution function equals $\frac{\Gamma(a,\lambda)}{\Gamma(a)}$ can be found in many introductory statistical textbooks. Using the fact that $\Gamma(a) - \Gamma(a,\lambda) = \gamma(a,\lambda)$ (Theorem 10.10), it follows that

$$\begin{aligned} E(w(N;\phi)) &= 1 - \tfrac{\Gamma(a,\lambda)}{\Gamma(a)} \\ &= \tfrac{\gamma(a,\lambda)}{\Gamma(a)}. \end{aligned}$$

From the definition of the weighted Poisson probability mass function, it follows that

$$\begin{aligned} f_w(n) &= \tfrac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\ &= \tfrac{I(n\geq a)\frac{e^{-\lambda}\lambda^n}{n!}}{\frac{\gamma(a,\lambda)}{\Gamma(a)}} \\ &= \tfrac{I(n\geq a)e^{-\lambda}\lambda^n\Gamma(a)}{n!\gamma(a,\lambda)}. \end{aligned}$$

From the definition of the probability generating function, it follows that

$$\begin{aligned} g(z) &= \sum_{k=0}^{\infty} f_w(k) z^k \\ &= \sum_{k=0}^{\infty} \frac{I(k\geq a)e^{-\lambda}\lambda^k\Gamma(a)}{k!\gamma(a,\lambda)} z^k \\ &= \frac{e^{-\lambda}\Gamma(a)}{e^{-\lambda z}\gamma(a,\lambda)} \sum_{k=a}^{\infty} \frac{e^{-\lambda z}(\lambda z)^k}{k!} \\ &= \frac{e^{-\lambda}\Gamma(a)}{e^{-\lambda z}\gamma(a,\lambda)} \left(1 - \sum_{k=0}^{a-1} \frac{e^{-\lambda z}(\lambda z)^k}{k!}\right) \\ &= \frac{e^{-\lambda}\Gamma(a)}{e^{-\lambda z}(\Gamma(a)-\Gamma(a,\lambda))} \left(1 - \frac{\Gamma(a,\lambda z)}{\Gamma(a)}\right) \\ &= \frac{e^{\lambda(z-1)}\gamma(a,\lambda z)}{\gamma(a,\lambda)}. \end{aligned}$$

From the definition of the expected value, it follows that

$$E(N^w) = \lim_{z\to^{-}1} \frac{\partial}{\partial z} g(z).$$

However, by using Theorem 10.10 for the derivative of the lower incomplete gamma function, it follows that

$$\begin{aligned} \frac{\partial}{\partial z} g(z) &= \frac{\partial}{\partial z} \frac{e^{\lambda(z-1)}\gamma(a,\lambda z)}{\gamma(a,\lambda)} \\ &= \frac{\lambda e^{\lambda(z-1)}\gamma(a,\lambda z)+e^{\lambda(z-1)}(\lambda z)^{a-1}e^{-\lambda z}\lambda}{\gamma(a,\lambda)} \\ &= \frac{\lambda e^{\lambda(z-1)}\gamma(a,\lambda z)+e^{-\lambda}(\lambda z)^{a-1}\lambda}{\gamma(a,\lambda)}. \end{aligned}$$

It then follows that

$$\begin{aligned} E(N^w) &= \lim_{z\to^{-}1} \frac{\lambda e^{\lambda(z-1)}\gamma(a,\lambda z)+e^{-\lambda}(\lambda z)^{a-1}\lambda}{\gamma(a,\lambda)} \\ &= \frac{e^{-\lambda}\lambda^a+\lambda\gamma(a,\lambda)}{\gamma(a,\lambda)}. \end{aligned}$$

From the definition of the variance, it follows that

$$Var(N^w) = \lim_{z\to^{-}1} \frac{\partial^2}{\partial z^2} g(z) + E(N^w) - (E(N^w))^2.$$

However,

$$\begin{aligned} \frac{\partial^2}{\partial z^2} g(z) &= \frac{\partial}{\partial z} \frac{\lambda e^{\lambda(z-1)}\gamma(a,\lambda z)+e^{-\lambda}(\lambda z)^{a-1}\lambda}{\gamma(a,\lambda)} \\ &= \frac{\lambda^2 e^{\lambda(z-1)}\gamma(a,\lambda z)+\lambda e^{\lambda(z-1)}(\lambda z)^{a-1}e^{-\lambda z}\lambda+e^{-\lambda}(a-1)(\lambda z)^{a-2}\lambda^2}{\gamma(a,\lambda)} \\ &= \frac{\lambda^2 e^{\lambda(z-1)}\gamma(a,\lambda z)+\lambda e^{-\lambda}(\lambda z)^{a-1}\lambda+e^{-\lambda}(a-1)(\lambda z)^{a-2}\lambda^2}{\gamma(a,\lambda)}. \end{aligned}$$

Consequently, the variance can be expressed as

$$\begin{aligned} Var\left(N^w\right) &= \lim_{z\to -1}\tfrac{\lambda^2 e^{\lambda(z-1)}\gamma(a,\lambda z)+\lambda e^{-\lambda}(\lambda z)^{a-1}\lambda+e^{-\lambda}(a-1)(\lambda z)^{a-2}\lambda^2}{\gamma(a,\lambda)} \\ &+ E\left(N^w\right)-\left(E\left(N^w\right)\right)^2 \\ &= \tfrac{e^{-\lambda}\left(\lambda^{1+a}+\lambda^a(a-1)+e^{\lambda}\lambda^2\gamma(a,\lambda)\right)}{\gamma(a,\lambda)}+E\left(N^w\right)-\left(E\left(N^w\right)\right)^2. \end{aligned}$$

□

Theorem 10.25. *If the weight function used in the weighted Poisson probability mass function is chosen as $w(n;\phi)=I(n\leq b)$ then*

$$\begin{aligned} E\left(w\left(N;\phi\right)\right) &= \tfrac{\Gamma(1+b,\lambda)}{\Gamma(1+b)}. \\ f_w\left(n\right) &= \tfrac{I(n\leq b)e^{-\lambda}\lambda^n\Gamma(1+b)}{n!\Gamma(1+b,\lambda)}. \\ g\left(z\right) &= \tfrac{e^{\lambda(z-1)}\Gamma(1+b,\lambda z)}{\Gamma(1+b,\lambda)}. \\ E\left(N^w\right) &= \lambda-\tfrac{e^{-\lambda}\lambda^{1+b}}{\Gamma(1+b,\lambda)}. \\ Var\left(N^w\right) &= \lambda\left(\lambda-\tfrac{e^{-\lambda}\lambda^b(\lambda+b)}{\Gamma(1+b,\lambda)}\right)+E\left(N^w\right)-\left(E\left(N^w\right)\right)^2. \end{aligned}$$

Proof. From the definition of the normalising constant, it follows that

$$\begin{aligned} E\left(w\left(N;\phi\right)\right) &= \textstyle\sum_{k=0}^{\infty} w\left(k\right)f\left(k\right) \\ &= \textstyle\sum_{k=0}^{\infty} I\left(k\leq b\right)\tfrac{e^{-\lambda}\lambda^k}{k!} \\ &= \textstyle\sum_{k=0}^{b}\tfrac{e^{-\lambda}\lambda^k}{k!}. \end{aligned}$$

$\sum_{k=0}^{b}\frac{e^{-\lambda}\lambda^k}{k!}$ is the cumulative distribution function of a Poisson distribution with parameter λ at point b. Thus, it follows that

$$E\left(w\left(N;\phi\right)\right) = \tfrac{\Gamma(1+b,\lambda)}{\Gamma(1+b)}.$$

From the definition of the weighted Poisson probability mass function, it follows that

$$\begin{aligned} f_w\left(n\right) &= \tfrac{w(n)f(n)}{\sum_{k=0}^{\infty}w(k)f(k)} \\ &= \tfrac{I(k\leq b)\frac{e^{-\lambda}\lambda^n}{n!}}{\frac{\Gamma(1+b,\lambda)}{\Gamma(1+b)}} \\ &= \tfrac{I(n\leq b)e^{-\lambda}\lambda^n\Gamma(1+b)}{n!\Gamma(1+b,\lambda)}. \end{aligned}$$

From the definition of the probability generating function, it follows that

$$
\begin{aligned}
g(z) &= \textstyle\sum_{k=0}^{\infty} f_w(k)\, z^k \\
&= \textstyle\sum_{k=0}^{\infty} \frac{I(k\leq b)e^{-\lambda}\lambda^k\Gamma(1+b)}{k!\Gamma(1+b,\lambda)} z^k \\
&= \frac{e^{-\lambda}\Gamma(1+b)}{\Gamma(1+b,\lambda)} \textstyle\sum_{k=0}^{b} \frac{\lambda^k}{k!} z^k \\
&= \frac{e^{-\lambda}\Gamma(1+b)}{\Gamma(1+b,\lambda)e^{-\lambda z}} \textstyle\sum_{k=0}^{b} \frac{e^{-\lambda z}(\lambda z)^k}{k!} \\
&= \frac{e^{-\lambda}\Gamma(1+b)}{\Gamma(1+b,\lambda)e^{-\lambda z}} \frac{\Gamma(1+b,\lambda z)}{\Gamma(1+b)} \\
&= \frac{e^{\lambda(z-1)}\Gamma(1+b,\lambda z)}{\Gamma(1+b,\lambda)}.
\end{aligned}
$$

From the definition of the expected value, it follows that

$$E(N^w) = \lim_{z\to^-1} \frac{\partial}{\partial z} g(z).$$

However, by using Theorem 10.10 for the derivative of the upper incomplete gamma function, it follows that

$$
\begin{aligned}
\frac{\partial}{\partial z} g(z) &= \frac{\partial}{\partial z} \frac{e^{\lambda(z-1)}\Gamma(1+b,\lambda z)}{\Gamma(1+b,\lambda)} \\
&= \frac{\lambda e^{\lambda(z-1)}\Gamma(1+b,\lambda z) - e^{\lambda(z-1)}(\lambda z)^b e^{-\lambda z}\lambda}{\Gamma(1+b,\lambda)} \\
&= \frac{\lambda e^{\lambda(z-1)}\Gamma(1+b,\lambda z) - e^{-\lambda}(\lambda z)^b\lambda}{\Gamma(1+b,\lambda)}.
\end{aligned}
$$

It then follows that

$$
\begin{aligned}
E(N^w) &= \lim_{z\to^-1} \frac{\lambda e^{\lambda(z-1)}\Gamma(1+b,\lambda z) - e^{-\lambda}(\lambda z)^b\lambda}{\Gamma(1+b,\lambda)} \\
&= \lambda - \frac{e^{-\lambda}\lambda^{1+b}}{\Gamma(1+b,\lambda)}.
\end{aligned}
$$

From the definition of the variance, it follows that

$$Var(N^w) = \lim_{z\to^-1} \frac{\partial^2}{\partial z^2} g(z) + E(N^w) - (E(N^w))^2.$$

However,

$$
\begin{aligned}
\frac{\partial^2}{\partial z^2} g(z) &= \frac{\partial}{\partial z} \frac{\lambda e^{\lambda(z-1)}\Gamma(1+b,\lambda z) - e^{-\lambda}(\lambda z)^b\lambda}{\Gamma(1+b,\lambda)} \\
&= \frac{\lambda^2 e^{\lambda(z-1)}\Gamma(1+b,\lambda z) - \lambda^2 e^{\lambda(z-1)}(\lambda z)^b e^{-\lambda z} - e^{-\lambda}\lambda^{b+1} b z^{b-1}}{\Gamma(1+b,\lambda)} \\
&= \frac{\lambda^2 e^{\lambda(z-1)}\Gamma(1+b,\lambda z) - \lambda^2 e^{-\lambda}(\lambda z)^b - e^{-\lambda}\lambda^{b+1} b z^{b-1}}{\Gamma(1+b,\lambda)}.
\end{aligned}
$$

Consequently, the variance can be expressed as

$$\begin{aligned} Var\left(N^w\right) &= \lim_{z\to-1}\frac{\lambda^2 e^{\lambda(z-1)}\Gamma(1+b,\lambda z)-\lambda^2 e^{-\lambda}(\lambda z)^b-e^{-\lambda}\lambda^{b+1}bz^{b-1}}{\Gamma(1+b,\lambda)} \\ &+ E\left(N^w\right)-\left(E\left(N^w\right)\right)^2 \\ &= \lambda\left(\lambda-\frac{e^{-\lambda}\lambda^b(\lambda+b)}{\Gamma(1+b,\lambda)}\right)+E\left(N^w\right)-\left(E\left(N^w\right)\right)^2. \end{aligned}$$

□

Theorem 10.26. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi)=I(n\geq a)\,I(n\leq b)$ *then*

$$\begin{aligned} E\left(w\left(N;\phi\right)\right) &= \frac{\Gamma(a)\Gamma(1+b,\lambda)-\Gamma(1+b)\Gamma(a,\lambda)}{\Gamma(a)\Gamma(1+b)}. \\ f_w\left(n\right) &= \frac{\Gamma(a)\Gamma(1+b)}{\Gamma(a)\Gamma(1+b,\lambda)-\Gamma(1+b)\Gamma(a,\lambda)}\frac{I(n\geq a)I(n\leq b)e^{-\lambda}\lambda^n}{n!}. \\ g\left(z\right) &= \frac{e^{\lambda(z-1)}\left(\Gamma(1+b)\Gamma(a,\lambda z)-\Gamma(a)\Gamma(1+b,\lambda z)\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)}. \\ E\left(N^w\right) &= \frac{e^{-\lambda}\left(\lambda\Gamma(a)\left(\lambda^b-e^{\lambda}\Gamma(1+b,\lambda)\right)-\Gamma(1+b)\left(\lambda^a-e^{\lambda}\lambda\Gamma(a,\lambda)\right)\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)}. \\ Var\left(N^w\right) &= \frac{e^{-\lambda}\left(\left(\lambda^{2+b}+b\lambda^{1+b}\right)\Gamma(a)+\left(\lambda^a-\lambda^{1+a}-a\lambda^a\right)\Gamma(1+b)+e^{\lambda}\lambda^2\left(\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)\right)\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)} \\ &+ E\left(N^w\right)-\left(E\left(N^w\right)\right)^2. \end{aligned}$$

Proof. From the definition of the normalising constant, it follows that

$$\begin{aligned} E\left(w\left(N;\phi\right)\right) &= \sum_{k=0}^{\infty} w\left(k\right)f\left(k\right) \\ &= \sum_{k=0}^{\infty} I\left(n\geq a\right)I\left(n\leq b\right)\frac{e^{-\lambda}\lambda^k}{k!} \\ &= \sum_{k=a}^{b}\frac{e^{-\lambda}\lambda^k}{k!}. \end{aligned}$$

Since $\sum_{k=a}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!}+\sum_{k=0}^{b}\frac{e^{-\lambda}\lambda^k}{k!}-\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!}=\sum_{k=a}^{b}\frac{e^{-\lambda}\lambda^k}{k!}$ the above summation can be split into three parts.

$$\begin{aligned} E\left(w\left(N;\phi\right)\right) &= \sum_{k=a}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!}+\sum_{k=0}^{b}\frac{e^{-\lambda}\lambda^k}{k!}-\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!} \\ &= \frac{\gamma(a,\lambda)}{\Gamma(a)}+\frac{\Gamma(1+b,\lambda)}{\Gamma(1+b)}-1 \\ &= \frac{\gamma(a,\lambda)\Gamma(1+b)+\Gamma(1+b,\lambda)\Gamma(a)-\Gamma(a)\Gamma(1+b)}{\Gamma(a)\Gamma(1+b)} \\ &= \frac{(\Gamma(a)-\Gamma(a,\lambda))\Gamma(1+b)+\Gamma(1+b,\lambda)\Gamma(a)-\Gamma(a)\Gamma(1+b)}{\Gamma(a)\Gamma(1+b)} \\ &= \frac{\Gamma(a)\Gamma(1+b,\lambda)-\Gamma(1+b)\Gamma(a,\lambda)}{\Gamma(a)\Gamma(1+b)}. \end{aligned}$$

From the definition of the weighted Poisson probability mass function, it follows that

$$
\begin{aligned}
f_w(n) &= \frac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\
&= \frac{I(n\geq a)I(n\leq b)\frac{e^{-\lambda}\lambda^n}{n!}}{\frac{\Gamma(a)\Gamma(1+b,\lambda)-\Gamma(1+b)\Gamma(a,\lambda)}{\Gamma(a)\Gamma(1+b)}} \\
&= \frac{\Gamma(a)\Gamma(1+b)}{\Gamma(a)\Gamma(1+b,\lambda)-\Gamma(1+b)\Gamma(a,\lambda)}\frac{I(n\geq a)I(n\leq b)e^{-\lambda}\lambda^n}{n!}.
\end{aligned}
$$

From the definition of the probability generating function, it follows that

$$
\begin{aligned}
g(z) &= \sum_{k=0}^{\infty} f_w(k)\, z^k \\
&= \sum_{k=0}^{\infty} \frac{\Gamma(a)\Gamma(1+b)}{\Gamma(a)\Gamma(1+b,\lambda)-\Gamma(1+b)\Gamma(a,\lambda)}\frac{I(n\geq a)I(n\leq b)e^{-\lambda}\lambda^n}{n!} z^k \\
&= \frac{\Gamma(a)\Gamma(1+b)e^{-\lambda}}{\Gamma(a)\Gamma(1+b,\lambda)-\Gamma(1+b)\Gamma(a,\lambda)} \sum_{k=a}^{b} \frac{(\lambda z)^k}{k!} \\
&= \frac{\Gamma(a)\Gamma(1+b)e^{-\lambda}}{\Gamma(a)\Gamma(1+b,\lambda)-\Gamma(1+b)\Gamma(a,\lambda)e^{-\lambda z}} \sum_{k=a}^{b} \frac{e^{-\lambda z}(\lambda z)^k}{k!}.
\end{aligned}
$$

Using the same reasoning as the normalising constant derivation, it follows that

$$
\begin{aligned}
g(z) &= \frac{\Gamma(a)\Gamma(1+b)e^{-\lambda}}{\Gamma(a)\Gamma(1+b,\lambda)-\Gamma(1+b)\Gamma(a,\lambda)e^{-\lambda z}} \left(\sum_{k=a}^{\infty} \frac{e^{-\lambda z}(\lambda z)^k}{k!}\right. \\
&+ \left.\sum_{k=0}^{b} \frac{e^{-\lambda z}(\lambda z)^k}{k!} - \sum_{k=0}^{\infty} \frac{e^{-\lambda z}(\lambda z)^k}{k!}\right) \\
&= \frac{\Gamma(a)\Gamma(1+b)e^{-\lambda}}{\Gamma(a)\Gamma(1+b,\lambda)-\Gamma(1+b)\Gamma(a,\lambda)e^{-\lambda z}} \left(\left(1-\sum_{k=0}^{a-1} \frac{e^{-\lambda z}(\lambda z)^k}{k!}\right)\right. \\
&+ \left.\sum_{k=0}^{b} \frac{e^{-\lambda z}(\lambda z)^k}{k!} - \sum_{k=0}^{\infty} \frac{e^{-\lambda z}(\lambda z)^k}{k!}\right) \\
&= \frac{\Gamma(a)\Gamma(1+b)e^{-\lambda}}{\Gamma(a)\Gamma(1+b,\lambda)-\Gamma(1+b)\Gamma(a,\lambda)e^{-\lambda z}} \left(\left(1-\frac{\Gamma(a,\lambda z)}{\Gamma(a)}\right)+\frac{\Gamma(1+b,\lambda z)}{\Gamma(1+b)}-1\right) \\
&= \frac{e^{\lambda(z-1)}\Gamma(a)\Gamma(1+b)}{\Gamma(a)\Gamma(1+b,\lambda)-\Gamma(1+b)\Gamma(a,\lambda)} \left(\frac{\Gamma(a,\lambda z)\Gamma(1+b)+\Gamma(a)\Gamma(1+b,\lambda z)}{\Gamma(a)\Gamma(1+b)}\right) \\
&= \frac{e^{\lambda(z-1)}(\Gamma(1+b)\Gamma(a,\lambda z)-\Gamma(a)\Gamma(1+b,\lambda z))}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)}.
\end{aligned}
$$

From the definition of the expected value, it follows that

$$
E(N^w) = \lim_{z\to -1} \frac{\partial}{\partial z} g(z).
$$

However by using Theorem 10.10 for the derivative of the upper incomplete gamma function, it follows that

$$\begin{aligned}
\frac{\partial}{\partial z}g\left(z\right) &= \frac{\partial}{\partial z}\frac{e^{\lambda(z-1)}(\Gamma(1+b)\Gamma(a,\lambda z)-\Gamma(a)\Gamma(1+b,\lambda z))}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)} \\
&= \frac{\lambda e^{\lambda(z-1)}(\Gamma(1+b)\Gamma(a,\lambda z)-\Gamma(a)\Gamma(1+b,\lambda z))+e^{\lambda(z-1)}\left(-\Gamma(1+b)(\lambda z)^{a-1}e^{-\lambda z}\lambda-\Gamma(a)(\lambda z)^{b}e^{-\lambda z}\lambda\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)} \\
&= \frac{\lambda e^{\lambda(z-1)}(\Gamma(1+b)\Gamma(a,\lambda z)-\Gamma(a)\Gamma(1+b,\lambda z))-e^{-\lambda}\left(\Gamma(1+b)(\lambda z)^{a-1}\lambda+\Gamma(a)(\lambda z)^{b}\lambda\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)}.
\end{aligned}$$

It then follows that

$$\begin{aligned}
E\left(N^{w}\right) &= \lim_{z\to-1}\frac{\lambda e^{\lambda(z-1)}(\Gamma(1+b)\Gamma(a,\lambda z)-\Gamma(a)\Gamma(1+b,\lambda z))-e^{-\lambda}\left(\Gamma(1+b)(\lambda z)^{a-1}\lambda+\Gamma(a)(\lambda z)^{b}\lambda\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)} \\
&= \frac{\lambda(\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda))-e^{-\lambda}\left(\Gamma(1+b)(\lambda)^{a-1}\lambda+\Gamma(a)(\lambda)^{b}\lambda\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)} \\
&= \frac{e^{-\lambda}\left(\lambda\Gamma(a)\left(\lambda^{b}-e^{\lambda}\Gamma(1+b,\lambda)\right)-\Gamma(1+b)\left(\lambda^{a}-e^{\lambda}\lambda\Gamma(a,\lambda)\right)\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)}.
\end{aligned}$$

From the definition of the variance, it follows that

$$Var\left(N^{w}\right) = \lim_{z\to-1}\frac{\partial^2}{\partial z^2}g\left(z\right)+E\left(N^{w}\right)-\left(E\left(N^{w}\right)\right)^{2}.$$

However,

$$\begin{aligned}
\frac{\partial^2}{\partial z^2}g\left(z\right) &= \frac{\partial}{\partial z}\frac{\lambda e^{\lambda(z-1)}(\Gamma(1+b)\Gamma(a,\lambda z)-\Gamma(a)\Gamma(1+b,\lambda z))-e^{-\lambda}\left(\Gamma(1+b)(\lambda z)^{a-1}\lambda+\Gamma(a)(\lambda z)^{b}\lambda\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)} \\
&= \frac{\lambda^{2}e^{\lambda(z-1)}(\Gamma(1+b)\Gamma(a,\lambda z)-\Gamma(a)\Gamma(1+b,\lambda z))+\lambda e^{\lambda(z-1)}\left(-\Gamma(1+b)(\lambda z)^{a-1}e^{\lambda z}\lambda+\Gamma(a)(\lambda z)^{b}e^{\lambda z}\lambda\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)} \\
&- \frac{e^{-\lambda}\left((a-1)\Gamma(1+b)(\lambda z)^{a-2}\lambda^{2}+\Gamma(a)b(\lambda z)^{b-1}\lambda^{2}\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)} \\
&= \frac{\lambda^{2}e^{\lambda(z-1)}(\Gamma(1+b)\Gamma(a,\lambda z)-\Gamma(a)\Gamma(1+b,\lambda z))+\lambda^{2}e^{-\lambda}\left(\Gamma(a)(\lambda z)^{b}-\Gamma(1+b)(\lambda z)^{a-1}\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)} \\
&- \frac{e^{-\lambda}\left((a-1)\Gamma(1+b)(\lambda z)^{a-2}\lambda^{2}+\Gamma(a)b(\lambda z)^{b-1}\lambda^{2}\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)}.
\end{aligned}$$

Consequently, the variance can be expressed as

$$\begin{aligned}
Var\left(N^{w}\right) &= \lim_{z\to-1}\frac{\lambda^{2}e^{\lambda(z-1)}(\Gamma(1+b)\Gamma(a,\lambda z)-\Gamma(a)\Gamma(1+b,\lambda z))+\lambda^{2}e^{-\lambda}\left(\Gamma(a)(\lambda z)^{b}-\Gamma(1+b)(\lambda z)^{a-1}\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)} \\
&- \lim_{z\to-1}\frac{e^{-\lambda}\left((a-1)\Gamma(1+b)(\lambda z)^{a-2}\lambda^{2}+\Gamma(a)b(\lambda z)^{b-1}\lambda^{2}\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)}+E\left(N^{w}\right)-\left(E\left(N^{w}\right)\right)^{2} \\
&= \frac{\lambda^{2}(\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda))+\lambda^{2}e^{-\lambda}\left(\Gamma(a)\lambda^{b}-\Gamma(1+b)\lambda^{a-1}\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)} \\
&- \frac{e^{-\lambda}\left((a-1)\Gamma(1+b)\lambda^{a}+\Gamma(a)b\lambda^{b+1}\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)}+E\left(N^{w}\right)-\left(E\left(N^{w}\right)\right)^{2} \\
&= \frac{e^{-\lambda}\left(\left(\lambda^{2+b}+b\lambda^{1+b}\right)\Gamma(a)+\left(\lambda^{a}-\lambda^{1+a}-a\lambda^{a}\right)\Gamma(1+b)+e^{\lambda}\lambda^{2}(\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda))\right)}{\Gamma(1+b)\Gamma(a,\lambda)-\Gamma(a)\Gamma(1+b,\lambda)} \\
&+ E\left(N^{w}\right)-\left(E\left(N^{w}\right)\right)^{2}.
\end{aligned}$$

□

Theorem 10.27. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = \begin{pmatrix} n \\ a \end{pmatrix}$ *where* $a \in \{0, 1, 2, ...\}$ *then*

$$\begin{aligned} E(w(N;\phi)) &= \tfrac{\lambda^a}{a!}. \\ f_w(n) &= \tfrac{e^{-\lambda}\lambda^{n-a}}{(n-a)!}. \\ g(z) &= e^{\lambda(z-1)}z^a. \\ E(N^w) &= a+\lambda. \\ Var(N^w) &= \lambda. \end{aligned}$$

Proof. From the definition of the normalising constant, it follows that

$$\begin{aligned} E(w(N;\phi)) &= \textstyle\sum_{k=0}^{\infty} w(k)\, f(k) \\ &= \textstyle\sum_{k=0}^{\infty} \begin{pmatrix} k \\ a \end{pmatrix} \tfrac{e^{-\lambda}\lambda^k}{k!} \\ &= \textstyle\sum_{k=a}^{\infty} \begin{pmatrix} k \\ a \end{pmatrix} \tfrac{e^{-\lambda}\lambda^k}{k!} \\ &= \textstyle\sum_{k=a}^{\infty} \tfrac{k!}{a!(k-a)!} \tfrac{e^{-\lambda}\lambda^k}{k!} \\ &= \tfrac{e^{-\lambda}}{a!} \textstyle\sum_{k=a}^{\infty} \tfrac{\lambda^k}{(k-a)!}. \end{aligned}$$

By reparameterising $m = k - a$, it follows that

$$\begin{aligned} E(w(N;\phi)) &= \tfrac{e^{-\lambda}}{a!} \textstyle\sum_{m=0}^{\infty} \tfrac{\lambda^{m+a}}{(m)!} \\ &= \tfrac{e^{-\lambda}\lambda^a}{a!} \textstyle\sum_{m=0}^{\infty} \tfrac{\lambda^m}{(m)!} \\ &= \tfrac{e^{-\lambda}\lambda^a}{a!} e^{\lambda} \\ &= \tfrac{\lambda^a}{a!}. \end{aligned}$$

From the definition of the weighted Poisson probability mass function, it follows that

$$
\begin{aligned}
f_w(n) &= \frac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\
&= \frac{\begin{pmatrix} n \\ a \end{pmatrix} \frac{e^{-\lambda}\lambda^n}{n!}}{\frac{\lambda^a}{a!}} \\
&= \frac{n!}{a!(n-a)!}\frac{e^{-\lambda}\lambda^n}{n!}\frac{a!}{\lambda^a} \\
&= \frac{e^{-\lambda}\lambda^{n-a}}{(n-a)!}.
\end{aligned}
$$

From the definition of the probability generating function, it follows that

$$
\begin{aligned}
g(z) &= \sum_{k=0}^{\infty} f_w(k)\, z^k \\
&= \sum_{k=a}^{\infty} \frac{e^{-\lambda}\lambda^{k-a}}{(k-a)!} z^k.
\end{aligned}
$$

By reparameterising $m = k - a$, it follows that

$$
\begin{aligned}
g(z) &= e^{-\lambda} z^a \sum_{m=0}^{\infty} \frac{(\lambda z)^m}{m!} \\
&= e^{\lambda(z-1)} z^a.
\end{aligned}
$$

From the definition of the expected value, it follows that

$$
E(N^w) = \lim_{z\to -1} \frac{\partial}{\partial z} g(z).
$$

However,

$$
\begin{aligned}
\frac{\partial}{\partial z} g(z) &= \frac{\partial}{\partial z}\left(e^{\lambda(z-1)} z^a\right) \\
&= \lambda e^{\lambda(z-1)} z^a + e^{\lambda(z-1)} a z^{a-1} \\
&= e^{\lambda(z-1)}\left(\lambda z^a + a z^{a-1}\right).
\end{aligned}
$$

It then follows that

$$
\begin{aligned}
E(N^w) &= \lim_{z\to -1} e^{\lambda(z-1)}\left(\lambda z^a + a z^{a-1}\right) \\
&= a + \lambda.
\end{aligned}
$$

From the definition of the variance, it follows that

$$
Var(N^w) = \lim_{z\to -1} \frac{\partial^2}{\partial z^2} g(z) + E(N^w) - \left(E(N^w)\right)^2.
$$

However,

$$
\begin{aligned}
\frac{\partial^2}{\partial z^2} g(z) &= \frac{\partial}{\partial z}\left(e^{\lambda(z-1)}\left(\lambda z^a + a z^{a-1}\right)\right) \\
&= \lambda e^{\lambda(z-1)}\left(\lambda z^a + a z^{a-1}\right) + e^{\lambda(z-1)}\left(\lambda a z^{a-1} + a(a-1) z^{a-2}\right) \\
&= e^{\lambda(z-1)} z^{a-2}\left(\lambda^2 z^2 + 2\lambda a z + a(a-1)\right).
\end{aligned}
$$

Consequently, the variance can be expressed as

$$\begin{aligned} Var\left(N^w\right) &= \lim_{z\to-1} e^{\lambda(z-1)} z^{a-2}\left(\lambda^2 z^2 + 2\lambda a z + a\left(a-1\right)\right) \\ &+ E\left(N^w\right) - \left(E\left(N^w\right)\right)^2 \\ &= \lambda^2 + 2\lambda a + a\left(a-1\right) + a + \lambda - \left(a+\lambda\right)^2 \\ &= \lambda. \end{aligned}$$

□

Theorem 10.28. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w\left(n;\phi\right) = \varepsilon\frac{n!}{e^{-\lambda}\lambda^n} I\left(n=0\right) + \left(1-\varepsilon\right)$ *then*

$$\begin{aligned} E\left(w\left(N;\phi\right)\right) &= 1. \\ f_w\left(n\right) &= \left(\varepsilon\frac{n!}{e^{-\lambda}\lambda^n} I\left(n=0\right) + \left(1-\varepsilon\right)\right)\frac{e^{-\lambda}\lambda^n}{n!}. \\ g\left(z\right) &= \varepsilon + e^{\lambda(z-1)}\left(1-\varepsilon\right). \\ E\left(N^w\right) &= \lambda\left(1-\varepsilon\right). \\ Var\left(N^w\right) &= \lambda\left(1-\varepsilon\right)\left(1+\varepsilon\lambda\right). \end{aligned}$$

Proof. From the definition of the normalising constant, it follows that

$$\begin{aligned} E\left(w\left(N;\phi\right)\right) &= \sum_{k=0}^{\infty} w\left(k\right) f\left(k\right) \\ &= \sum_{k=0}^{\infty}\left(\varepsilon\frac{k!}{e^{-\lambda}\lambda^k} I\left(k=0\right) + \left(1-\varepsilon\right)\right)\frac{e^{-\lambda}\lambda^k}{k!} \\ &= \varepsilon\sum_{k=0}^{\infty} I\left(n=0\right) + \left(1-\varepsilon\right)\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!} \\ &= \varepsilon + \left(1-\varepsilon\right) \\ &= 1. \end{aligned}$$

From the definition of the weighted Poisson probability mass function, it follows that

$$\begin{aligned} f_w\left(n\right) &= \frac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\ &= \left(\varepsilon\frac{n!}{e^{-\lambda}\lambda^n} I\left(n=0\right) + \left(1-\varepsilon\right)\right)\frac{e^{-\lambda}\lambda^n}{n!}. \end{aligned}$$

From the definition of the probability generating function, it follows that

$$
\begin{aligned}
g(z) &= \textstyle\sum_{k=0}^{\infty} f_w(k)\, z^k \\
&= \textstyle\sum_{k=0}^{\infty} \left(\varepsilon \frac{k!}{e^{-\lambda}\lambda^k} I(k=0) + (1-\varepsilon)\right) \frac{e^{-\lambda}\lambda^k}{k!} z^k \\
&= \varepsilon \textstyle\sum_{k=0}^{\infty} \frac{k!}{e^{-\lambda}\lambda^k} I(k=0) \frac{e^{-\lambda}(\lambda z)^k}{k!} + (1-\varepsilon) \sum_{k=0}^{\infty} \frac{e^{-\lambda}(\lambda z)^k}{k!} \\
&= \varepsilon \textstyle\sum_{k=0}^{\infty} \frac{(\lambda z)^k}{\lambda^k} I(k=0) + (1-\varepsilon) e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda z)^k}{k!} \\
&= \varepsilon + e^{\lambda(z-1)}(1-\varepsilon).
\end{aligned}
$$

From the definition of the expected value, it follows that

$$
\begin{aligned}
E(N^w) &= \lim_{z\to^-1} \tfrac{\partial}{\partial z} g(z) \\
\tfrac{\partial}{\partial z} g(z) &= \tfrac{\partial}{\partial z}\left(\varepsilon + e^{\lambda(z-1)}(1-\varepsilon)\right) \\
&= \lambda e^{\lambda(z-1)}(1-\varepsilon).
\end{aligned}
$$

It then follows that

$$
\begin{aligned}
E(N^w) &= \lim_{z\to^-1} \lambda e^{\lambda(z-1)}(1-\varepsilon) \\
&= \lambda(1-\varepsilon).
\end{aligned}
$$

From the definition of the variance, it follows that

$$
Var(N^w) = \lim_{z\to^-1} \tfrac{\partial^2}{\partial z^2} g(z) + E(N^w) - (E(N^w))^2.
$$

However,

$$
\begin{aligned}
\tfrac{\partial^2}{\partial z^2} g(z) &= \tfrac{\partial}{\partial z}\left(\lambda e^{\lambda(z-1)}(1-\varepsilon)\right) \\
&= \lambda^2 e^{\lambda(z-1)}(1-\varepsilon).
\end{aligned}
$$

Consequently, the variance can be expressed as

$$
\begin{aligned}
Var(N^w) &= \lim_{z\to^-1} \lambda^2 e^{\lambda(z-1)}(1-\varepsilon) + E(N^w) - (E(N^w))^2 \\
&= \lambda^2(1-\varepsilon) + \lambda(1-\varepsilon) - (\lambda(1-\varepsilon))^2 \\
&= \lambda(1-\varepsilon)(1+\varepsilon\lambda).
\end{aligned}
$$

□

Theorem 10.29. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w(n;\phi) = (a)_n$ *then*

$$\begin{aligned} E(w(N;\phi)) &= e^{-\lambda}(1-\lambda)^{-a}. \\ f_w(n) &= \tfrac{(1-\lambda)^a \lambda^n (a)_n}{n!}. \\ g(z) &= (1-\lambda)^a (1-\lambda z)^{-a}. \\ E(N^w) &= \tfrac{a\lambda}{1-\lambda}. \\ Var(N^w) &= \tfrac{a\lambda}{(\lambda-1)^2}. \end{aligned}$$

Proof. From the definition of the normalising constant, it follows that

$$\begin{aligned} E(w(N;\phi)) &= \textstyle\sum_{k=0}^{\infty} w(k) f(k) \\ &= \textstyle\sum_{k=0}^{\infty} (a)_n \frac{e^{-\lambda}\lambda^k}{k!} \\ &= e^{-\lambda} \textstyle\sum_{k=0}^{\infty} \frac{\Gamma(a+k)}{\Gamma(a)} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} {}_1F_0(a;\lambda) \\ &= e^{-\lambda}(1-\lambda)^{-a}. \end{aligned}$$

From the definition of the weighted Poisson probability mass function, it follows that

$$\begin{aligned} f_w(n) &= \tfrac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\ &= \tfrac{(a)_n \frac{e^{-\lambda}\lambda^n}{n!}}{e^{-\lambda}(1-\lambda)^{-a}} \\ &= \tfrac{(1-\lambda)^a \lambda^n (a)_n}{n!}. \end{aligned}$$

From the definition of the probability generating function, it follows that

$$\begin{aligned} g(z) &= \textstyle\sum_{k=0}^{\infty} f_w(k) z^k \\ &= \textstyle\sum_{k=0}^{\infty} \frac{(1-\lambda)^a \lambda^k (a)_k}{k!} z^k \\ &= (1-\lambda)^a \textstyle\sum_{k=0}^{\infty} \frac{\Gamma(a+k)}{\Gamma(a)} \frac{(\lambda z)^k}{k!} \\ &= (1-\lambda)^a \, {}_1F_0(a;\lambda z) \\ &= (1-\lambda)^a (1-\lambda z)^{-a}. \end{aligned}$$

From the definition of the expected value, it follows that

$$\begin{aligned} E\left(N^{w}\right) &= \lim_{z\to -1}\tfrac{\partial}{\partial z}g\left(z\right) \\ \tfrac{\partial}{\partial z}g\left(z\right) &= \tfrac{\partial}{\partial z}\left(\left(1-\lambda\right)^{a}\left(1-\lambda z\right)^{-a}\right) \\ &= \left(1-\lambda\right)^{a}a\lambda\left(1-\lambda z\right)^{-a-1}. \end{aligned}$$

It then follows that

$$\begin{aligned} E\left(N^{w}\right) &= \lim_{z\to -1}\left(1-\lambda\right)^{a}a\lambda\left(1-\lambda z\right)^{-a-1} \\ &= \left(1-\lambda\right)^{a}a\lambda\left(1-\lambda\right)^{-a-1} \\ &= \tfrac{a\lambda}{1-\lambda}. \end{aligned}$$

From the definition of the variance, it follows that

$$Var\left(N^{w}\right) = \lim_{z\to -1}\tfrac{\partial^{2}}{\partial z^{2}}g\left(z\right)+E\left(N^{w}\right)-\left(E\left(N^{w}\right)\right)^{2}.$$

However,

$$\begin{aligned} \tfrac{\partial^{2}}{\partial z^{2}}g\left(z\right) &= \tfrac{\partial}{\partial z}\left(\left(1-\lambda\right)^{a}a\lambda\left(1-\lambda z\right)^{-a-1}\right) \\ &= \left(1-\lambda\right)^{a}a\left(1+a\right)\lambda^{2}\left(1-\lambda z\right)^{-a-2}. \end{aligned}$$

Consequently, the variance can be expressed as

$$\begin{aligned} Var\left(N^{w}\right) &= \lim_{z\to -1}\left(1-\lambda\right)^{a}a\left(1+a\right)\lambda^{2}\left(1-\lambda z\right)^{-a-2}+E\left(N^{w}\right)-\left(E\left(N^{w}\right)\right)^{2} \\ &= \left(1-\lambda\right)^{a}a\left(1+a\right)\lambda^{2}\left(1-z\right)^{-a-2}+\tfrac{a\lambda}{1-\lambda}-\left(\tfrac{a\lambda}{1-\lambda}\right)^{2} \\ &= \tfrac{a\lambda}{\left(\lambda-1\right)^{2}}. \end{aligned}$$

□

Theorem 10.30. *If the weight function used in the weighted Poisson probability mass function is chosen as* $w\left(n;\phi\right)=\left(n\right)_{a}$ *then*

$$\begin{aligned} E\left(w\left(N;\phi\right)\right) &= e^{-\lambda}\lambda\Gamma\left(1+a\right){}_{1}F_{1}\left(1+a;2;\lambda\right). \\ f_{w}\left(n\right) &= \tfrac{\lambda^{n-1}\left(n\right)_{a}}{n!\Gamma\left(1+a\right){}_{1}F_{1}\left(1+a;2;\lambda\right)}. \\ g\left(z\right) &= \tfrac{z\,{}_{1}F_{1}\left(1+a;2;\lambda z\right)}{{}_{1}F_{1}\left(1+a;2;\lambda\right)}. \\ E\left(N^{w}\right) &= \tfrac{2\,{}_{1}F_{1}\left(1+a;2;\lambda\right)+\left(1+a\right)\lambda\,{}_{1}F_{1}\left(2+a;3;\lambda\right)}{2\,{}_{1}F_{1}\left(1+a;2;\lambda\right)}. \\ Var\left(N^{w}\right) &= \tfrac{\left(1+a\right)\lambda\left(6\,{}_{1}F_{1}\left(2+a;3;\lambda\right)+\left(2+a\right)\lambda\,{}_{1}F_{1}\left(3+a;4;\lambda\right)\right)}{6\,{}_{1}F_{1}\left(1+a;2;\lambda\right)}+E\left(N^{w}\right)-\left(E\left(N^{w}\right)\right)^{2}. \end{aligned}$$

Proof. From the definition of the normalising constant, it follows that

$$\begin{aligned}
E\left(w\left(N;\phi\right)\right) &= \textstyle\sum_{k=0}^{\infty} w\left(k\right) f\left(k\right) \\
&= \textstyle\sum_{k=0}^{\infty} \left(k\right)_a \frac{e^{-\lambda}\lambda^k}{k!} \\
&= e^{-\lambda} \textstyle\sum_{k=0}^{\infty} \frac{\Gamma(a+k)}{\Gamma(k)} \frac{\lambda^k}{k!} \\
&= e^{-\lambda}\lambda \textstyle\sum_{k=0}^{\infty} \frac{\Gamma(a+k)k}{\Gamma(k+1)} \frac{\lambda^{k-1}}{k!} \\
&= e^{-\lambda}\lambda \textstyle\sum_{k=1}^{\infty} \frac{\Gamma(a+k)}{\Gamma(k+1)} \frac{\lambda^{k-1}}{(k-1)!}.
\end{aligned}$$

By reparameterising $m = k - 1$, it follows that

$$\begin{aligned}
E\left(w\left(N;\phi\right)\right) &= e^{-\lambda}\lambda \textstyle\sum_{m=0}^{\infty} \frac{\Gamma(a+m+1)}{\Gamma(m+2)} \frac{\lambda^m}{m!} \\
&= e^{-\lambda}\lambda\Gamma\left(1+a\right) \textstyle\sum_{m=0}^{\infty} \frac{\Gamma(a+m+1)}{\Gamma(a+1)} \frac{\Gamma(2)}{\Gamma(m+2)} \frac{\lambda^m}{m!} \\
&= e^{-\lambda}\lambda\Gamma\left(1+a\right) {}_1F_1\left(1+a;2;\lambda\right).
\end{aligned}$$

From the definition of the weighted Poisson probability mass function, it follows that

$$\begin{aligned}
f_w\left(n\right) &= \frac{w(n)f(n)}{\sum_{k=0}^{\infty} w(k)f(k)} \\
&= \frac{\frac{e^{-\lambda}\lambda^n}{n!}}{e^{-\lambda}\lambda\Gamma(1+a){}_1F_1(1+a;2;\lambda)} \\
&= \frac{\lambda^{n-1}(n)_a}{n!\Gamma(1+a){}_1F_1(1+a;2;\lambda)}.
\end{aligned}$$

From the definition of the probability generating function, it follows that

$$\begin{aligned}
g\left(z\right) &= \textstyle\sum_{k=0}^{\infty} f_w\left(k\right) z^k \\
&= \textstyle\sum_{k=0}^{\infty} \frac{\lambda^{k-1}(k)_a}{k!\Gamma(1+a){}_1F_1(1+a;2;\lambda)} z^k \\
&= \frac{z}{{}_1F_1(1+a;2;\lambda)} \textstyle\sum_{k=0}^{\infty} \frac{\Gamma(k+a)(\lambda z)^{k-1}}{\Gamma(k)\Gamma(1+a)k!} \\
&= \frac{z}{{}_1F_1(1+a;2;\lambda)} \textstyle\sum_{k=0}^{\infty} \frac{k\Gamma(k+a)(\lambda z)^{k-1}}{\Gamma(k+1)\Gamma(1+a)k!} \\
&= \frac{z}{{}_1F_1(1+a;2;\lambda)} \textstyle\sum_{k=0}^{\infty} \frac{\Gamma(k+a)(\lambda z)^{k-1}}{\Gamma(k+1)\Gamma(1+a)(k-1)!}.
\end{aligned}$$

By reparameterising $m = k - 1$, it follows that

$$
\begin{aligned}
g(z) &= \frac{z}{{}_1F_1(1+a;2;\lambda)} \sum_{m=0}^{\infty} \frac{\Gamma(m+a+1)}{\Gamma(m+2)\Gamma(1+a)} \frac{(z\lambda)^m}{m!} \\
&= \frac{z}{{}_1F_1(1+a;2;\lambda)} \sum_{m=0}^{\infty} \frac{\Gamma(m+a+1)\Gamma(2)}{\Gamma(a+1)\Gamma(m+2)} \frac{(\lambda z)^m}{m!} \\
&= \frac{z\,{}_1F_1(1+a;2;\lambda z)}{{}_1F_1(1+a;2;\lambda)}.
\end{aligned}
$$

From the definition of the expected value, it follows that

$$
E(N^w) = \lim_{z\to -1} \frac{\partial}{\partial z} g(z).
$$

However,

$$
\frac{\partial}{\partial z} g(z) = \frac{\partial}{\partial z} \frac{z\,{}_1F_1(1+a;2;\lambda z)}{{}_1F_1(1+a;2;\lambda)}.
$$

By using the result in Theorem 10.6, it follows that

$$
\begin{aligned}
\frac{\partial}{\partial z} g(z) &= \frac{{}_1F_1(1+a;2;\lambda z)}{{}_1F_1(1+a;2;\lambda)} + \frac{\frac{1+a}{2}\,{}_1F_1(2+a;3;\lambda z)}{{}_1F_1(1+a;2;\lambda)} \\
&= \frac{2\,{}_1F_1(1+a;2;\lambda z) + (1+a)\lambda z\,{}_1F_1(2+a;3;\lambda z)}{2\,{}_1F_1(1+a;2;\lambda)}.
\end{aligned}
$$

It then follows that

$$
\begin{aligned}
E(N^w) &= \lim_{z\to -1} \frac{2\,{}_1F_1(1+a;2;\lambda z) + (1+a)\lambda z\,{}_1F_1(2+a;3;\lambda z)}{2\,{}_1F_1(1+a;2;\lambda)} \\
&= \frac{2\,{}_1F_1(1+a;2;\lambda) + (1+a)\lambda\,{}_1F_1(2+a;3;\lambda)}{2\,{}_1F_1(1+a;2;\lambda)}.
\end{aligned}
$$

From the definition of the variance, it follows that

$$
Var(N^w) = \lim_{z\to -1} \frac{\partial^2}{\partial z^2} g(z) + E(N^w) - (E(N^w))^2.
$$

However,

$$
\begin{aligned}
\frac{\partial^2}{\partial z^2} g(z) &= \frac{\partial}{\partial z} \frac{2\,{}_1F_1(1+a;2;\lambda z) + (1+a)\lambda z\,{}_1F_1(2+a;3;\lambda z)}{2\,{}_1F_1(1+a;2;\lambda)} \\
&= \frac{(1+a)\lambda(6\,{}_1F_1(2+a;3;\lambda z) + (2+a)\lambda z\,{}_1F_1(3+a;4;\lambda z))}{6\,{}_1F_1(1+a;2;\lambda)}.
\end{aligned}
$$

Consequently, the variance can be expressed as

$$
\begin{aligned}
Var(N^w) &= \lim_{z\to -1} \frac{(1+a)\lambda(6\,{}_1F_1(2+a;3;\lambda z) + (2+a)\lambda z\,{}_1F_1(3+a;4;\lambda z))}{6\,{}_1F_1(1+a;2;\lambda)} + E(N^w) - (E(N^w))^2 \\
&= \frac{(1+a)\lambda(6\,{}_1F_1(2+a;3;\lambda) + (2+a)\lambda\,{}_1F_1(3+a;4;\lambda))}{6\,{}_1F_1(1+a;2;\lambda)} + E(N^w) - (E(N^w))^2.
\end{aligned}
$$

□

10.3 Additional weighted Poisson distribution fits

10.3.1 Weekly sales figures

For item 409, of the 28 weight functions that were tested, 9 models perform better than the Poisson, and 1 additional model may (the AIC and AICc disagree with the BIC). The best fit is achieved when $w(n;\phi) = (19.7293)_n$ and $\lambda = 0.683935$. This is shown in Figure 10.1. Although the weighted Poisson is able to capture the behaviour in tails of the distribution more accurately than the Poisson, the model still seems to be unable to model the observed distribution with a high degree of accuracy.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	$(0.56222; 0.805649)$	$(0.493803; 0.766735)$	$(0.637741; 0.720464)$
a	$(8.71696; 30.7417)$	$(13.3244; 42.8965)$	$(16.6434; 24.1238)$

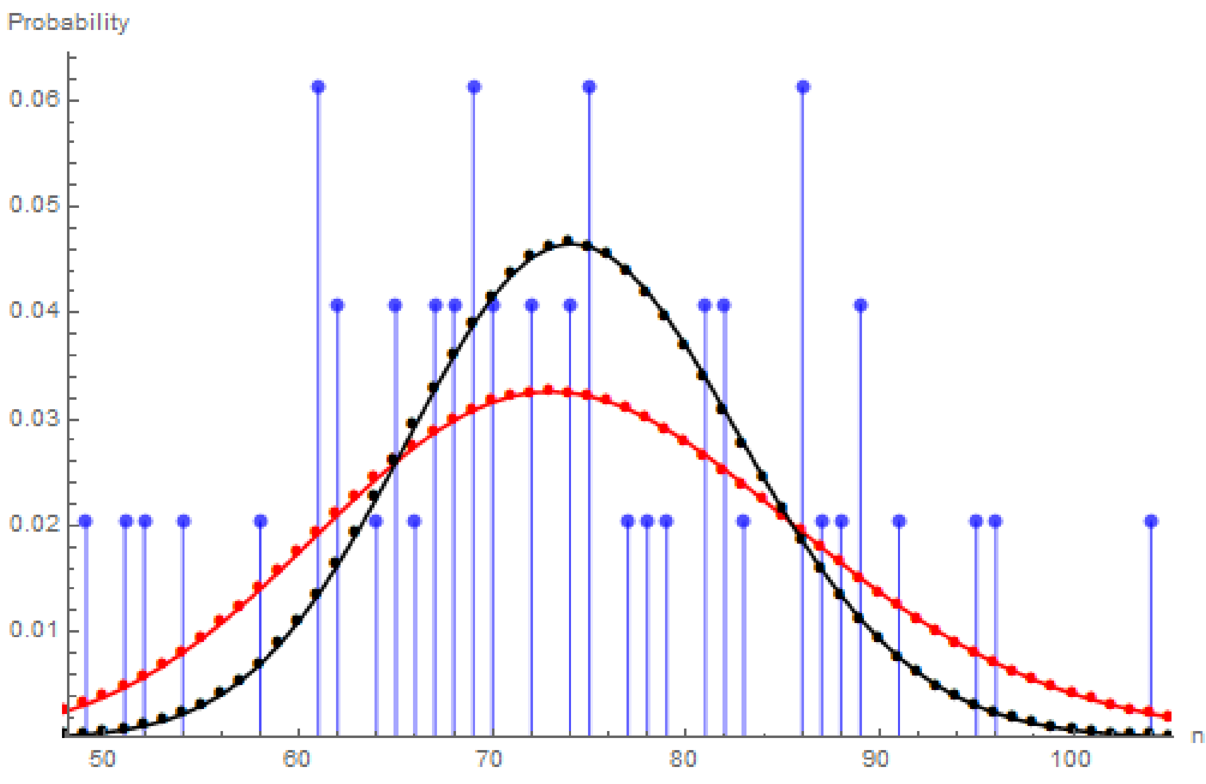


Figure 10.1: Weekly sales - Item 409

For item 726, of the 14 weight functions that were tested, only 1 model may perform better than the Poisson (the AIC and AICc disagree with the BIC). This fit is achieved when $w(n;\phi) = n + 0.400104$ and $\lambda = 0.540675$, and is shown in Figure 10.2. Although the AIC and AICc disagree with the BIC, the weighted Poisson fit appears to be better since it is much closer to the observed probabilities in four out of five observed frequencies, and only slightly further away at one. In general, it appears as if the model gives a better fit at the mean as well as in the tails.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	$(0.240878; 0.830472)$	$(0.098893; 1.01923)$	$(0.453359; 0.638422)$
ε	$(0; 0.964982)$	$(0.116912; 7.59677)$	$(0.263534; 0.661297)$

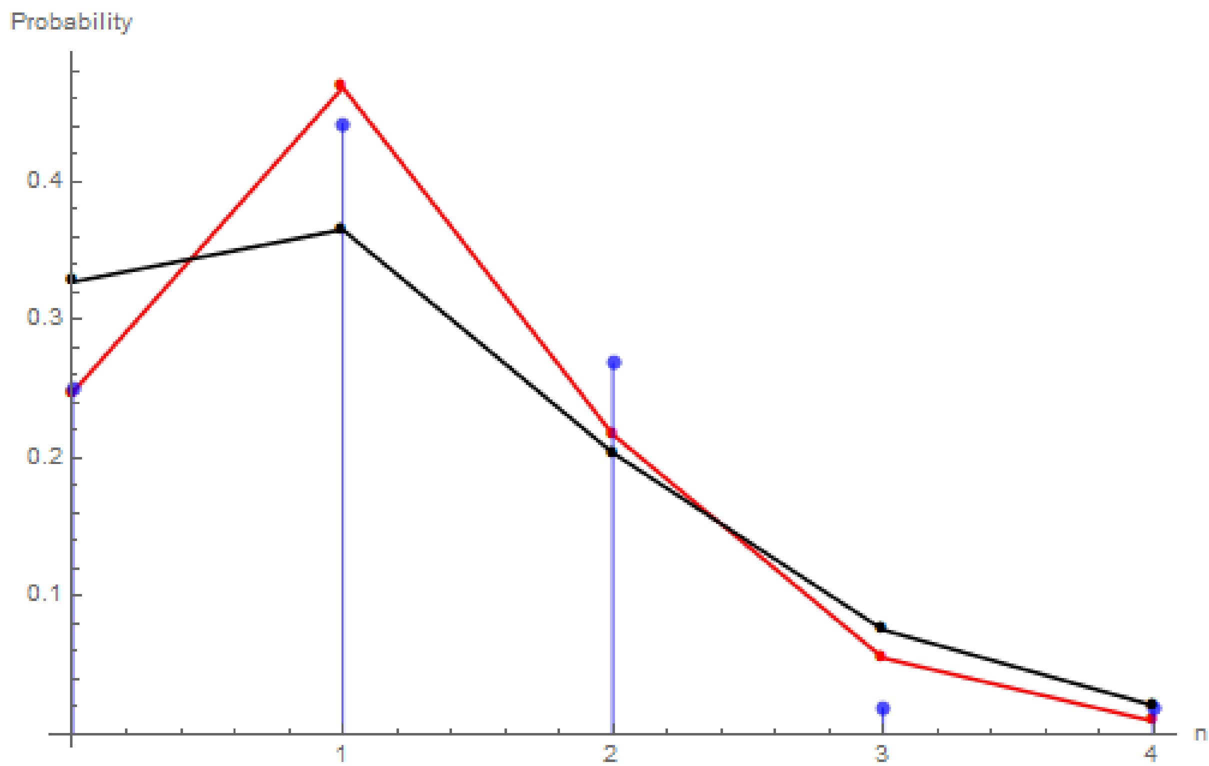


Figure 10.2: Weekly sales - Item 726

10.3.2 Airplane accidents

For the fatalities per incident, of the 12 weight functions that were tested, 6 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = \frac{\Gamma(0.822916+n)}{n!\Gamma(0.822916)} \frac{Beta(252204.8229,0.822409+n)}{Beta(252204.8229,0.822409)}$ and $\lambda = 245318$. This is shown in Figure 10.3. Clearly the weighted Poisson gives a better fit to the data relative to the Poisson.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	$(88399.3; 402237)$	$(2065.29; 323419)$	$(311.51; 104585)$
a	$(90830.6; 413578)$	$(2048.53; 332785)$	$(270.076; 107810)$
b	$(0.756138; 0.888679)$	$(0.742345; 0.842701)$	$(0.131846; 1.31792)$
r	$(0.756625; 0.889206)$	$(0.698649; 0.837322)$	$(0.35423; 1.54961)$

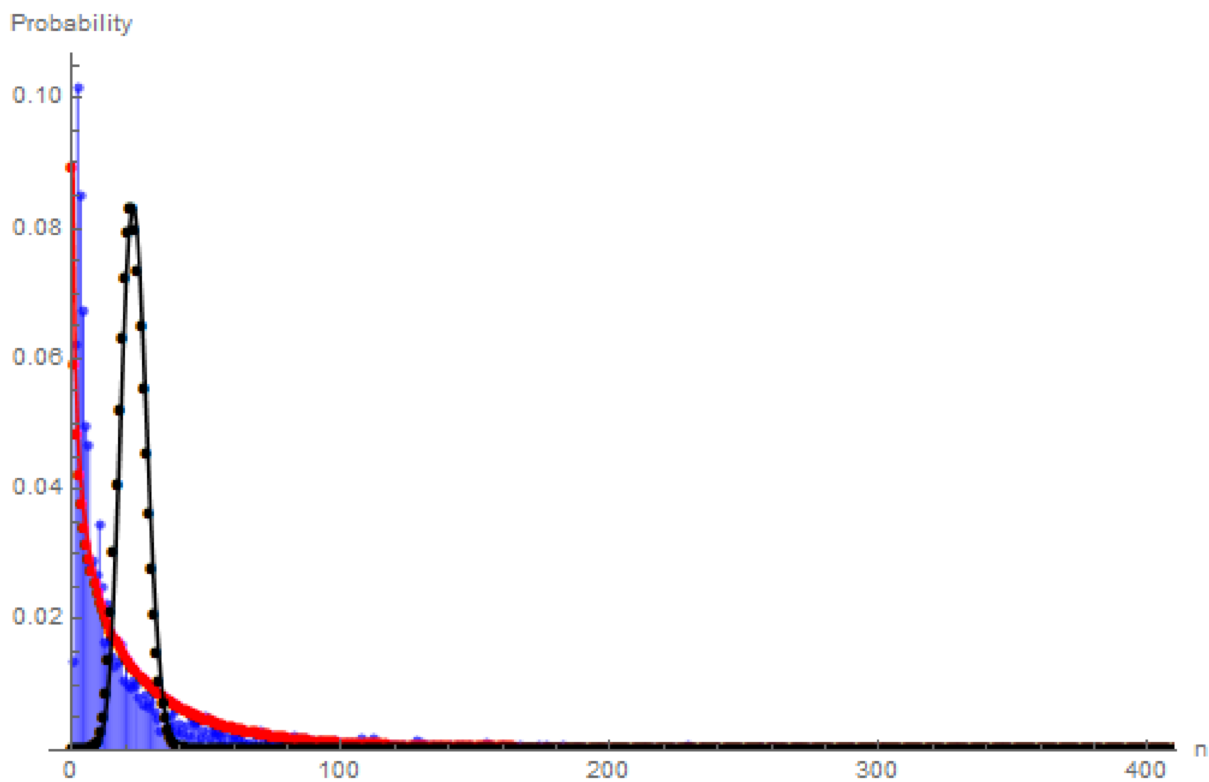


Figure 10.3: Airplane accidents - Fatalities per incident

Due to the long tail of the data, seeing the details of the fit may not be clear. Thus Figure 10.4 shows a cropped view of the plot.

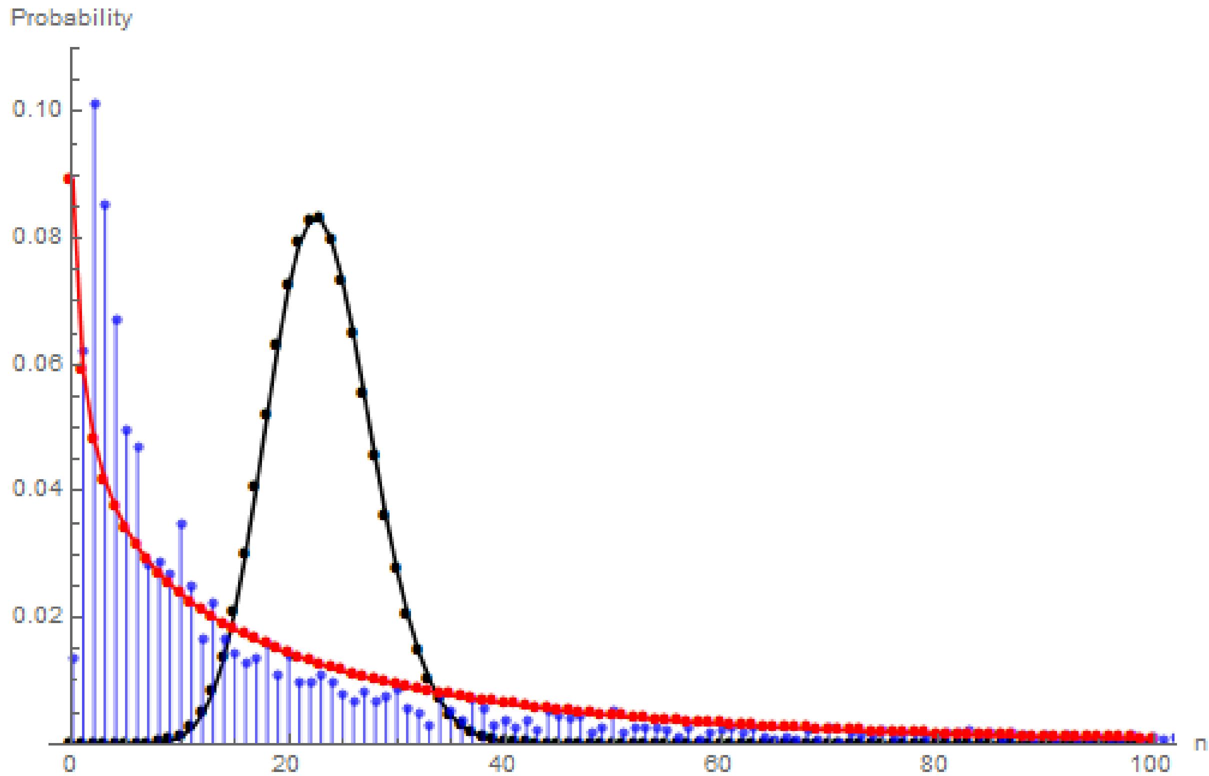


Figure 10.4: Airplane accidents - Cropped fatalities per incident

For the number of incidents per month, of the 14 weight functions that were tested, 3 models perform better than the Poisson, and 3 additional models may (the AIC and AICc disagree with the BIC). The best fit is achieved when $w(n;\phi) = (34.9643)_n$ and $\lambda = 0.148743$. This is shown in Figure 10.5. For this plot, the weighted Poisson appears to give a better fit close to the mean as well as in the parts of the tail of the distribution.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	(0.0504211; 0.247065)	(0.0315781; 0.237445)	(0.0361087; 0.240409)
a	(7.8418; 62.0868)	(19.0901; 136.867)	(19.3759; 163.146)

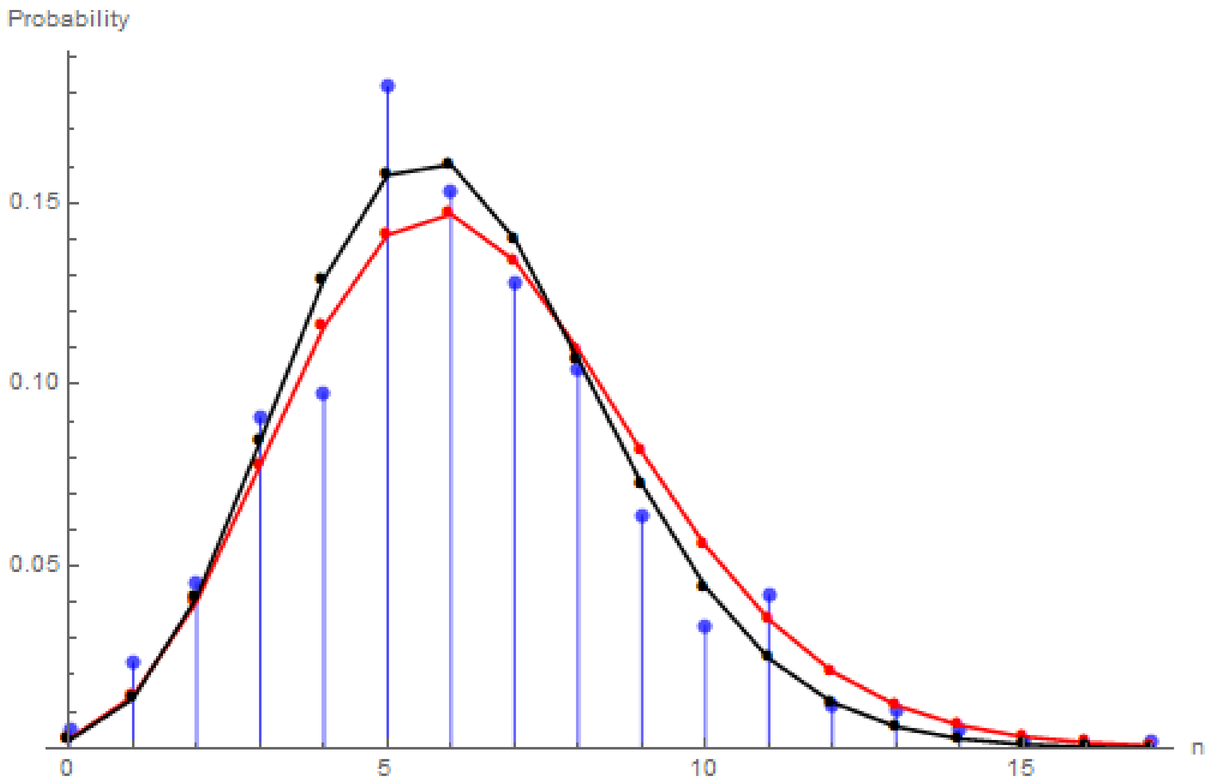


Figure 10.5: Airplane accidents - Incidents per month

10.3.3 USA mass shootings

For the number of people injured per incident, of the 12 weight functions that were tested, 5 models perform better than the Poisson, and 1 additional model may (the AIC and AICc disagree with the BIC). The best fit is achieved when $w(n;\phi) = \frac{\Gamma(2.37867+n)}{n!\Gamma(2.37867)} \frac{Beta(19156.67867,2.35224+n)}{Beta(19154.3,2.35224)}$ and $\lambda = 9526.54$. This is shown in Figure 10.6. Visually it appears as if the Poisson distribution gives a better fit for this dataset even though based on the AIC, AICc and BIC this is not the case. In fact, the weighted Poisson's AIC is roughly 25% smaller than that of the Poisson. The reason for this seeming contradiction is the influence of outliers. There have been relatively few shootings where more than twenty people have been injured, but these occurrences are enough to ensure that the thicker tail of the weighted Poisson distribution gives a better overall fit.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	$(0; 19376.5)$	$(2.55522; 160563)$	$(6.85417; 4557.64)$
a	$(0; 40470.6)$	$(0.001; 194802)$	$(2.30547; 8467.86)$
b	$(1.83189; 2.87258)$	$(1.05666; 428.689)$	$(0.734741; 23.2924)$
r	$(1.84965; 2.9077)$	$(0.40318; 339.651)$	$(0.385286; 5.22879)$

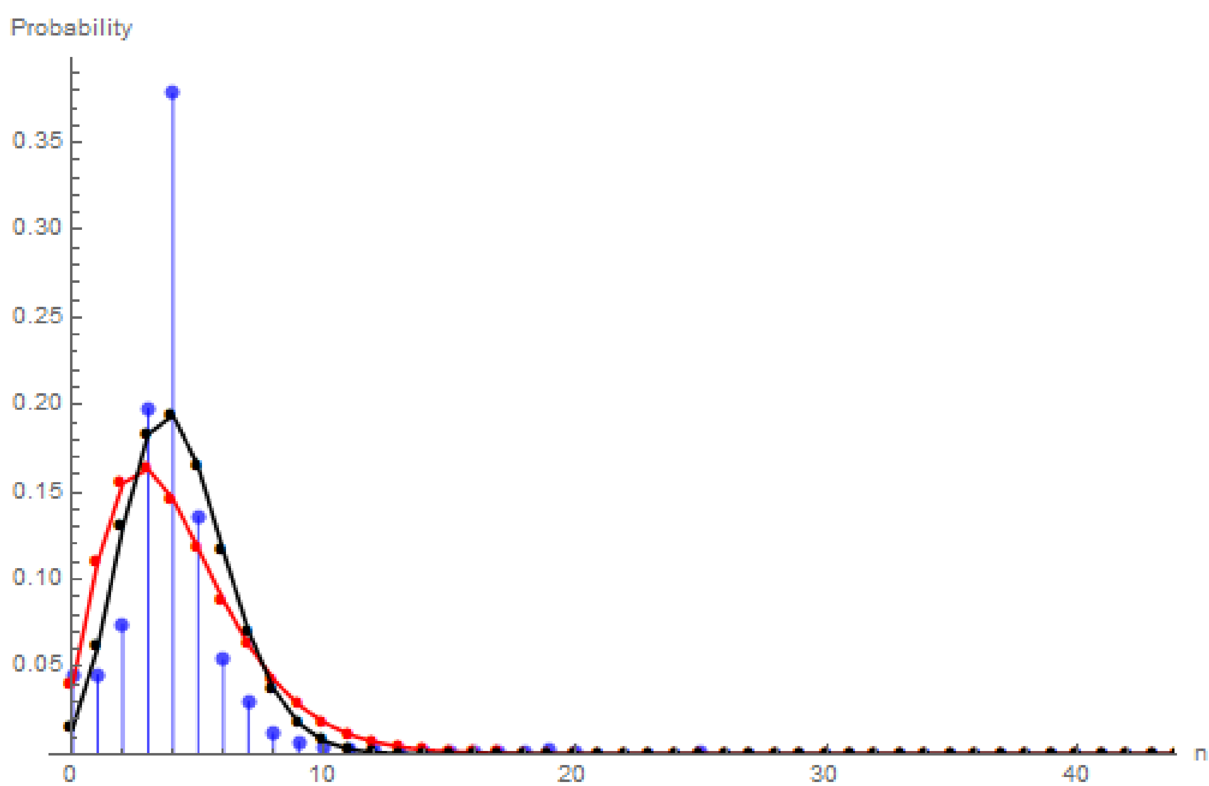


Figure 10.6: Mass shootings - Injuries per incident

For the number of people injured per day, of the 12 weight functions that were tested, 7 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = \frac{\Gamma(0.0470452+n)}{n!\Gamma(0.0470452)} \frac{Beta(5728.197,3.10309+n)}{Beta(5728.15,3.10309)}$ and $\lambda = 4565.16$. This is shown in Figure 10.7. In this case, the weighted Poisson distribution clearly gives a better fit to the data, especially in the thicker tails of the distribution.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	(2917.62; 6212.7)	(115.975; 65163.3)	(22.9515; 132601)
a	(3656.65; 7799.64)	(100.18; 68724.8)	(11.8711; 16485.5)
b	(2.74708; 3.45909)	(0.01; 0.025)	(0.0405052; 4.14726)
r	(0.036617; 0.0574734)	(0.01; 0.0366)	(0.0237813; 3.02111)

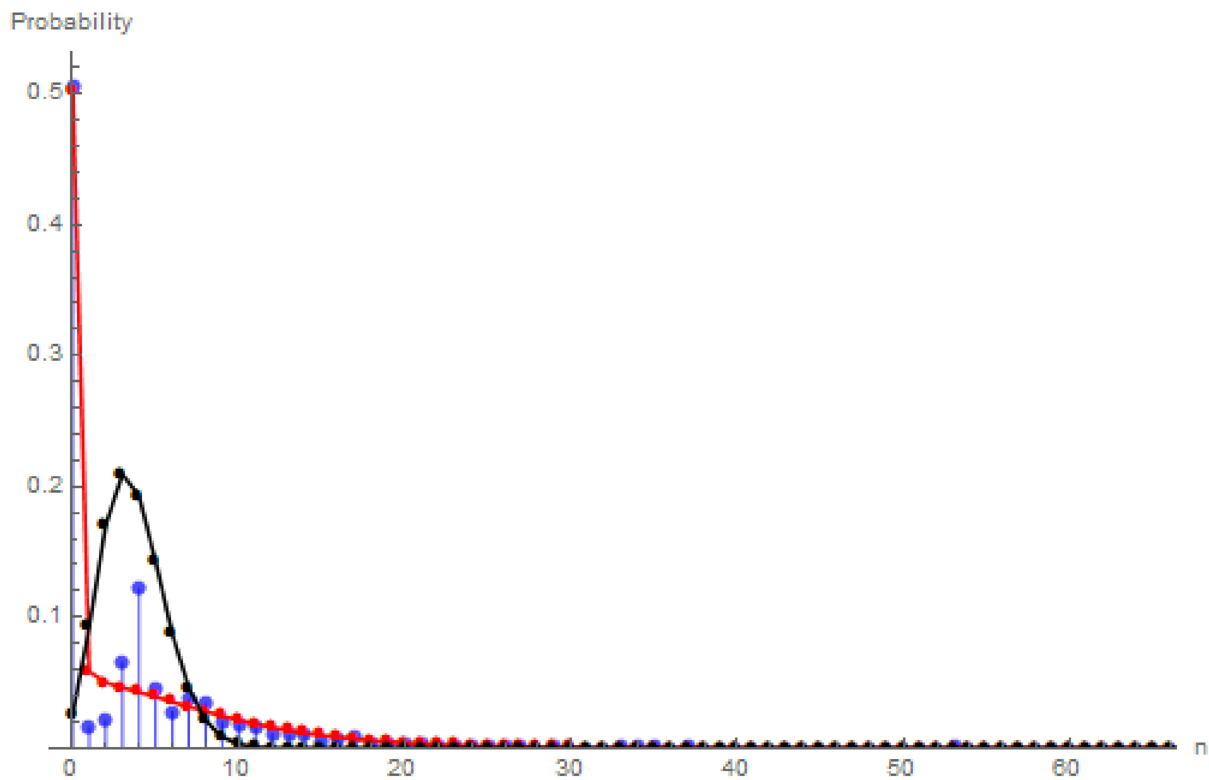


Figure 10.7: Mass shootings - Injuries per day

For the number of people injured per month, of the 23 weight functions that were tested, 10 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = (4.7508)_n$ and $\lambda = 0.95932$. This is shown in Figure 10.8. In this case, the mean of the data is 112.033 and the variance is 4610.74. The expected value and variance of the Poisson distribution is 112.033. For the weighted Poisson distribution the expected value is 112.034 and the variance is 2754.03. Clearly the weighted Poisson distribution gives a better fit to the data, especially in the thicker tails of the distribution. However, neither distribution can be said to give a good fit.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	$(0.944505; 0.974134)$	$(0.910151; 0.978254)$	$(0.953218; 0.964407)$
a	$(3.03747; 6.46417)$	$(2.8018; 10.3215)$	$(4.17055; 5.41539)$

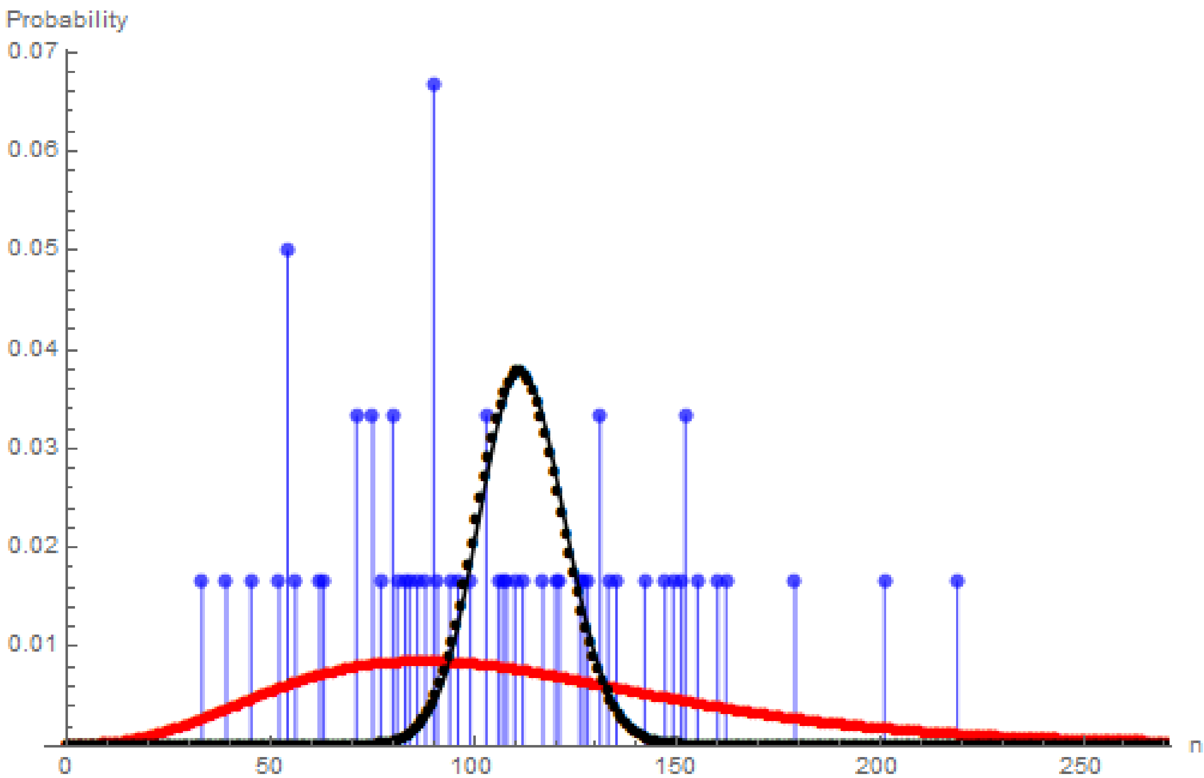


Figure 10.8: Mass shootings - Injuries per month

For the number of fatalities per incident, of the 14 weight functions that were tested, 8 models perform better than the Poisson. The best fit is achieved when $w\,(n;\phi) = (0.778335)_n$ and $\lambda = 0.594393$. This is shown in Figure 10.9. The weighted Poisson distribution gives a substantially better fit than the Poisson.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	(0.556888; 0.631899)	(0.492291; 0.673708)	(0.519699; 0.657046)
a	(0.671146; 0.885523)	(0.599329; 1.0886)	(0.608221; 1.00297)

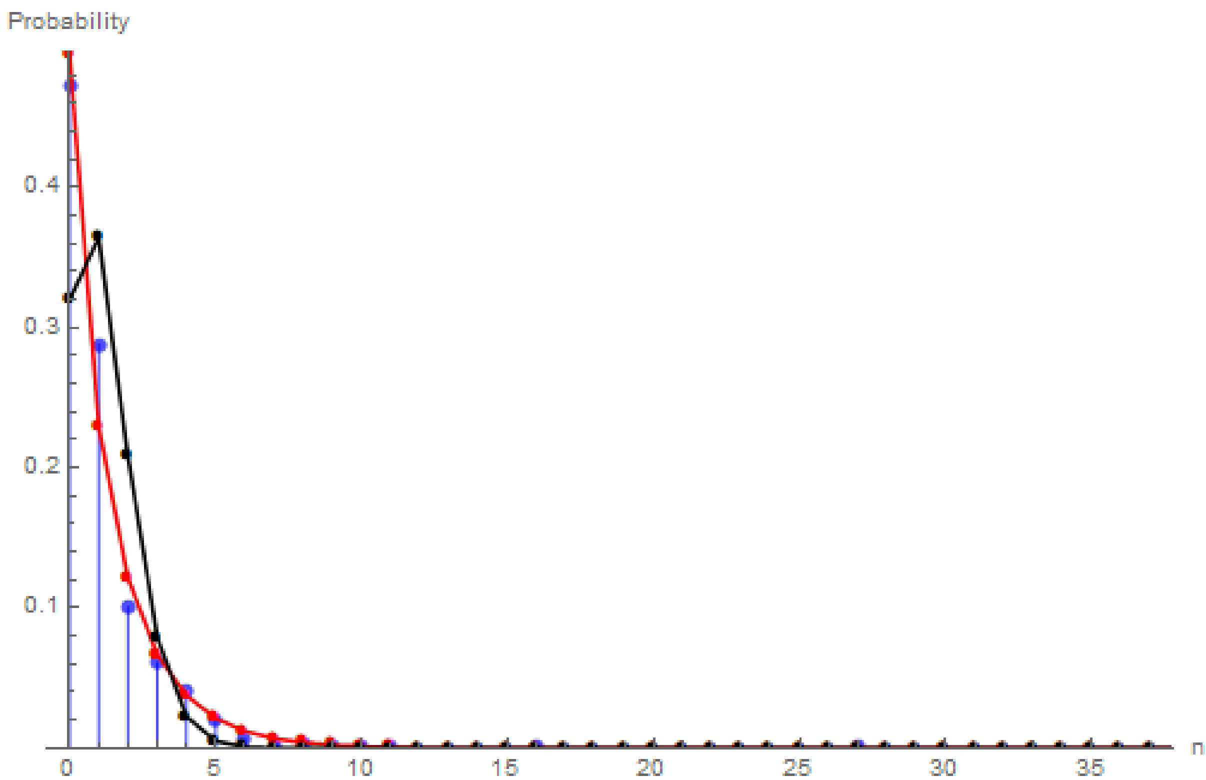


Figure 10.9: Mass shootings - Fatalities per incident

For the number of fatalities per day, of the 14 weight functions that were tested, 7 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = (0.291193)_n$ and $\lambda = 0.77284$. This is shown in Figure 10.10. In this case, the weighted Poisson distribution clearly gives a better fit to the data.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	$(0.745399; 0.800281)$	$(0.727203; 0.810189)$	$(0.7133311; 0.819651)$
a	$(0.25543; 0.326955)$	$(0.250118; 0.345572)$	$(0.2288; 0.37079)$

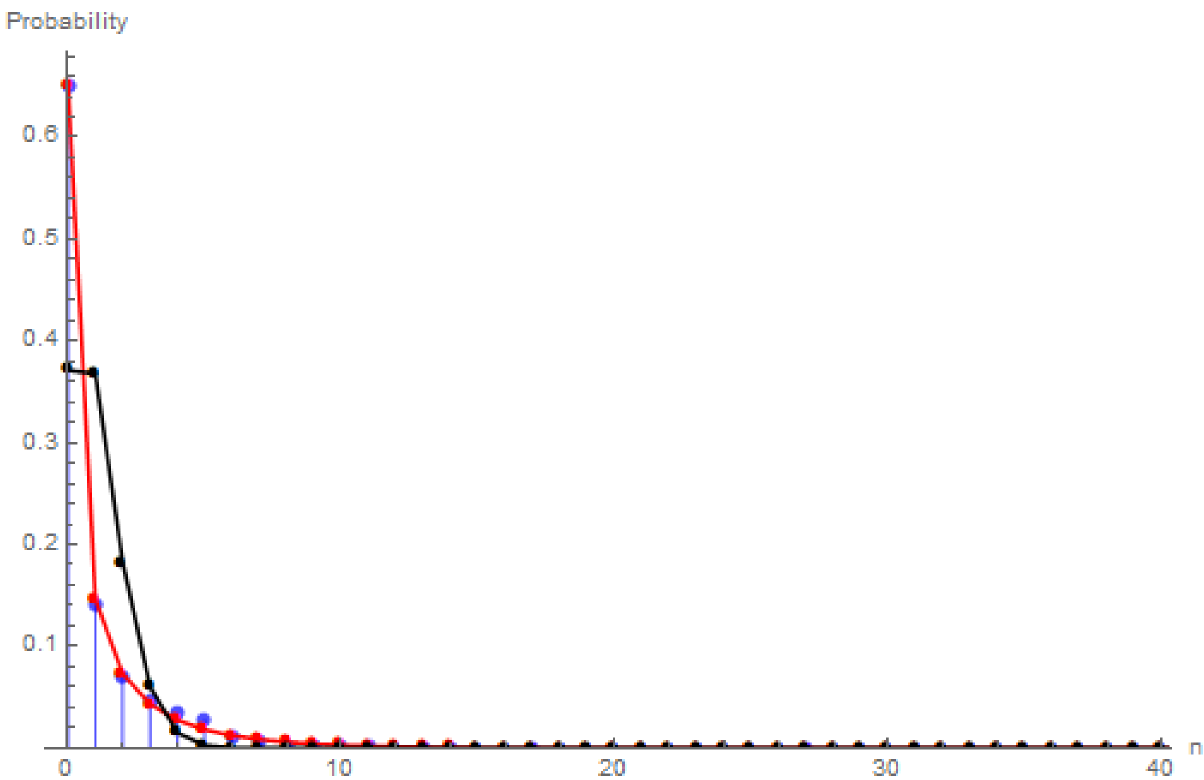


Figure 10.10: Mass shootings - Fatalities per day

For the number of fatalities per month, of the 26 weight functions that were tested, 9 models perform better than the Poisson, and 3 additional models may (the AIC and AICc disagree with the BIC). The best fit is achieved when $w(n;\phi) = (7.36203)_n$ and $\lambda = 0.803742$. This is shown in Figure 10.11. In this case, the weighted Poisson distribution gives a better fit to the data, although the model is still far from ideal.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	(0.73372; 0.873764)	(0.609418; 0.878808)	(0.774038; 0.826452)
a	(4.18496; 10.5391)	(4.43696; 18.0328)	(6.30695; 8.68692)

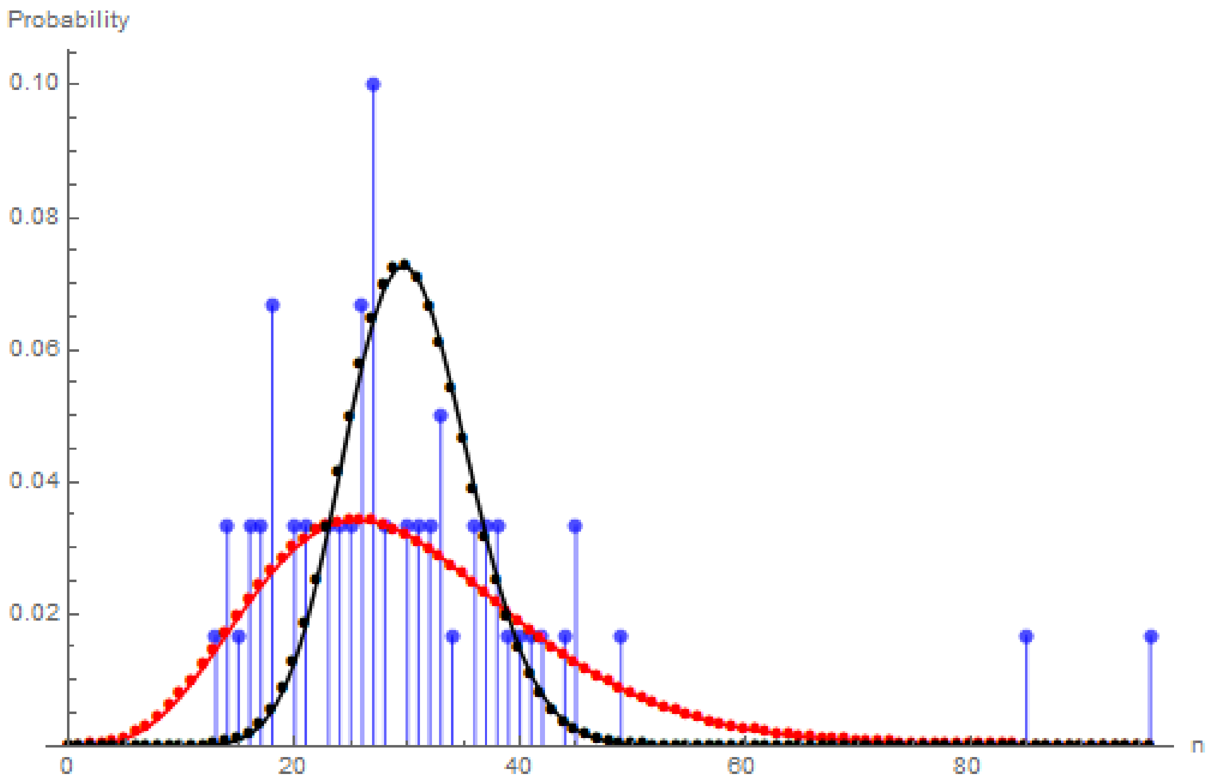


Figure 10.11: Mass shootings - Fatalities per month

For the number of incidents per month, of the 14 weight functions that were tested, 3 models perform better than the Poisson, and 3 additional models may (the AIC and AICc disagree with the BIC). The best fit is achieved when $w(n;\phi) = (14.3382)_n$ and $\lambda = 0.648321$. This is shown in Figure 10.12. In this case, the weighted Poisson distribution gives a better fit to the data, although the model is far from perfect.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	(0.520868; 0.775787)	(0.50007; 0.724876)	(0.598068; 0.688329)
a	(6.41148; 22.265)	(9.98108; 26.0307)	(11.934; 17.6548)

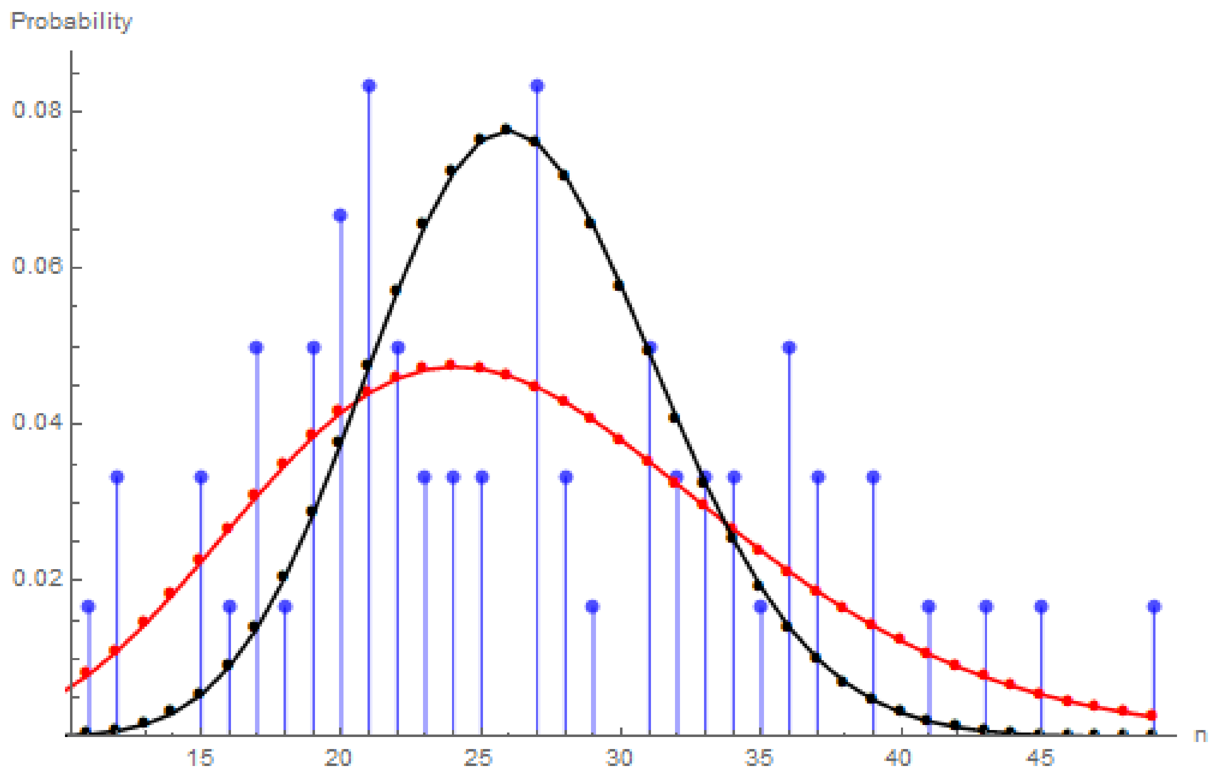


Figure 10.12: Mass shootings - Incidents per month

In all three of the above "per month" analyses, the fits of the distributions are not ideal. This could be due to the relative sparsity in the number of months in the dataset.

10.3.4 Vehicle accidents in Great Britain

For the number of casualties per hour, of the 10 weight functions that were tested, 5 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = (0.998156)_n$ and $\lambda = 0.981357$. This is shown in Figure 10.13. In this case, the weighted Poisson distribution gives a good fit to the data.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	(0.980712; 0.982001)	(0.98051; 0.982225)	(0.97816; 0.983616)
a	(0.970047; 1.02627)	(0.957641; 1.03497)	(0.889065; 1.12995)

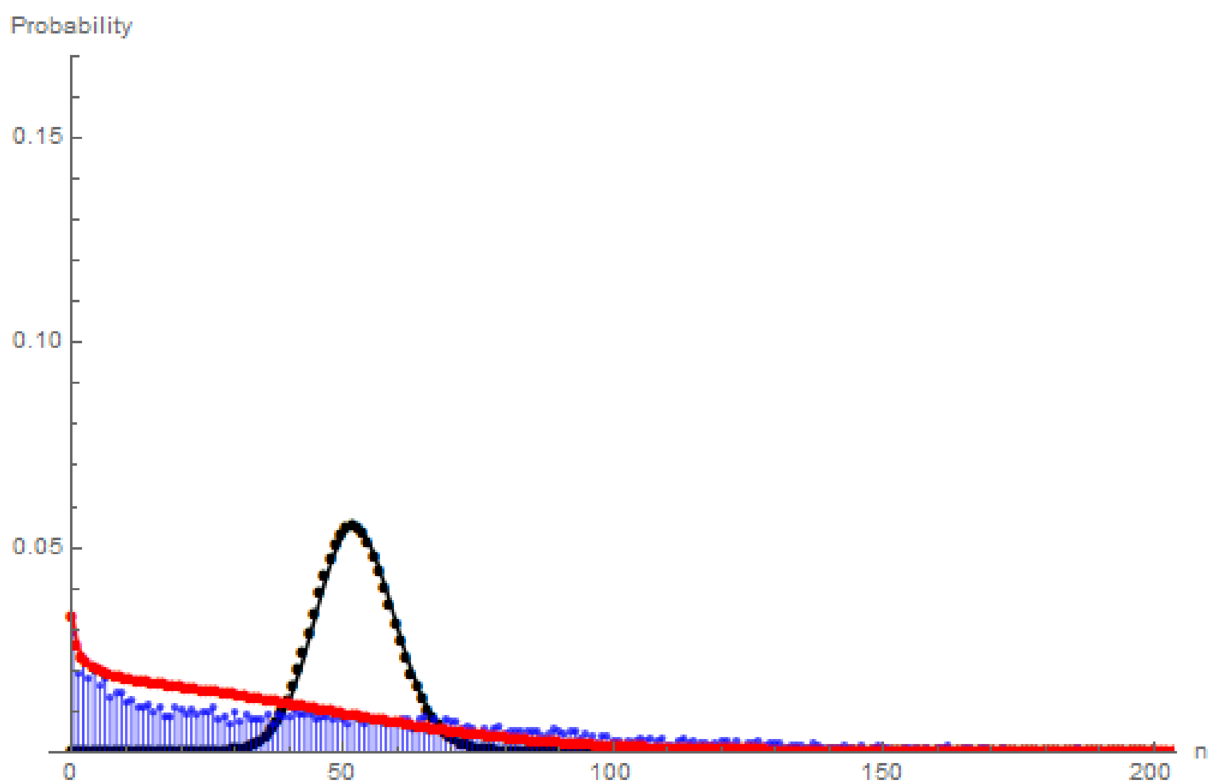


Figure 10.13: Britain accidents - Casualties per hour

For the number of vehicles per hour, of the 13 weight functions that were tested, 7 models perform better than the Poisson. The best occurs when $w(n;\phi) = \left(\frac{\Gamma(26974+n)}{n!\Gamma(26974)} \frac{Beta(27820,73964+n)}{Beta(846,73964)}\right)^{-1}$ and $\lambda = 19416$. This is shown in Figure 10.14. In this case, the weighted Poisson distribution gives a good fit to the data.

The weight function that results in the best fit contains quite a few gamma functions with extremely large parameters. This results in extremely slow calculations (even after optimising code and parallelising processing). As a result, calculating accurate confidence intervals, while theoretically possible, quickly becomes intractable. For this reason the confidence intervals for this specific application have been excluded.

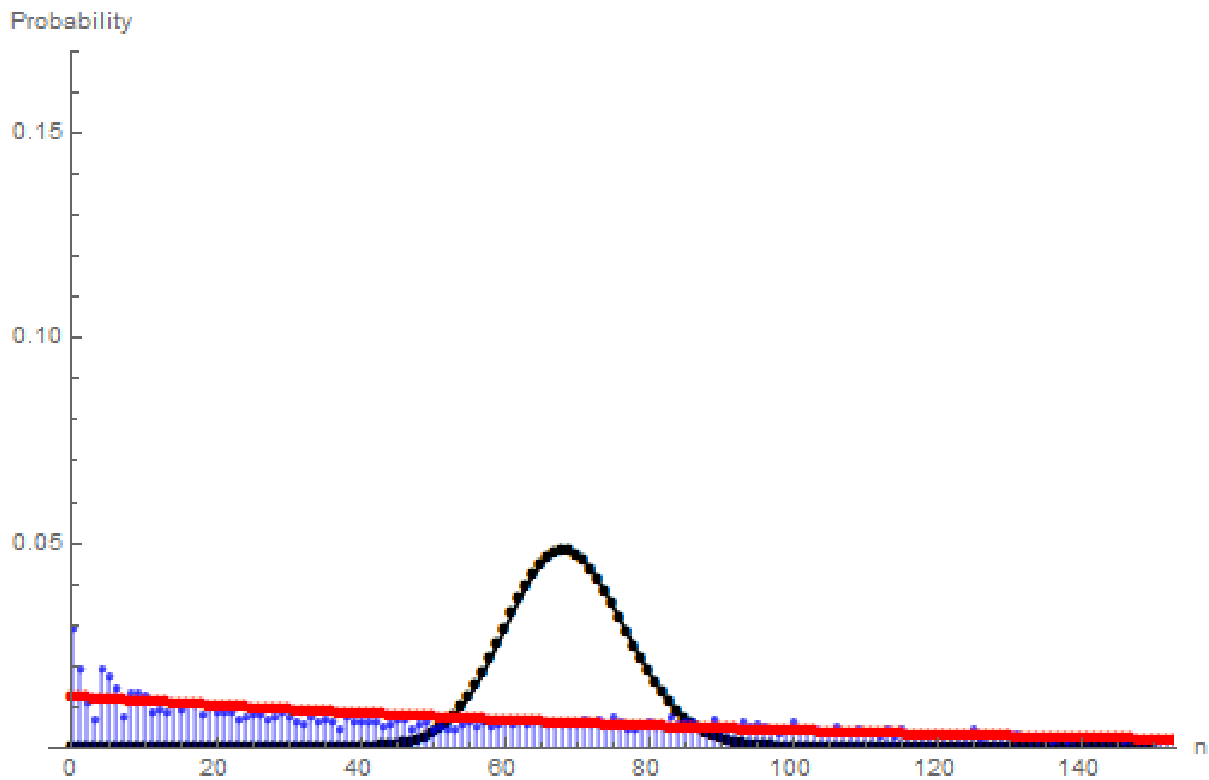


Figure 10.14: Britain accidents - Vehicles per hour

For the number of casualties per incident, of the 28 weight functions that were tested, 24 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = a \times \frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$, $a = 23606565.3873$ and $\lambda = 13919592.9622$. This is shown in Figure 10.15. In this case, the weighted Poisson distribution gives a very good fit to the data. The fit is near-perfect.

The confidence intervals for the weighted Poisson parameters are as follows:

	Non-parametric bootstrap	Parametric bootstrap
λ	(7575030; 36708075)	(3.48211; 16295772)
a	(12924999; 62246590)	(3.21385; 28046485)

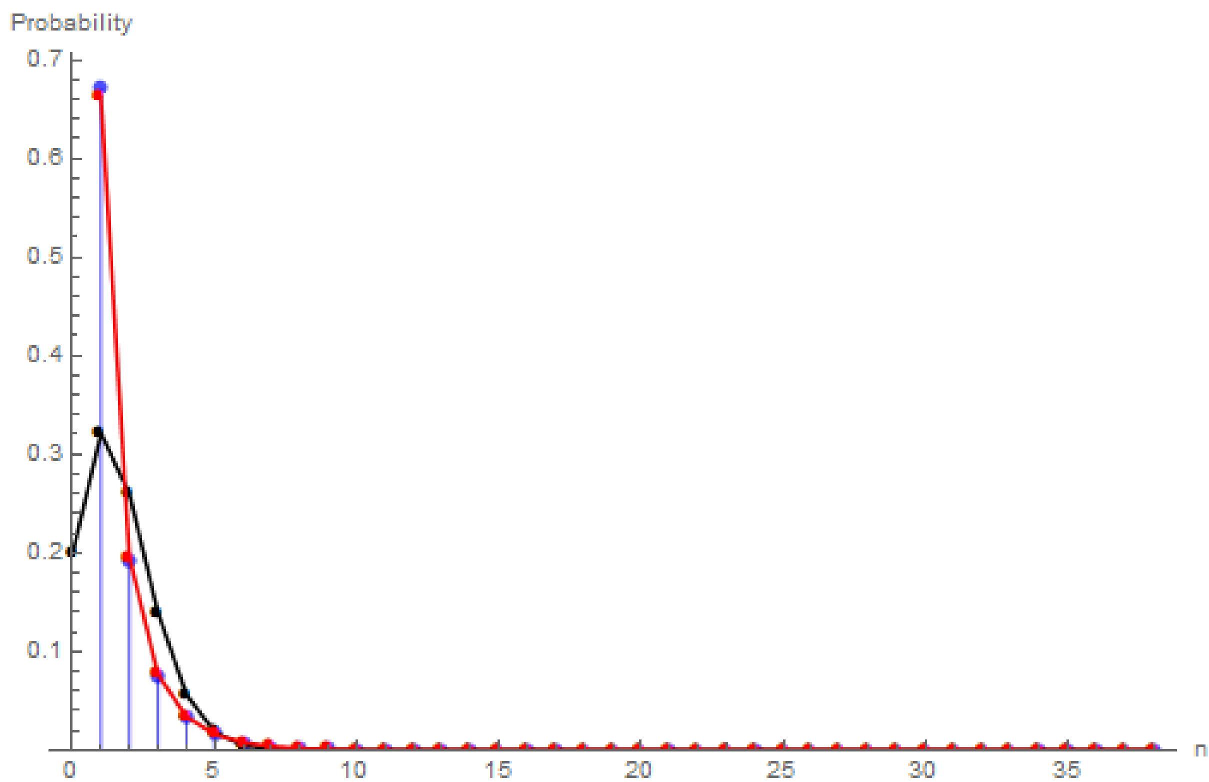


Figure 10.15: Britain accidents - Casualties per incident

10.3.5 Vehicle accidents in Canada

For the number of vehicles per incident, of the 27 weight functions that were tested, 23 models perform better than the Poisson. The best fit occurs if $w(n;\phi) = (n + 0.001)(n - 0.324713)^2$ and $\lambda = 0.340524$. This is shown in Figure 10.16. In this case, the weighted Poisson distribution gives a very good fit to the data.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	(0.339489; 0.341559)	(0.183299; 0.810372)	(0.270302; 0.420862)
a	(0; 0.00000001)	(0.00000001; 0.000000015)	(0.00000001; 0.000000018)
b	(0.322944; 0.326479)	(0.0000001; 0.66193)	(0.0349641; 0.500586)

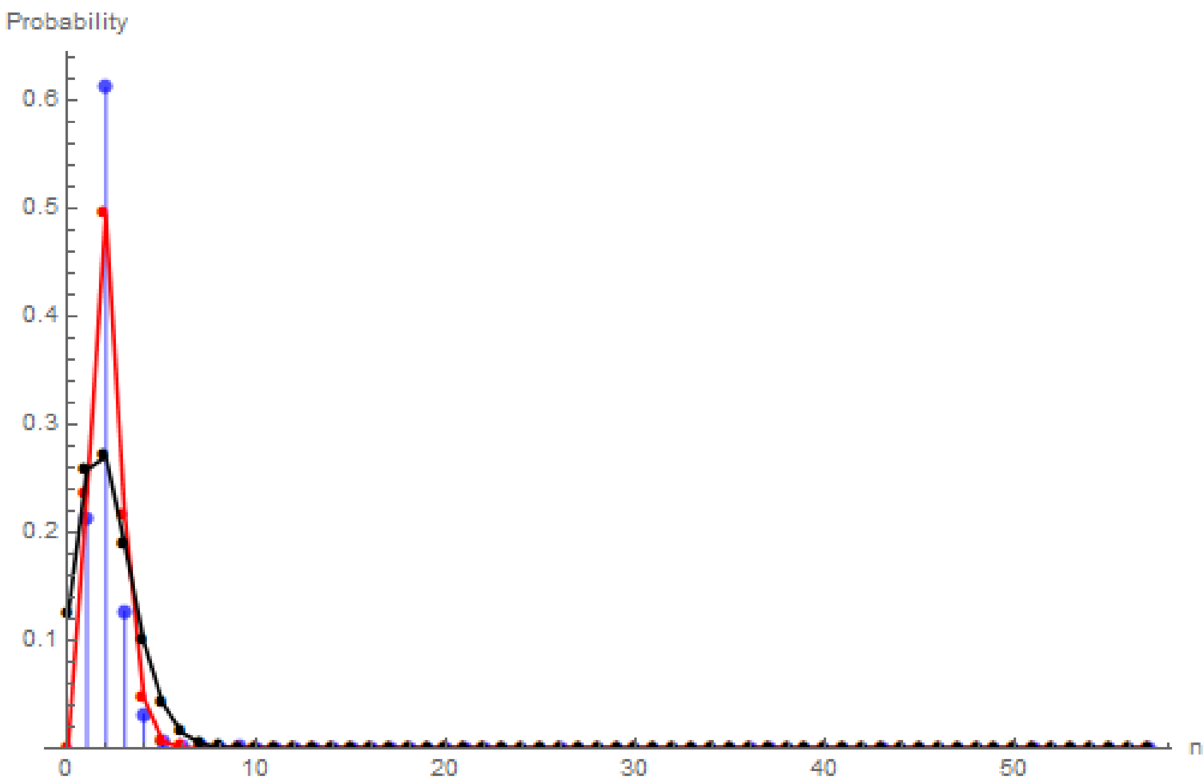


Figure 10.16: Canada accidents - Vehicles per incident

For the number of incidents per hour, of the 26 weight functions that were tested, 10 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = (1.43859)_n$ and $\lambda = 0.990283$. This is shown in Figure 10.17. In this case, the weighted Poisson distribution gives a decent fit to the data. To better visualise the shape of the weighted Poisson distribution a cropped graph is provided in Figure 10.18.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	$(0.98963; 0.990935)$	$(0.989165; 0.991296)$	$(0.988855; 0.991453)$
a	$(1.35645; 1.52074)$	$(1.31013; 1.60878)$	$(1.2946; 1.62693)$

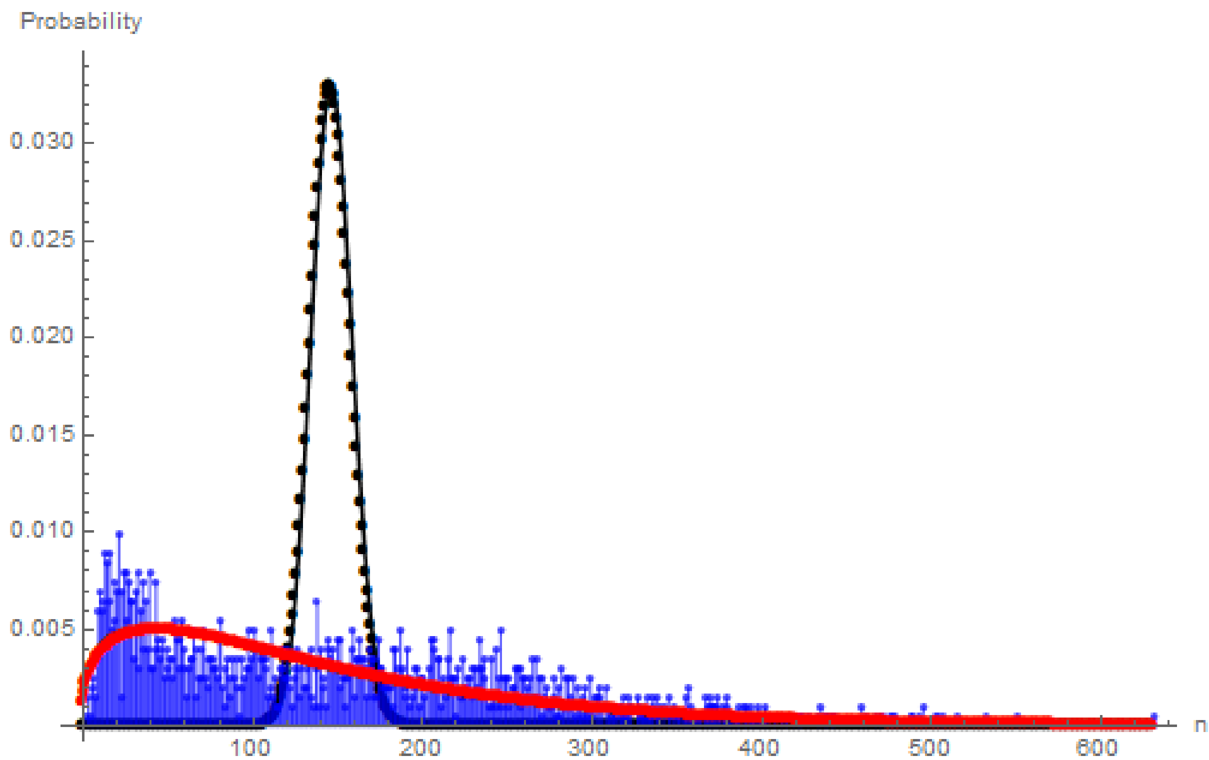


Figure 10.17: Canada accidents - Incidents per hour

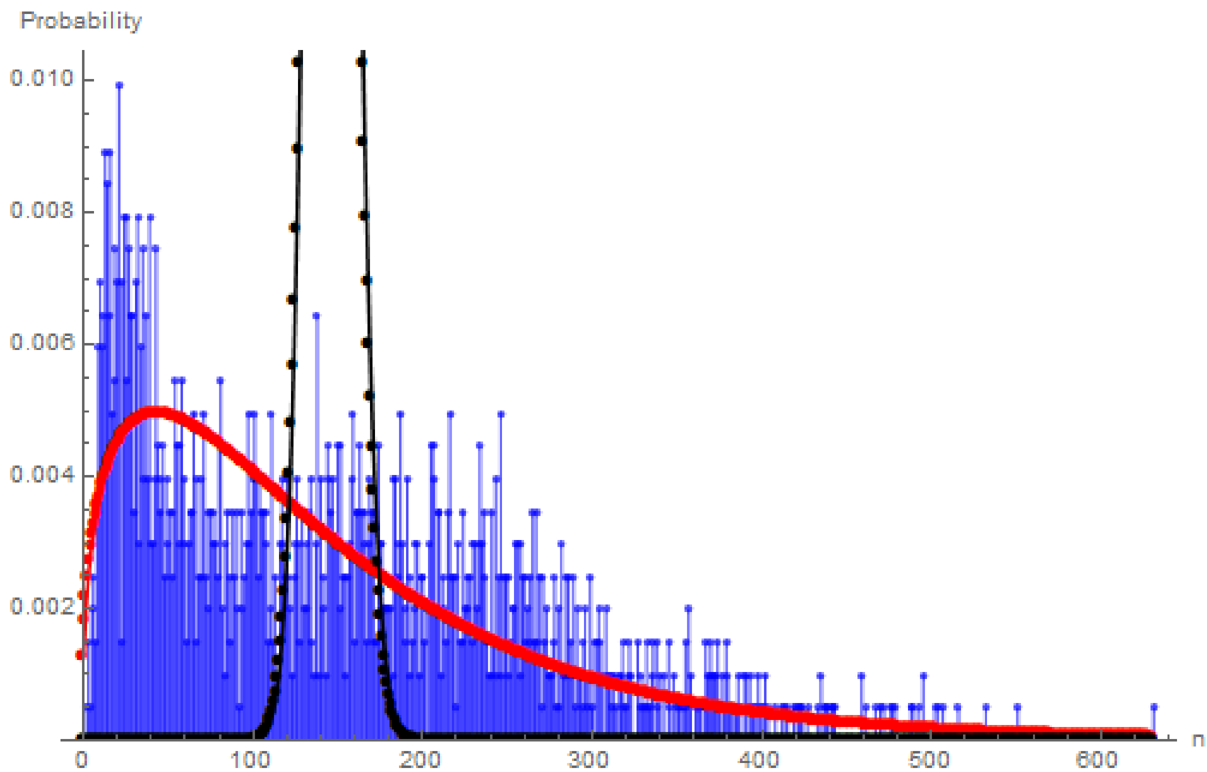


Figure 10.18: Canada accidents - Incidents per hour cropped

For the number of vehicles per hour, of the 27 weight functions that were tested, 12 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = (1.15352)_n$ and $\lambda = 0.996256$. This is shown in Figure 10.19. In this case, the weighted Poisson distribution gives a decent fit to the data.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	$(0.995999; 0.996513)$	$(0.995679; 0.996728)$	$(0.995678; 0.996712)$
a	$(1.08948; 1.21756)$	$(1.03408; 1.2974)$	$(1.04076; 1.29827)$

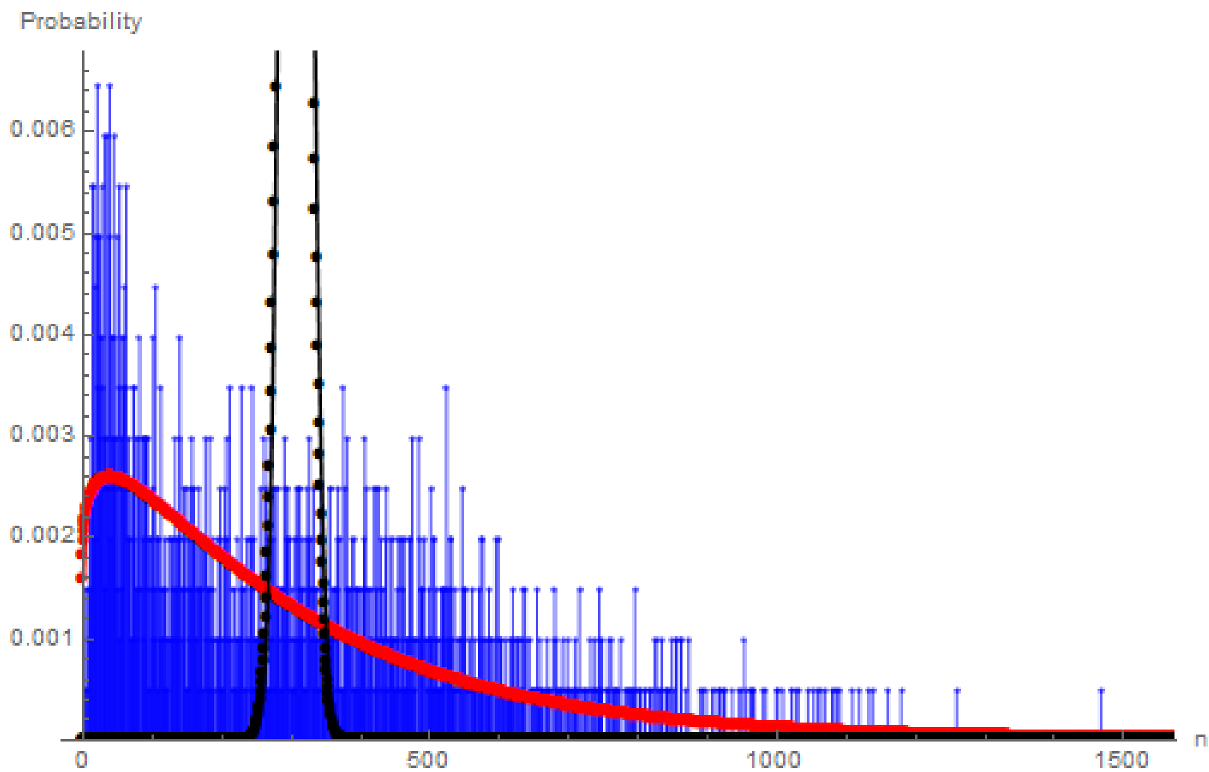


Figure 10.19: Canada accidents - Vehicles per hour

10.3.6 USA gun violence

For the number of fatalities per incident, of the 13 weight functions that were tested, 3 models perform better than the Poisson. The best fit occurs when $w(n;\phi) = \frac{\Gamma(2.47+n)}{n!\Gamma(2.47)} \frac{Beta(1894.47+n)}{Beta(1892,2.47)}$ and $\lambda = 78.1346$. This is shown in Figure 10.20. In this case, both the Poisson and weighted Poisson distributions give very good fits to the data. The AIC of the weighted Poisson distribution is only 0.1% smaller than that of the Poisson distribution.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	$(0; 184.332)$	$(0.0103529; 499.332)$	$(0.0107802; 112.431)$
a	$(0; 4473.66)$	$(0.01; 2928.57)$	$(0.01; 1203.33)$
b	$(2.11297; 2.82715)$	$(0.508997; 84.7026)$	$(0.389424; 99.713)$
r	$(2.11298; 2.82716)$	$(0.924693; 93.7917)$	$(0.770778; 116.815)$

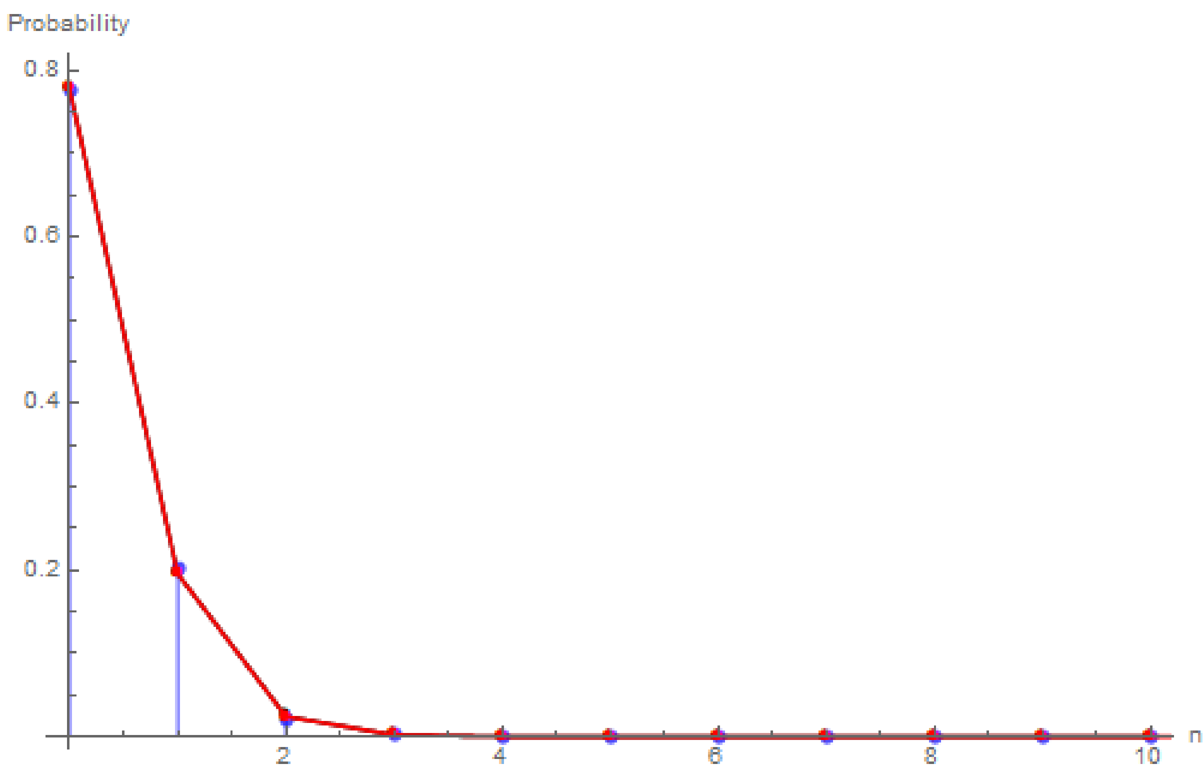


Figure 10.20: USA gun violence - Casualties per incident

For the number of injuries per day, of the 27 weight functions that were tested, 10 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = (14.3355)_n$ and $\lambda = 0.841052$. This is shown in Figure 10.21. In this case, the weighted Poisson distribution gives a good fit to the data.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	$(0.829742; 0.852362)$	$(0.828722; 0.852457)$	$(0.818613; 0.858138)$
a	$(13.1402; 15.5308)$	$(13.136; 15.6432)$	$(12.5645; 16.8175)$

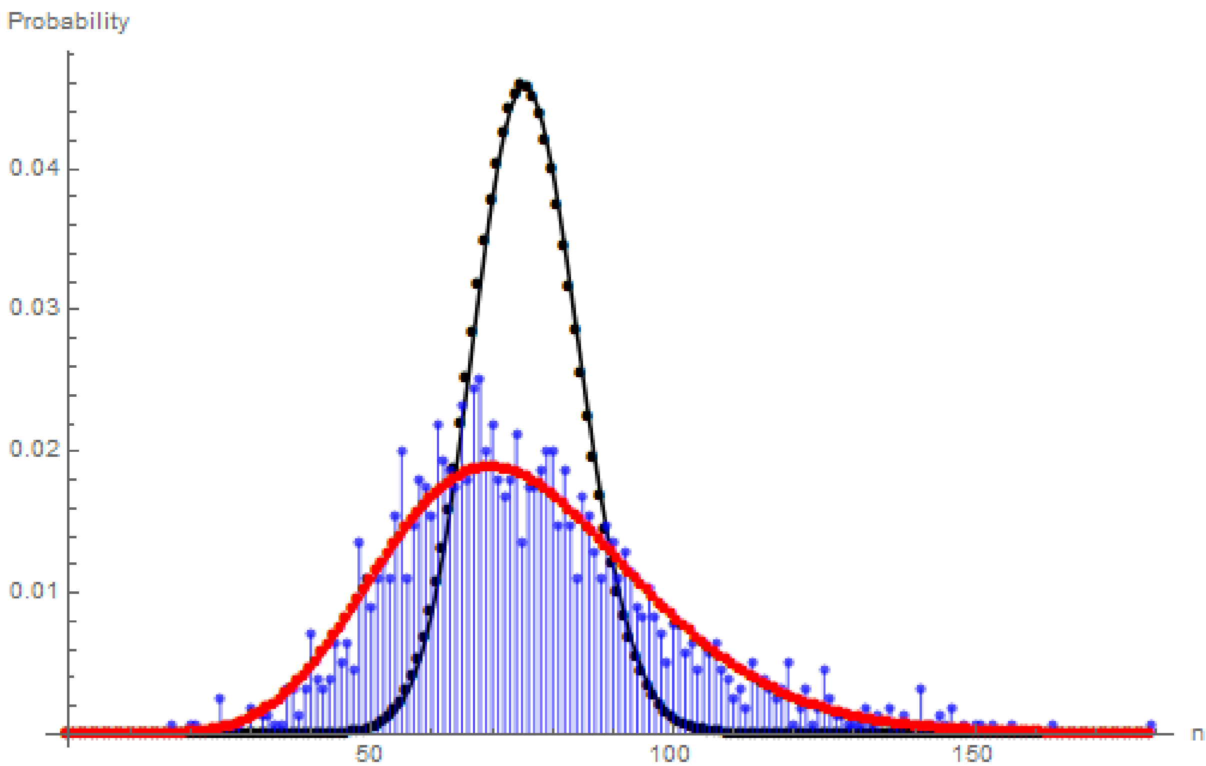


Figure 10.21: USA gun violence - Injuries per day

10.3.7 English Premier League matches

For the number of away team goals per game, of the 14 weight functions that were tested, 6 models perform better than the Poisson, and 1 additional model may (the AIC and AICc disagree with the BIC). The best fit is achieved when $w(n;\phi) = (10.0312)_n$ and $\lambda = 0.101221$. This is shown in Figure 10.22. In this case, both the Poisson and weighted Poisson distributions give a very good fit to the data. However, the weighted Poisson is clearly better.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	(0.0662003; 0.136241)	(0.0645825; 0.133055)	(0.0072111; 0.205716)
a	(6.17849; 13.8839)	(7.35292; 16.2707)	(4.33554; 152.284)

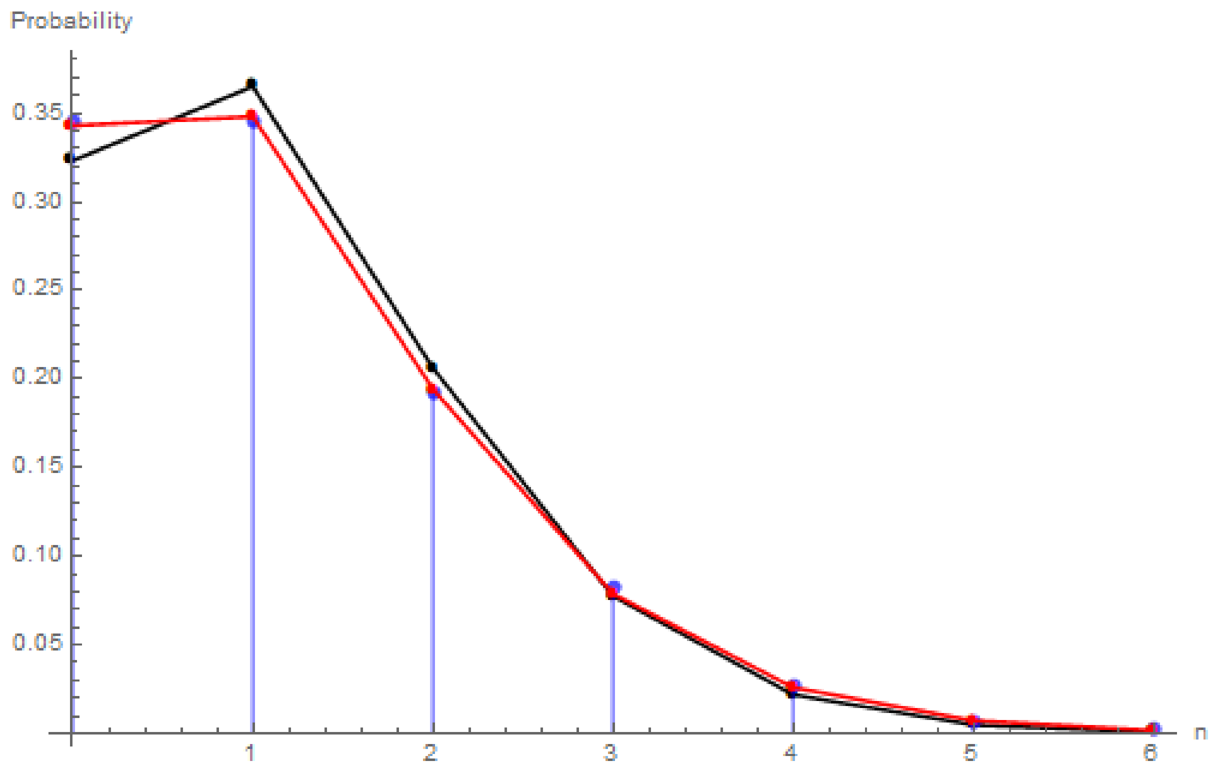


Figure 10.22: EPL games - Away team goals per game

For the number of home team shots on target per game, of the 13 weight functions that were tested, 8 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = (8.43405)_n$ and $\lambda = 0.440041$. This is shown in Figure 10.23. In this case the weighted Poisson distributions gives a very good fit to the data.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	(0.418597; 0.461485)	(0.417954; 0.459797)	(0.352096; 0.502905)
a	(7.70901; 9.15909)	(7.76895; 9.23608)	(6.55745; 12.0993)

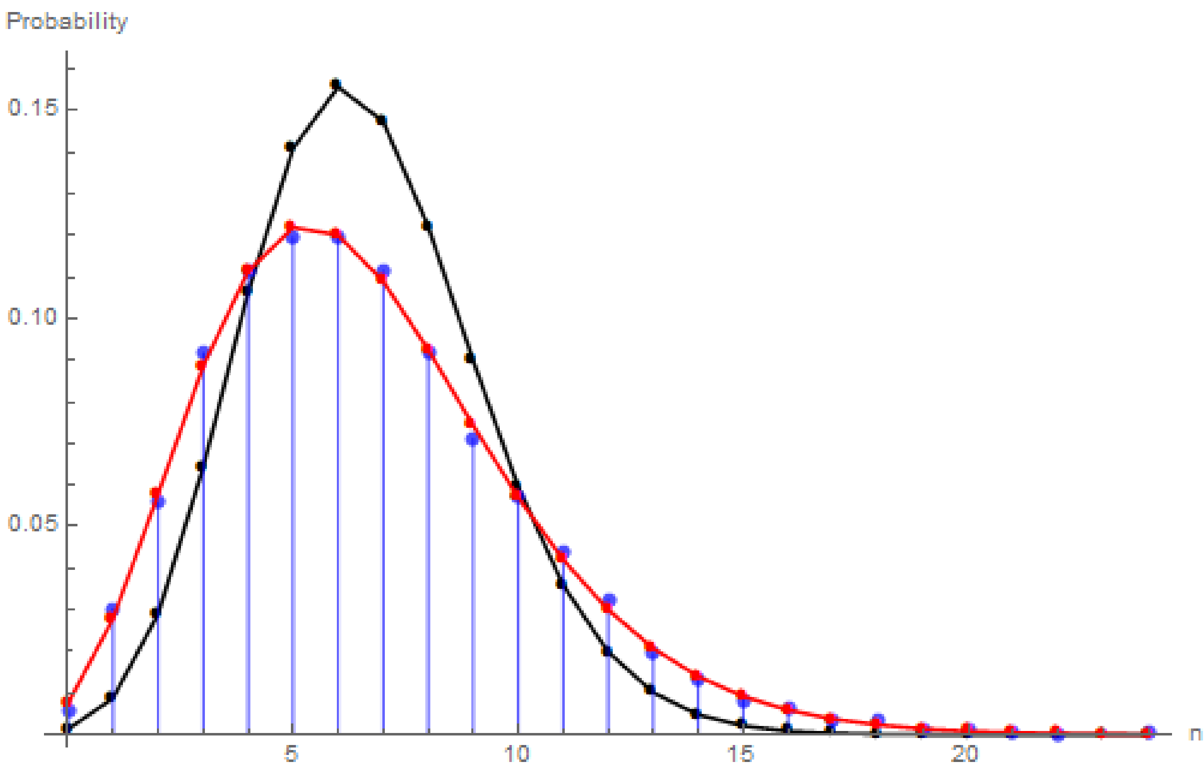


Figure 10.23: EPL games - Home team shots on target per game

For the number of away team shots on target per game, of the 14 weight functions that were tested, 7 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = (7.56268)_n$ and $\lambda = 0.401701$. This is shown in Figure 10.24. In this case, the weighted Poisson distributions gives a very good fit to the data.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information	Non-parametric bootstrap	Parametric bootstrap
λ	(0.378727; 0.424675)	(0.379324; 0.42346)	(0.315481; 0.467525)
a	(6.84868; 8.27668)	(6.91557; 8.29845)	(5.77729; 11.0422)

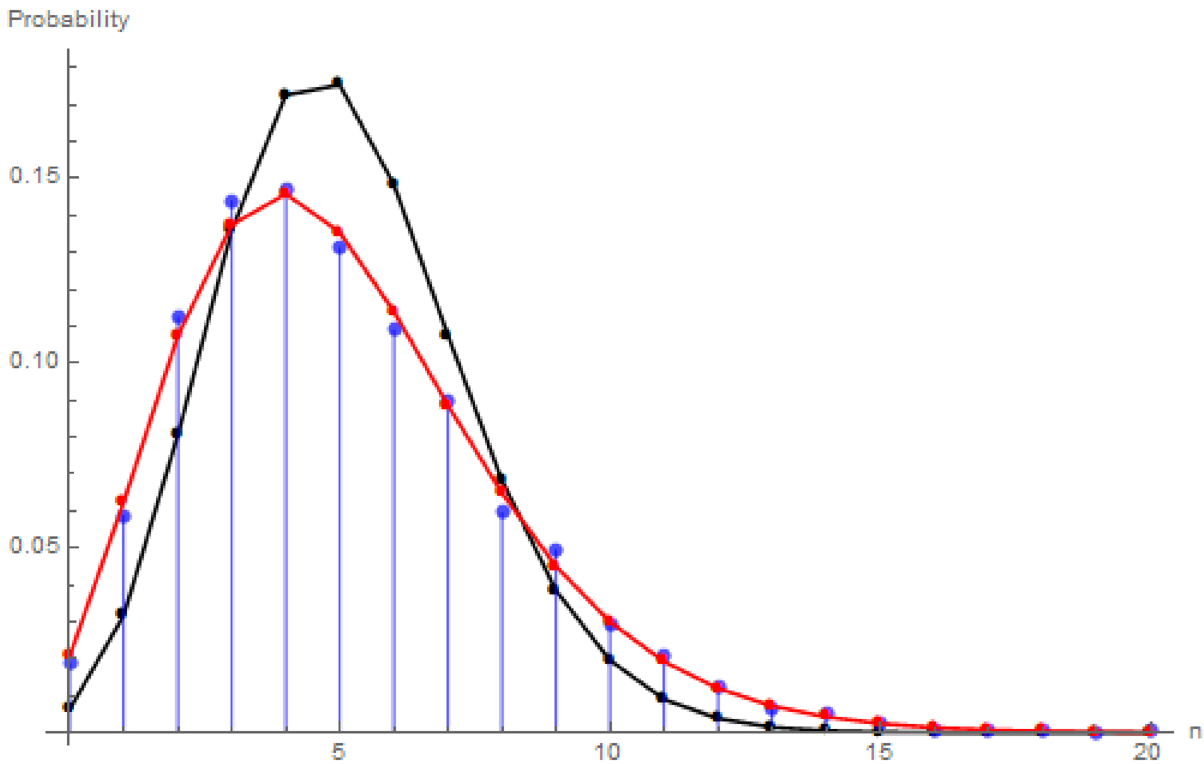


Figure 10.24: EPL games - Away team shots on target per game

10.3.8 USA flight delays

For the number of minutes spent taxiing before takeoff, of the 27 weight functions that were tested, 12 models perform better than the Poisson. The best fit is achieved when $w\left(n;\phi\right)=\frac{\Gamma(3.34795+n)}{n!\Gamma(3.34795)}\frac{Beta(2567.667,3.27899+n)}{Beta(2564.32,3.27899)}$ and $\lambda = 1884.32$. This is shown in Figure 10.25. In this case, the weighted Poisson distributions gives a better fit to the data, although the fitted model underestimates the number of observations close to the mode and overestimates the number of observations in the tails of the distribution.

The confidence intervals for the weighted Poisson parameters are as follows:

	Non-parametric bootstrap	Parametric bootstrap
λ	$(24.6591; 32626.3)$	$(61.3091; 11263.8)$
a	$(30.0165; 43080.3)$	$(52.8322; 15430.3)$
b	$(1.84894; 11.8597)$	$(0.599897; 9.5549)$
r	$(0.997799; 4.57705)$	$(0.31564; 6.41553)$

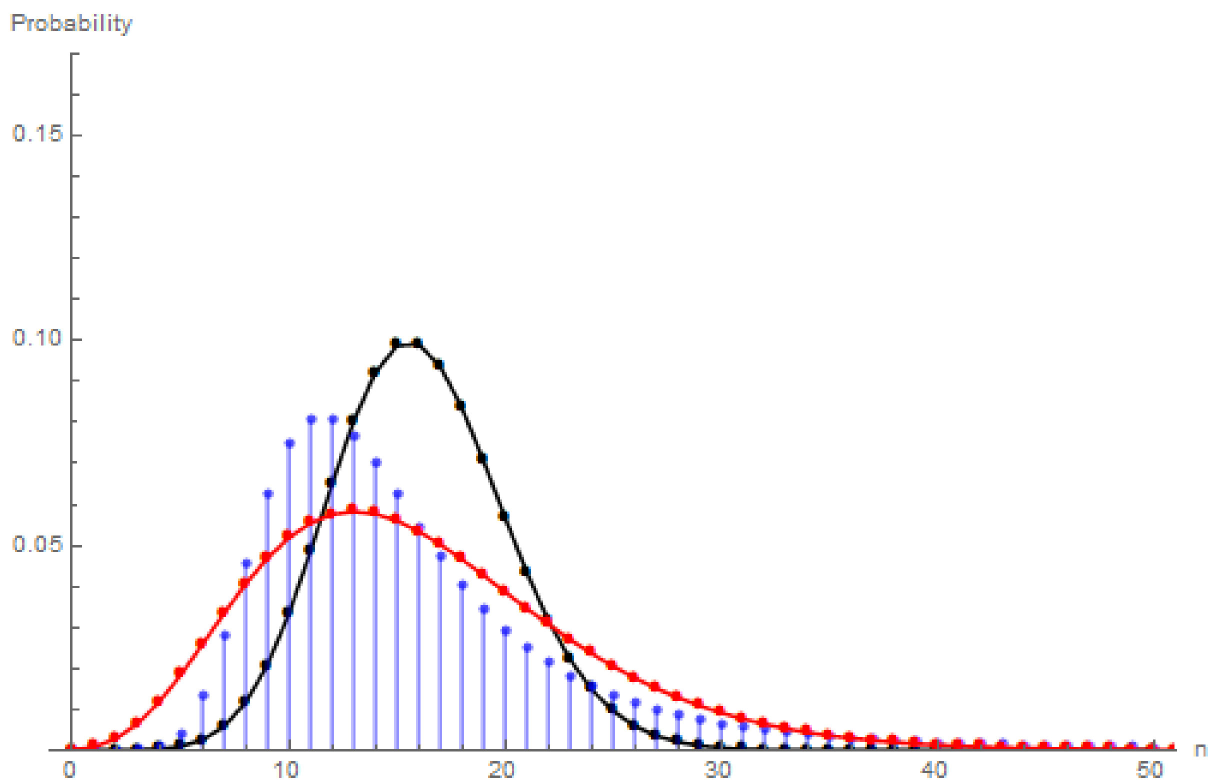


Figure 10.25: USA flights - Departure taxi time

For the number of minutes spent taxiing after landing, of the 27 weight functions that were tested, 16 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = \frac{\Gamma(2.51019+n)}{n!\Gamma(2.51019)} \frac{Beta(1603.52,2.47317+n)}{Beta(1601.01,2.47317)}$ and $\lambda = 1019.78$. This is shown in Figure 10.26. In this case, the weighted Poisson distributions gives a decent fit to the data, although the frequency associated with the mode is underestimated.

The confidence intervals for the weighted Poisson parameters are as follows:

	Non-parametric bootstrap	Parametric bootstrap
λ	$(55.5336; 29285.6)$	$(11.3282; 8157.55)$
a	$(83.9117; 39976.4)$	$(3.73751; 13421.6)$
b	$(1.56802; 5.04119)$	$(0.188981; 6.44642)$
r	$(1.18525; 4.16455)$	$(0.387331; 8.71784)$

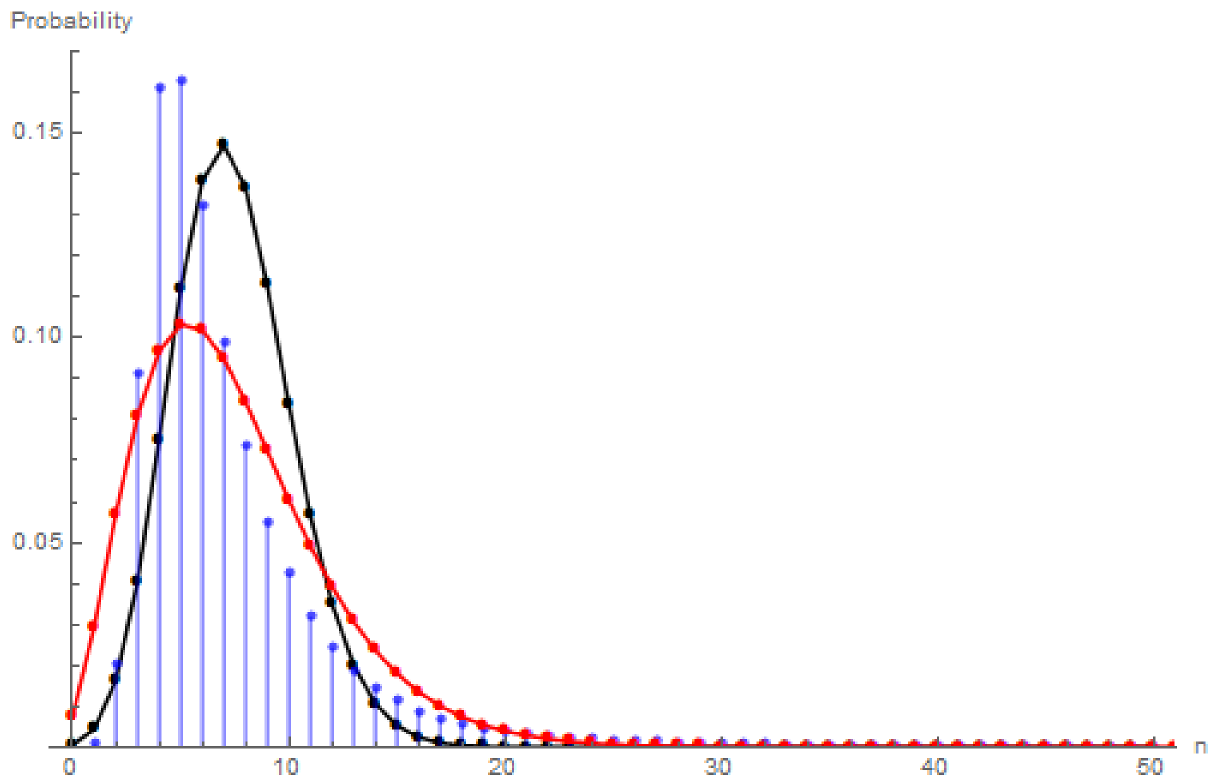


Figure 10.26: USA flights - Arrival taxi time

For the number of minutes delayed before departure, of the 26 weight functions that were tested, 12 models perform better than the Poisson. The best fit is achieved when $w(n;\phi) = 22903 \times \frac{\Gamma(n)\Gamma(22904)}{\Gamma(n+22904)}$ and $\lambda = 22896.9$. This is shown in Figure 10.27. In this case, the weighted Poisson distributions gives a decent fit to the data, although the probability of being delayed for 1 minute is dramatically overestimated.

The confidence intervals for the weighted Poisson parameters are as follows:

	Fisher Information
λ	(22896.4; 22897.4)
a	(22902.5; 22903.5)

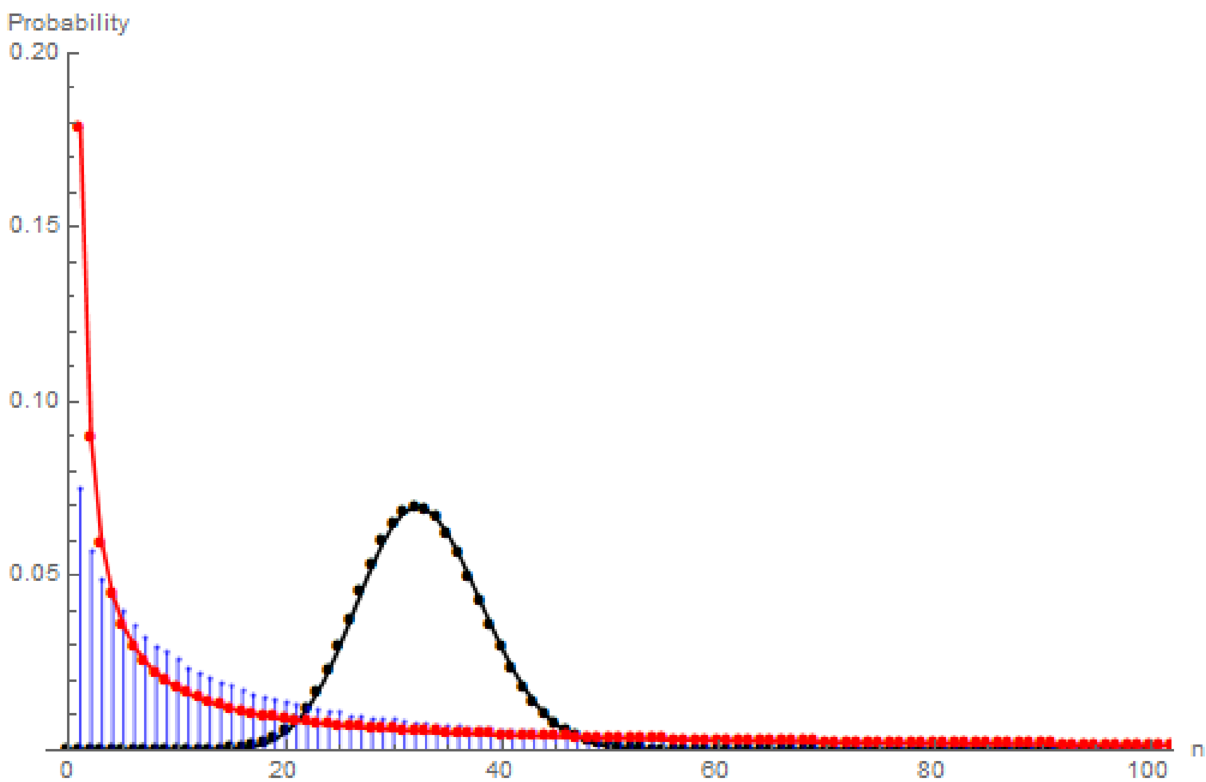


Figure 10.27: USA flights - Departure delay

10.4 Tables

10.4.1 Novel distributional data fits

Note:

- Some of the parameters of the weight functions have been relabeled. This is done to enable the tables to better fit onto a single page.

- Some parameter estimates are either extremely small, extremely large or extremely close to their boundaries. These values have all been rounded. For example, in the case of the zero-inflated Poisson distribution, it was often found that the fit of the distribution was extremely poor, $\hat{\varepsilon} = 2 \times 10^{-28}$. Even though $\hat{\varepsilon} > 0$, this value would be rounded to 0 in the tables below. Similarly, to avoid computational errors, variables that are strictly larger than 0 were often set minimum bounds of 0.00001. Purely for simplicity's sake these values will also be rounded to 0 since arbitrary precision could be achieved with no practical change in the fit of the weighted Poisson distribution.

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	42.6923						-227.19	456.38	456.46	458.331
$w(n) = n$	41.6923						-228.519	459.038	459.118	460.989
$w(n;\phi) = n^{-a}$	56.9258	1					-406.165	816.33	816.575	820.232
$w(n;\phi) = n + p$	42.6923					1078390	-227.19	458.38	458.624	462.282
$w(n;\phi) = an^3 + bn^2 + cn$	41.241	0.0277221	0	165.919			-228.215	464.43	465.281	482.235
$w(n;\phi) = (n+a)(n-b)^2$	48.8243	1324560	54.2155				-212.731	431.463	431.963	437.316
$w(n;\phi) = (n+a)(n^2 - bn + c)$	44.2058	14364.6334	132.26128	5110.2942			-225.2441	458.488	459.339	466.293
$w(n;\phi) = a + \frac{b-ac}{n+c}$	42.6929	8.81485	0.459659	1			-227.249	462.498	463.349	470.303
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	28.2316				23	1	-247.059	500.118	500.618	505.972
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	1.56039				97	0.958191	-355.457	716.914	717.414	722.768
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	125.28	79.4522					-201.99	407.979	408.224	411.882
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	46.9285	3	23				-221.997	449.994	450.494	455.848
$w(n;\phi) = \frac{-1}{ln(1-p)}\frac{p^n}{n}$	108.477					0.403007	-225.876	455.753	455.997	459.655
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	6358.1	8883.77	8.38997		8.60047		-200.069	408.139	408.99	415.944
$w(n;\phi) = Boole(n \geq a)$	42.6842	23					-227.1703	458.341	458.586	462.243
$w(n;\phi) = Boole(n \leq b)$	42.6923		73				-227.189	458.379	458.624	462.281
$w(n;\phi) = Boole(n \geq a)\, Boole(n \leq b)$	42.6845	23	73				-227.1699	460.34	460.84	466.193
$w(n;\phi) = \binom{n}{a}$	29.6923	13					-253.89	544.78	512.025	515.683
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	42.6923					0	-227.189	458.38	458.624	462.282
$w(n;\phi) = (a)_n$	0.683935	19.7293					-200.17	404.34	404.585	408.243
$w(n;\phi) = (n)_a$	42.6923	0					-227.19	458.38	458.624	462.282
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	30.1297				5	0.645281	-222.587	451.175	451.675	457.028
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	43.7136	13225.9	1.12249	13097.9			-225.876	459.753	460.604	467.558
$w(n;\phi) = \left(\frac{-1}{ln(1-p)}\frac{p^n}{n}\right)^{-1}$	0.601794					0.0144342	-228.519	461.038	461.282	464.94
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	41.6923	0					-228.519	461.038	461.282	464.94
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	121.236	0.072205	640.686		96.7254		-201.763	411.526	412.377	419.331
$w(n;\phi) = \binom{n}{a}^{-1}$	42.6923	0					-227.19	458.38	458.624	462.282
$w(n;\phi) = (a)_n^{-1}$	1	0					-12515.5	25034.9	25035.2	25038.8
$w(n;\phi) = (n)_a^{-1}$	125.28	80.4522					-201.99	407.979	408.224	411.882

Table 10.1: Weekly sales - Item 409

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n)=1$	1.11538						-66.3403	134.681	134.761	136.632
$w(n;\phi)=n+p$	0.540675					0.400104	-64.4861	132.972	133.217	136.875
$w(n;\phi)=(n+a)(n-b)^2$	0.213082	0.00000165	0.257714				-64.2744	134.549	135.049	140.403
$w(n;\phi)=(n+a)(n^2-bn+c)$	0.24	162903	0.391944	0.0947748			-64.2616	136.523	137.374	144.328
$w(n;\phi)=a+\frac{b-ac}{n+c}$	0.586767	1.39547	0.547372	10			-64.4915	136.983	137.834	144.788
$w(n;\phi)=\binom{n+r-1}{n}p^n(1-p)^r$	0.420488				7	0.59767	-64.4242	134.848	135.348	140.702
$w(n;\phi)=\binom{r}{n}p^n(1-p)^{r-n}$	0.773401				13	0.173523	-64.5231	135.046	135.546	140.9
$w(n;\phi)=\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	0.211438	39.6573	29.2015		30.081		-64.3878	134.776	135.276	140.629
$w(n;\phi)=Boole(n\le b)$	1.14132		4				-66.026	136.052	136.297	139.955
$w(n;\phi)=p\frac{n!}{e^{-\lambda}\lambda^n}Boole(n=0)+(1-p)$	1.11538					0	-66.3403	136.684	136.926	140.583
$w(n;\phi)=(a)_n$	0.000566858	1965.94					-66.3451	136.69	136.935	140.593
$w(n;\phi)=\left(\binom{n+r-1}{n}p^n(1-p)^r\right)^{-1}$	1.17465				2	1	-68.5086	143.017	143.517	148.871
$w(n;\phi)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$	0.541014	0	3.37762	0.400104			-64.4861	136.972	137.823	144.777
$w(n;\phi)=\left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	0.0628858	30.2442	1.04496		1.04577		-64.3863	136.773	137.624	144.578
$w(n;\phi)=(a)_n^{-1}$	1	0.404091					-65.8125	135.625	135.87	139.527

Table 10.2: Weekly sales - Item 726

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n)=1$	23.0116						-74390.2	148782	148782	148789
$w(n;\phi)=n+p$	23.0116					7.30×10^7	-74390.2	148784	148784	148797
$w(n;\phi)=(n+a)(n-b)^2$	24.2626	1.09×10^8	25.4228				-65820.3	131647	131647	131665
$w(n;\phi)=(n+a)(n^2-bn+c)$	22.4594	35.9042	38.9042	1398.82			-73532.4	147073	147073	147098
$w(n;\phi)=a+\frac{b-ac}{n+c}$	24.0116	0	665.221	1			-71728.3	143465	143465	143489
$w(n;\phi)=\binom{n+r-1}{n}p^n(1-p)^r$	23.0327				1	0.9999081	-74390.2	148786	148786	148805
$w(n;\phi)=\binom{r}{n}p^n(1-p)^{r-n}$	Computationally intractable									
$w(n;\phi)=\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	245318	252204	0.822409		0.822916		-14904.8	29817.6	29817.6	29842.4
$w(n;\phi)=Boole(n\le b)$	23.0116		583				-74390.2	148784	148784	148797
$w(n;\phi)=p\frac{n!}{e^{-\lambda}\lambda^n}Boole(n=0)+(1-p)$	23.0116					0	-74390.2	148784	148784	148797
$w(n;\phi)=(a)_n$	0.971239	0.681436					-14912.6	29829.1	29829.1	29841.5
$w(n;\phi)=\left(\binom{n+r-1}{n}p^n(1-p)^r\right)^{-1}$	27.3793				21	0.636568	-48600.9	97207.8	97207.8	97226.4
$w(n;\phi)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$	24.3737	9712.73	0.0004299	311869			-70495.7	140999	140999	141024
$w(n;\phi)=\left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	Convergence									
$w(n;\phi)=(a)_n^{-1}$	1		0.0045999				-499431	998866	998866	998878

Table 10.3: Airplane accidents - Fatalities per incident

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n)=1$	6.10943						-1420.39	2842.78	2842.78	2847.16
$w(n;\phi)=n+p$	6.10943					8.03×10^6	-1420.39	2844.78	2844.8	2853.55
$w(n;\phi)=(n+a)(n-b)^2$	3.65182	1.31×10^7	2.37104				-1542.26	3090.52	3090.57	3103.69
$w(n;\phi)=(n+a)(n^2-bn+c)$	10.0025	183.203	36.894	14873.7			-1952.72	3913.44	3913.51	3930.99
$w(n;\phi)=a+\frac{b-ac}{n+c}$	7.10358	0	3037.5	1			-1416.28	2840.57	2840.64	2858.11
$w(n;\phi)=\binom{n+r-1}{n}p^n(1-p)^r$	6.11554					0.999001	-1420.39	2846.78	2846.82	2859.94
$w(n;\phi)=\binom{r}{n}p^n(1-p)^{r-n}$	2.86169				32	0.351522	-1600.13	3206.27	3206.31	3219.43
$w(n;\phi)=\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	20.2736	106.103	44.2786		0.641189		-1416.07	2840.13	2840.2	2857.68
$w(n;\phi)=Boole(n\le b)$	6.11031	17					-1420.35	2844.69	2844.71	2853.47
$w(n;\phi)=p\frac{n!}{e^{-\lambda}\lambda^n}Boole(n=0)+(1-p)$	6.10943					0	-1420.39	2844.78	2844.8	2854.55
$w(n;\phi)=(a)_n$	0.148743	34.9643					-1416.38	2836.76	2836.78	2845.53
$w(n;\phi)=\left(\binom{n+r-1}{n}p^n(1-p)^r\right)^{-1}$	7.09647				2	0.998999	-1416.28	2838.57	2838.61	2851.73
$w(n;\phi)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$	7.08583	75.8755	81.352	2189.85			-1416.28	2840.55	2840.62	2858.1
$w(n;\phi)=\left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	7.0304	0	205.421		1.99333		-1416.28	2840.57	2840.64	2858.12
$w(n;\phi)=(a)_n^{-1}$	1	0.002117					-8007.85	16019.7	16019.7	16028.5

Table 10.4: Airplane accidents - Incidents per month

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	4.23934						-5051.88	10105.8	10105.8	10111.1
$w(n;\phi) = n + p$	4.23834					3.515×10^7	-5051.88	10107.8	10107.8	10118.5
$w(n;\phi) = (n+a)(n-b)^2$	2.20844	5.68×10^7	1.37183				-5385.64	10777.3	10777.3	10793.4
$w(n;\phi) = (n+a)(n^2 - bn + c)$	4.22476	85476.5	1.03014	2585			-5047.97	10103.9	10104	10125.4
$w(n;\phi) = a + \frac{b-ac}{n+c}$	5.20972	0	11.0138	1			-4966.01	9940.03	9940.05	9961.5
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	4.2384				1	1	-5051.88	10109.8	10109.8	10125.9
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	Computationally intractable									
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	9526.54	19154.3	2.35224		2.37867		-3738.96	7485.92	7485.92	7507.39
$w(n;\phi) = Boole(n \leq b)$	4.23834		443				-5051.88	10107.8	10107.8	10118.5
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	4.23832					0	-5051.88	10107.8	10107.8	10118.5
$w(n;\phi) = (a)_n$	0.538578	3.63116					-3759.92	7523.85	7523.86	7534.59
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	8.90156				6	1	-4837.3	9678.6	9678.61	9689.34
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	Convergence									
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	22865.2	24167	55700.6		82222.7		-4052.34	8112.68	8112.71	7134.16
$w(n;\phi) = (a)_n^{-1}$	1	0.024201					-15150.9	30305.7	30305.7	30316.5

Table 10.5: Mass shootings - Injuries per incident

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	3.68127						-9192.71	18387.4	18387.4	18392.9
$w(n;\phi) = n + p$	3.68127					2.39×10^9	-9192.71	18389.4	18389.4	18400.4
$w(n;\phi) = (n+a)(n-b)^2$	2.57429	1.023×10^8	2.52017				-7843.98	15694	15698	15710.5
$w(n;\phi) = (n+a)(n^2 - bn + c)$	2.96312	3.43268	6.43268	24.0814			-8762.18	17532.4	17532.4	17554.4
$w(n;\phi) = a + \frac{b-ac}{n+c}$	4.63587	0	401953	1			-8294.91	16597.8	16597.8	16619.8
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	3.68127				1	1	-9192.71	18391.4	18391.4	18407.9
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	Computationally intractable									
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	4565.16	5728.15	3.10309		0.0470452		-3948.78	7905.57	7925.59	7927.61
$w(n;\phi) = Boole(n \leq b)$	3.68127		443				-9192.71	18389.4	18389.4	18400.4
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	3.68127					0	-9192.71	18389.4	18389.4	18400.4
$w(n;\phi) = (a)_n$	0.926499	0.92042					-4040.79	8085.59	8085.59	8096.61
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	16.6244				16	1	-6002.75	12011.5	12011.5	12028
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	8.61433	4631.92	6.11644	3748.83			-5546.01	11100	11100.1	11122.1
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	Convergence									
$w(n;\phi) = (a)_n^{-1}$	1	0.405254					-23544.2	47092.3	47092.3	47103.3

Table 10.6: Mass shootings - Injuries per day

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n)=1$	112.033						-1055.68	2113.17	2113.44	2115.46
$w(n)=n$	111.033						-1062.2	2126.39	2126.46	2128.49
$w(n;\phi)=n^{-a}$	160.077	30					-529.806	1063.61	1063.82	1067.8
$w(n;\phi)=n+p$	112.033					2.24×10^{10}	-1055.68	2115.37	2115.58	2119.56
$w(n;\phi)=an^3+bn^2+cn$	110.712	0.010865	0	702.956			-1061.13	2130.26	2130.98	2138.63
$w(n;\phi)=(n+a)(n-b)^2$	122.37	257622	131.722				-1000.89	2007.78	2008.21	2014.07
$w(n;\phi)=(n+a)(n^2-bn+c)$	111.19	2293.62	0	15658.6			-1056.23	2120.46	2121.19	2128.84
$w(n;\phi)=a+\frac{b-ac}{n+c}$	112.914	0.0100995	2599.32	16			-1050.72	2109.44	2110.16	2117.81
$w(n;\phi)=\binom{n+r-1}{n}p^n(1-p)^r$	41.6609				191	0.999662	-1531.66	3069.32	3069.75	3075.61
$w(n;\phi)=\binom{r}{n}p^n(1-p)^{r-n}$	Computationally intractable									
$w(n;\phi)=a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	3240.51	3090.74					-344.827			
$w(n;\phi)=\frac{ab^a}{n^{a+1}}$	Computationally intractable									
$w(n;\phi)=\frac{-1}{ln(1-p)}\frac{p^n}{n}$	146.305					0.772647	-1049.17	2102.35	2102.56	2106.53
$w(n;\phi)=\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	110254	114698	2.27168		3.34894		-318.573	645.145	645.873	653.523
$w(n;\phi)=Boole(n\geq a)$	112.033	33					-1055.68	2115.37	2115.58	2119.56
$w(n;\phi)=Boole(n\geq a)\,Boole(n\leq b)$	112.033	33	532				-1055.68	2117.17	2117.8	2123.65
$w(n;\phi)=\binom{n}{a}$	80.0333	32					-1364.47	2732.93	2733.14	2737.12
$w(n;\phi)=p\frac{n!}{e^{-\lambda}\lambda^n}Boole(n=0)+(1-p)$	112.033					0	-1055.68	2115.37	2115.58	2119.56
$w(n;\phi)=(a)_n$	0.95932	4.75082					-318.171	640.343	640.553	644.532
$w(n;\phi)=(n)_a$	112.033	0					-1055.68	2115.37	2115.58	2119.56
$w(n;\phi)=\left(\binom{n+r-1}{n}p^n(1-p)^r\right)^{-1}$	71.9513				2	0.63655	-1049.29	2104.58	2105.01	2110.87
$w(n;\phi)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$	113.042	6424.37	0.00117072	227907			-1049.17	2106.35	2107.07	2114.72
$w(n;\phi)=\left(\frac{-1}{ln(1-p)}\frac{p^n}{n}\right)^{-1}$	1.49224					0.0134396	-1062.2	2128.39	2128.6	2132.58
$w(n;\phi)=\left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	111.033	0					-1062.2	2128.39	2128.6	2132.58
$w(n;\phi)=\left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	5384.65	1884.69	6679.2		33151.2		-333.008	674.016	674.743	682.393
$w(n;\phi)=\binom{n}{a}^{-1}$	112.033	0					-1055.68	2115.37	2115.58	2119.56
$w(n;\phi)=(a)_n^{-1}$	1	0.00001					-51858.7	103721	103722	103726
$w(n;\phi)=(n)_a^{-1}$	3240.51	3091.74					-344.827	693.654	693.865	697.843

Table 10.7: Mass shootings -Injuries per month

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n)=1$	1.14061						-2858.79	5719.59	5719.59	5724.96
$w(n;\phi)=n+p$	1.14061					2.74×10^{7}	-2858.79	5721.59	5721.6	5732.33
$w(n;\phi)=(n+a)(n-b)^2$	1.29369	1.86215	2.55274				-2768.056	5543.11	5543.13	5559.22
$w(n;\phi)=(n+a)(n^2-bn+c)$	0.85154	2.32335	3.05372	6.96681			-2763.04	5534.09	5534.11	5555.56
$w(n;\phi)=a+\frac{b-ac}{n+c}$	1.53609	0	42.6905	2			-2775.18	5558.36	5558.36	5579.83
$w(n;\phi)=\binom{n+r-1}{n}p^n(1-p)^r$	1.14061				1	1	-2858.79	5723.59	5723.6	5739.7
$w(n;\phi)=\binom{r}{n}p^n(1-p)^{r-n}$	0.652784				87	0.0344308	-3679.83	7365.66	7365.68	7381.77
$w(n;\phi)=\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	15902.2	26485.7	0.874543		0.8741		-2339.55	4687.1	4687.12	4708.58
$w(n;\phi)=Boole(n\leq b)$	1.14061						-2858.79	5721.59	5721.6	5721.33
$w(n;\phi)=p\frac{n!}{e^{-\lambda}\lambda^n}Boole(n=0)+(1-p)$	1.14061					0	-2858.79	5721.59	5721.6	5721.33
$w(n;\phi)=(a)_n$	0.594393	0.778335					-2339.86	4683.71	4683.72	4694.45
$w(n;\phi)=\left(\binom{n+r-1}{n}p^n(1-p)^r\right)^{-1}$	8.79587				25	0.611163	-2426.7	4857.41	4857.41	4868.14
$w(n;\phi)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$	2.70857	599.884	109.958	600.533			-2591.12	5190.24	5190.27	5211.72
$w(n;\phi)=\left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	99140.3	385968	732877		380693		-2345.92	4699.083	4699.86	4721.31
$w(n;\phi)=(a)_n^{-1}$	1	0.773409					-3861.49	7726.99	7726.99	7737.72

Table 10.8: Mass shootings - Fatalities per incident

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	0.99069						-3446.38	6894.75	6894.75	6900.26
$w(n;\phi) = n + p$	0.99069					2.84×10^7	-3446.38	6896.75	6896.76	6907.77
$w(n;\phi) = (n+a)(n-b)^2$	1.05806	5.91×10^7	1.60114				-2896.17	5798.34	5798.35	5814.87
$w(n;\phi) = (n+a)(n^2 - bn + c)$	0.812589	1.34192	4.09836	7.30162			-3170.44	6348.87	6348.89	6370.91
$w(n;\phi) = a + \frac{b-ac}{n+c}$	1.58111	0.01000	305032	1			-3172.59	6353.17	6353.19	6375.21
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	0.99069				1	1	-3446.38	6898.75	6898.76	8915.28
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	0.264656				86	0.0717673	-4687.64	9381.29	9381.3	9397.82
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	23723.9	29986.8	0.368011		0.792581		-2345.9	4699.8	4699.82	4721.84
$w(n;\phi) = Boole(n \leq b)$	0.99069		62				-3446.38	6896.75	6896.76	6907.77
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	0.99069					0	-3446.38	6896.75	6896.76	6907.77
$w(n;\phi) = (a)_n$	0.77284	0.291193					-2346.05	4696.11	4696.12	4707.13
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	12.8065				24	0.998971	-2649.28	5304.57	5304.58	5321.1
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	3.72279	127.336	4.8648	20396.4			-2596.13	5200.26	5200.28	5222.3
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	35088	277523	220115		234510		-2519.58	5047.17	5047.19	5069.21
$w(n;\phi) = (a)_n^{-1}$	1	1.19751					-4732.6	9469.19	9469.2	9480.21

Table 10.9: Mass shootings - Fatalities per day

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	30.15						-320.808	643.615	643.684	645.709
$w(n) = n$	29.15						-324.849	653.698	653.909	657.887
$w(n;\phi) = n^{-a}$	43.2861	11					-485.048	972.095	972.164	974.19
$w(n;\phi) = n + p$	30.1499					316932	-320.808	645.645	645.804	649.804
$w(n;\phi) = an^3 + bn^2 + cn$	28.8351	0.00406793	0.00026508	16.2694			-324.079	656.158	656.885	664.535
$w(n;\phi) = (n+a)(n-b)^2$	35.1377	1.97×10^6	39.4537				-309.305	624.61	625.038	630.893
$w(n;\phi) = (n+a)(n^2 - bn + c)$	29.8039	331.06	15.1988	4714.85			-319.794	647.589	648.316	655.966
$w(n;\phi) = a + \frac{b-ac}{n+c}$	31.1165	0.0100414	914.533	2			-317.281	642.562	643.28	650.94
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	60.15				1	1	-320.808	645.615	645.826	649.804
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	Computationally intractable									
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	146.386	110.417					-245.84	495.681	495.891	499.869
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	Computationally intractable									
$w(n;\phi) = \frac{-1}{ln(1-p)} \frac{p^n}{n}$	48.7061					0.640293	-316.803	637.606	637.817	641.795
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	26459.4	326410.2	3.88403		3.87992		-232.338	472.677	473.404	481.054
$w(n;\phi) = Boole(n \geq a)$	30.1471	13					-320.798	645.597	645.807	649.785
$w(n;\phi) = Boole(n \leq b)$	30.15		96				-320.808	645.615	645.826	649.804
$w(n;\phi) = Boole(n \geq a)\, Boole(n \leq b)$	30.1471	13	96				-320.798	647.597	648.025	653.88
$w(n;\phi) = \binom{n}{a}$	18.15	12					-406.981	817.962	818.172	822.15
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	30.15					0	-320.808	645.615	645.826	649.804
$w(n;\phi) = (a)_n$	0.803742	7.36203					-232.631	469.262	469.472	473.45
$w(n;\phi) = (n)_a$	30.15	0					-320.808	645.615	645.826	649.804
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	23.6505				8	0.636621	-299.181	604.362	604.791	610.645
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	31.181	773.637	0.0107979	6706.91			-316.803	641.606	642.334	649.984
$w(n;\phi) = \left(\frac{-1}{ln(1-p)} \frac{p^n}{n}\right)^{-1}$	28.8595					0.99033	-324.849	653.698	653.909	657.887
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	29.15	0					-324.849	653.698	653.909	657.887
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	189.838	11.7074	271905		174.107		-242.213	492.427	493.154	500.804
$w(n;\phi) = \binom{n}{a}^{-1}$	36.2011	5					-298.254	600.508	600.719	604.697
$w(n;\phi) = (a)_n^{-1}$	1	0					-9175.96	18355.9	18356.1	18360.1
$w(n;\phi) = (n)_a^{-1}$	126.387	111.417					-245.84	495.681	495.891	499.869

Table 10.10: Mass shootings - Fatalities per month

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	5.37894						-5255.57	10513.1	10513.1	10518.5
$w(n) = n$	4.37894						-5284.02	10570	10570.1	10575.4
$w(n;\phi) = n^{-a}$	5.35349						-7730.6	15465.2	15465.2	15475.9
$w(n;\phi) = n + p$	5.37894					3.51×10^7	-5255.57	10515.1	10515.1	10525.9
$w(n;\phi) = an^3 + bn^2 + cn$	4.36418	0.00727	0	24.9933			-5279.88	10567.8	10567.8	10589.2
$w(n;\phi) = (n+a)(n-b)^2$	3.39887	3.576×10^7	0.557962				-5345.11	10696.2	10696.2	10712.3
$w(n;\phi) = (n+a)(n^2 - bn + c)$	3.41962	2981.51	1.00008	0.250042			-5345.16	10698.3	10698.3	10719.8
$w(n;\phi) = a + \frac{b-ac}{n+c}$	5.02137	1.05×10^6	0.01	2			-5234.55	10471.1	10471.1	10492.6
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	5.37894				1	1	-5255.57	10517.1	10517.2	10533.2
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	Computationally intractable									
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	1.27584×10^8	1.36161×10^8					-4650.39	9304.79	9301.8	9315.53
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	19.5315	9	4				-3498.29	7002.58	7002.6	7018.69
$w(n;\phi) = \frac{-1}{ln(1-p)} \frac{p^n}{n}$	7.98941					0.82352	-5252.74	10509.5	10509.5	10520.2
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	20641.2	27455.8	1.53417		1.50215		-4040.13	8088.26	8088.28	8109.73
$w(n;\phi) = Boole(n \geq a)$	3.99914	4					-4694.22	9392.44	9392.45	9403.18
$w(n;\phi) = Boole(n \leq b)$	5.37894		501				-5255.57	10515.1	10515.1	10525.9
$w(n;\phi) = Boole(n \geq a)\, Boole(n \leq b)$	3.99914	4	500				-4694.22	9392.44	9392.45	9403.18
$w(n;\phi) = \binom{n}{a}$	5.37894	0					-5255.57	10515.1	10515.1	10525.9
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	5.37894					0	-5255.57	10515.1	10515.1	10525.9
$w(n;\phi) = (a)_n$	0.502206	5.33169					-3865.19	7734.38	7734.39	7745.12
$w(n;\phi) = (n)_a$	5.32812	0.0204444					-5248.12	10500.2	10500.2	10511
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	11.1393				7	1	-5183.78	10373.6	10373.6	10389.7
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	5.05654	1.05711	1.77551	0			-5231.38	10470.8	10470.8	10492.2
$w(n;\phi) = \left(\frac{-1}{ln(1-p)} \frac{p^n}{n}\right)^{-1}$	3.19286					0.72914	-5284.02	10572	10572.1	10582.8
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	4.37894	0					-5284.02	10572	10572.1	10582.8
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	52.9669	1.0141421	4664.97		55.4862		-4925.07	9858.14	9858.17	9879.62
$w(n;\phi) = \binom{n}{a}^{-1}$	5.37894	0					-5255.57	10515.1	10515.1	10525.9
$w(n;\phi) = (a)_n^{-1}$	1	0					-20413.1	40830.1	40830.1	80840.8
$w(n;\phi) = (n)_a^{-1}$	1.72556×10^8	1.8416×10^8					-4650.39	9304.79	9304.8	9315.53

Table 10.11: Mass shootings -Victims per incident

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	0.868565						-2387.48	4776.96	4776.96	4782.47
$w(n;\phi) = n + p$	0.868565					2.596×10^7	-2387.48	4778.96	4778.96	4789.98
$w(n;\phi) = (n+a)(n-b)^2$	0.248353	3.39×10^7	0.406871				-2451.08	4908.16	4908.17	4924.69
$w(n;\phi) = (n+a)(n^2 - bn + c)$	0.639885	1.16682	3.31916	5.02521			-2335.45	4678.91	4678.93	4700.95
$w(n;\phi) = a + \frac{b-ac}{n+c}$	1.41435	0	45.8972	1			-2347.37	4700.75	4700.76	4717.28
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	0.868565				1	1	-2387.48	4780.96	4780.97	4797.49
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	0.223217				33	0.15822	-2629.43	5264.85	5264.86	5281.38
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	5.86299	12.1811	1.22613		1.22613		-2333.45	4674.89	4674.91	4696.93
$w(n;\phi) = Boole(n \leq b)$	0.868783		6				-2387.41	4778.83	4778.84	4789.85
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	0.868586					0	-2387.48	4778.96	4778.96	4789.98
$w(n;\phi) = (a)_n$	0.294394	2.08179					-2333.77	4671.54	4671.55	4682.56
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	2.90984				5	0.998999	-2333.99	4673.97	4673.99	4690.5
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	1.71337	21.8731	11.0648	283.827			-2336.58	4681.16	4681.18	4703.2
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	2.63085	1.30812	53.68		5.04484		-2333.99	4675.99	4676.01	4698.03
$w(n;\phi) = (a)_n^{-1}$	1	0.905651					-2657.2	5318.4	5318.41	5329.42

Table 10.12: Mass shootings - Incidents per day

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	26.4333						-236.732	475.464	475.533	477.558
$w(n) = n$	25.4333						-238.919	479.837	479.936	481.932
$w(n;\phi) = n^{-a}$	36.2731	9					-415.142	834.285	834.495	939.474
$w(n;\phi) = n + p$	26.4333					1.4×10^7	-236.732	477.464	477.674	481.653
$w(n;\phi) = an^3 + bn^2 + cn$	24.9693	0.02811	0	60.4092			-238.431	480.862	481.072	485.051
$w(n;\phi) = (n+a)(n-b)^2$	28.043	188.714	29.6047				-224.964	455.928	456.357	462.211
$w(n;\phi) = (n+a)(n^2 - bn + c)$	25.9335	1002.01	2.62906	2118.53			-236.209	480.418	481.145	488.795
$w(n;\phi) = a + \frac{b-ac}{n+c}$	27.3949	0	499.363	2			-234.909	477.817	478.544	496.194
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	20.9492				8	0.999018	-248.624	503.249	503.677	509.532
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	Computationally intractable									
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	65.7355	36.5779					-214.656	433.312	433.523	437.501
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	Computationally intractable									
$w(n;\phi) = \frac{-1}{ln(1-p)}\frac{p^n}{n}$	59.0935					0.464933	-234.595	473.19	473.4	477.378
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	873.766	1216.87	5.83227		5.81087		-213.003	434.006	434.733	442.383
$w(n;\phi) = Boole(n \geq a)$	26.4293	11					-236.718	477.435	477.696	481.219
$w(n;\phi) = Boole(n \leq b)$	26.4334	49					-236.732	477.464	477.725	481.247
$w(n;\phi) = Boole(n \geq a)\, Boole(n \leq b)$	26.43	11	49				-236.716	479.432	479.965	485.108
$w(n;\phi) = \binom{n}{a}$	23.4333	3					-244.079	494.157	494.586	500.44
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	26.4333					0	-236.732	477.464	477.674	481.653
$w(n;\phi) = (a)_n$	0.648321	14.3382					-213.078	430.155	430.366	434.344
$w(n;\phi) = (n)_a$	26.4333	0					-236.732	477.464	477.674	481.653
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	18.7383				4	0.636635	-231.365	468.729	469.158	475.012
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	27.4739	1288.17	0.129728	48849.7			-234.595	477.19	477.917	485.567
$w(n;\phi) = \left(\frac{-1}{ln(1-p)}\frac{p^n}{n}\right)^{-1}$	0.936754					0.0368317	-238.919	481.837	482.048	486.026
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	25.4333	1.72*10^-9					238.919	481.837	482.048	486.026
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	69.9403	2.11859	921.197		48.2353		-214.207	436.414	437.141	444.791
$w(n;\phi) = \binom{n}{a}^{-1}$	26.4333	0					-236.732	477.464	477.674	481.653
$w(n;\phi) = (a)_n^{-1}$	1	0					-7523.52	15051	15051.3	15055.2
$w(n;\phi) = (n)_a^{-1}$	65.7355	37.5779					-214.656	433.312	433.523	437.501

Table 10.13: Mass shootings - Incidents per month

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	32.5683						-114035	228073	228073	228080
$w(n;\phi) = n + p$	32.5683					1.77×10^8	-114035	228075	228075	228089
$w(n;\phi) = (n+a)(n-b)^2$	32.1584	4×10^8	32.4529				-97171.9	194360	194350	194371
$w(n;\phi) = (n+a)(n^2 - bn + c)$	32.0235	459.83	2.62997	3251.3			-113492	226993	226993	227021
$w(n;\phi) = a + \frac{b-ac}{n+c}$	33.5375	0	6512.19	2			-110837	221753	221753	221782
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	32.6023				1	0.998957	-114035	228077	228077	228098
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	Computationally intractable									
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	1373.38	1351.36	0.877599		0.887543		-39038.6	78085.3	78085.3	78113.6
$w(n;\phi) = Boole(n \leq b)$	32.5683		164				-114035	228075	228075	228089
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	32.5683					0	-114035	228075	228075	228089
$w(n;\phi) = (a)_n$	0.963976	1.2171					-39324.7	78653.4	78653.4	78667.6
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	26.28				10	0.632214	-91785.6	183575	183575	183589
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	33.5884	28961.2	9807.78	24560.5			-109867	219742	219742	219770
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	1617.54	0.414822	4604.42		2515.88		-39091	78190	78190	78208.3
$w(n;\phi) = (a)_n^{-1}$	1	0.00784581					-1.6243×10^6	3.2486×10^6	3.2486×10^6	3.24862×10^6

Table 10.14: Britain accidents - Incidents per hour

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	52.5414						-208595	417192	417192	417199
$w(n;\phi) = n + p$	52.5414					1.89×10^8	-208595	417194	417194	417208
$w(n;\phi) = (n+a)(n-b)^2$	51.5402	5.54799×10^8	51.5396				-187328	374661	374661	374682
$w(n;\phi) = (n+a)(n^2 - bn + c)$	Convergence									
$w(n;\phi) = a + \frac{b-ac}{n+c}$	34.9097	1	5.89619	1			-241216	482440	482440	482468
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	52.5414				1	1	-208.595	417196	417196	417217
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	Computationally intractable									
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	425.216	336.303	0.330653		0.237527		-64281.7	128571	128571	128600
$w(n;\phi) = Boole(n \leq b)$	52.5414		1697				-208595	417194	417194	417208
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	52.5414					0	-208595	417194	417194	417208
$w(n;\phi) = (a)_n$	0.981357	0.998156					-43546.5	87096.9	87096.9	87111.1
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	43.6181				17	0.636376	-162481	324968	324968	324989
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	53.5862	21354.7	0.00156815	2618.06			-199403	398814	398814	398842
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	Convergence									
$w(n;\phi) = (a)_n^{-1}$	Convergence									

Table 10.15: Britain accidents - Casualties per hour

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	68.7079						-246757	493515	493515	493522
$w(n;\phi) = n + p$	68.7079					1.82×10^8	-246757	493517	493517	493531
$w(n;\phi) = (n+a)(n-b)^2$	66.9842	5.5648×10^8	66.6207				-222827	445661	445661	445682
$w(n;\phi) = (n+a)(n^2 - bn + c)$	25.9007	0.273147	0.395746	5.49594			-448824	897656	897656	897684
$w(n;\phi) = a + \frac{b-ac}{n+c}$	69.7079	0.01	260075	1			-242086	484180	484180	484208
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	68.7971				1	0.998703	-246757	493519	493519	493541
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	Computationally intractable									
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	6250.74	6231.48	0.0710545		2.06448		-46141.6	92291.2	92291.2	92319.5
$w(n;\phi) = Boole(n \leq b)$	68.7079		1746				-246757	493517	493517	493531
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	68.7079					0	-246757	493517	493517	493531
$w(n;\phi) = (a)_n$	0.985755	0.992867					-45876.9	91757.8	91757.8	91772
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	55.789				20	0.636077	-190919	381845	381845	381866
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	69.8865	5182.54	0.000731412	1258.57			-237712	475431	475431	475460
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	19416.2	846.226	73964.1		26974.3		-45798.2	91604.3	91604.3	91632.6
$w(n;\phi) = (a)_n^{-1}$	1	0.00666583					-4.3551×10^6	8.71025×10^6	8.71025×10^6	8.71025×10^6

Table 10.16: Britain accidents - Vehicles per hour

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	1.61324						-423711	847424	847424	847434
$w(n) = n$	0.613239						-354265	708532	708532	708543
$w(n;\phi) = n^{-a}$	6.84661	5					-396959	793922	793922	793943
$w(n;\phi) = n + p$	0.613239					0	-354265	708534	708534	708555
$w(n;\phi) = an^3 + bn^2 + cn$	0.546157	0.0102276	0	0.310151			-351943	703.895	703.895	703.937
$w(n;\phi) = (n+a)(n-b)^2$	0.352571	2.1×10^7	0				-373093	746192	746192	746224
$w(n;\phi) = (n+a)(n^2 - bn + c)$	0.431924	0	3	3.87129			-340708	681423	681423	681465
$w(n;\phi) = a + \frac{b-ac}{n+c}$	0.861735	6.8774×10^6	0.0000011	1			-345790	691588	691588	691630
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	0.672225				4	1	-407945	815896	815896	815927
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	0.912061				47	0.0790988	-423968	847942	847942	847973
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	1.39196*10^7	2.36066×10^7					-299951	599907	599907	599928
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	4.50033	2	1				-307844	615694	615694	615725
$w(n;\phi) = \frac{-1}{ln(1-p)} \frac{p^n}{n}$	1.9144					0.913869	-324130	648264	648264	648285
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	318.43	4866.42	6.0371		6.03686		-401359	802726	802726	802769
$w(n;\phi) = Boole(n \geq a)$	1.04702	1					-377793	675589	675589	675611
$w(n;\phi) = Boole(n \leq b)$	1.61324		38				-423711	847426	847426	847447
$w(n;\phi) = Boole(n \geq a)\, Boole(n \leq b)$	1.04702	1	38				-337793	675591	675591	675623
$w(n;\phi) = \binom{n}{a}$	0.613239	1					-354.265	708534	708534	708555
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	Convergence									
$w(n;\phi) = (a)_n$	0.0583323	26.0428					-422526	845057	845057	845078
$w(n;\phi) = (n)_a$	1.04702	0					-337793	675589	675589	675611
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	0.67225				4	1	-407945	815896	815896	815927
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	1.02575	0.576635	0.0651226	0.001			-340753	681514	681514	681557
$w(n;\phi) = \left(\frac{-1}{ln(1-p)} \frac{p^n}{n}\right)^{-1}$	0.513662					0.837622	-354362	708534	708534	708555
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	0.613238	0					-354265	708534	708534	708555
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	0.631434	0.01	0.01		107.51		-356436	712881	712881	712923
$w(n;\phi) = \binom{n}{a}^{-1}$	1.74951	1					-324130	648264	648264	648285
$w(n;\phi) = (a)_n^{-1}$	1	0.001					-409889	819782	819782	819803
$w(n;\phi) = (n)_a^{-1}$	1.36191×10^7	2.30963×10^7					-299951	599907	599907	599928

Table 10.17: Britain accidents -Casualties per incident

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	2.10965						-416619	833239	833239	833250
$w(n) = n$	1.10965						-354893	709789	709789	709800
$w(n;\phi) = n^{-a}$	1.73899	1					-564853	1.12971×10^6	1.12971×10^6	1.12973×10^6
$w(n;\phi) = n + p$	1.10965					0	-354893	709791	709791	709812
$w(n;\phi) = an^3 + bn^2 + cn$	0.43485	75557	1.8425	2.09252			-334040	668089	668089	668131
$w(n;\phi) = (n+a)(n-b)^2$	0.229596	0	0.655111				-321998	644001	644001	644033
$w(n;\phi) = (n+a)(n^2 - bn + c)$	0.231486	0	1.29826	0.42137			-321992	643992	643992	644034
$w(n;\phi) = a + \frac{b-ac}{n+c}$	Convergence									
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	0.300066				17	1	-369313	738631	738631	738663
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	0.501504				45	0.20713	-365564	731134	731134	731166
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	2.67996	0					-394801	789606	7890606	789628
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	4.04092	1	1				-424314	848633	848633	848654
$w(n;\phi) = \frac{-1}{ln(1-p)}\frac{p^n}{n}$	3.18041					0.842364	-394785	789575	789575	789596
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	0.049853	290.15	329.59		320.36		-366356	732698	732698	732740
$w(n;\phi) = Boole(n \geq a)$	1.73899	1					-372034	744072	744072	744093
$w(n;\phi) = Boole(n \leq b)$	2.10965		37				-416619	833241	833241	833263
$w(n;\phi) = Boole(n \geq a)\, Boole(n \leq b)$	1.73899						-372134	744072	744072	744093
$w(n;\phi) = \binom{n}{a}$	1.10965	1					-354893	709789	709789	709800
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	Convergence									
$w(n;\phi) = (a)_n$	0.0134907	154.184					-417744	835493	835493	835514
$w(n;\phi) = (n)_a$	0.0528901	58.0781					-341853	683709	683709	683730
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	2.94252				2	0.998721	-445083	890172	890172	890204
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	1.13961	0.000108859	0.00456108	0100144577			-355062	710131	710131	10173
$w(n;\phi) = \left(\frac{-1}{ln(1-p)}\frac{p^n}{n}\right)^{-1}$	0.959759					0.864922	-354893	709791	709791	709812
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	0.0528901	57.0781					-341853	683709	683709	683730
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	0.122043	14.3606	0		0		-335465	670938	670938	670980
$w(n;\phi) = \binom{n}{a}^{-1}$	2.67907	1					-394785	789575	789575	789596
$w(n;\phi) = (a)_n^{-1}$	1	0					-468025	936053	936053	936074
$w(n;\phi) = (n)_a^{-1}$	1.74	0					-372067	744138	744138	744160

Table 10.18: Britain accidents - Vehicles per incident

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	2.09274						-465223	930447	930447	930458
$w(n) = n$	1.09274						-408602	817206	817206	817217
$w(n;\phi) = n^{-a}$	1.7168	1					-607832	1.21567×10^6	1.21567×10^6	1.21569×10^6
$w(n;\phi) = n + p$	1.09274					0	-408602	817208	817208	817229
$w(n;\phi) = an^3 + bn^2 + cn$	0.425756	3.30051	0	0			-404136	808281	808281	808323
$w(n;\phi) = (n+a)(n-b)^2$	0.340524	0	0.324713				-403086	806179	806179	806211
$w(n;\phi) = (n+a)(n^2 - bn + c)$	0.340524	0	0.649426	0.105438			-403086	806181	806181	806223
$w(n;\phi) = a + \frac{b-ac}{n+c}$	1.22966	3020.8	0.00100023	6			-408137	816282	816282	816324
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	0.744359				8	0.763716	-436292	872589	872589	872621
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	0.953299					0.0768837	-449553	899113	899113	899144
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	2.65186	0.0001			73		-432653	865309	865309	865330
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	4.00689	1	1				-454507	909019	909019	909051
$w(n;\phi) = \frac{-1}{ln(1-p)}\frac{p^n}{n}$	3.05935					0.866514	-432642	865289	865289	865310
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	12.3477	363.762	11.7833		11.0575		-430720	861448	861448	861490
$w(n;\phi) = Boole(n \geq a)$	1.7168	1					-417739	835483	835483	835504
$w(n;\phi) = Boole(n \leq b)$	2.09274		57				-465223	930449	930449	930471
$w(n;\phi) = Boole(n \geq a)\, Boole(n \leq b)$	1.7168	1	57				-417739	835485	835485	835516
$w(n;\phi) = \binom{n}{a}$	1.09274	1					-408602	817208	817208	817229
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	Convergence									
$w(n;\phi) = (a)_n$	0.0413737	48.4887					-464182	928368	928368	928389
$w(n;\phi) = (n)_a$	0.829444	1.802					-407633	815271	815271	815292
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	2.92712				2	1	-490.204	980414	980414	980446
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	1.23285	0.000142232	0.000849372	0.00125779			-408271	816550	816550	816592
$w(n;\phi) = \left(\frac{-1}{ln(1-p)}\frac{p^n}{n}\right)^{-1}$	0.74982					0.64515	-408602	817208	817208	817229
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	0.829444	0.890198					-407633	815271	815271	815292
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	0.685721	1.01525	0		0		-404445	808897	808897	808940
$w(n;\phi) = \binom{n}{a}^{-1}$	2.65097	1					-432642	865289	865289	865310
$w(n;\phi) = (a)_n^{-1}$	1	0					-553070	1.10614×10^6	1.10614×10^6	1.10617×10^6
$w(n;\phi) = (n)_a^{-1}$	1.71781	0					-417761	835527	835527	835548

Table 10.19: Canada accidents - Vehicles per incident

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	146.61						-91412.1	182826	182826	182832
$w(n) = n$	145.61						-92198.7	184399	184399	184405
$w(n;\phi) = n^{-a}$	178.83	30					-77888.2	155780	155780	155792
$w(n;\phi) = n + p$	146.61					4.68659×10^7	-91412.1	182828	182828	182839
$w(n;\phi) = an^3 + bn^2 + cn$	145.103	0.310127	0	19328.9			-92083.5	184175	184175	174197
$w(n;\phi) = (n+a)(n-b)^2$	148.777	248121	150.381				-84752.7	169511	169511	169528
$w(n;\phi) = (n+a)(n^2 - bn + c)$	146.417	223376	384.375	115549			-91158.2	182324	182324	172347
$w(n;\phi) = a + \frac{b-ac}{n+c}$	147.603	0	544.899	2			-90669.5	181347	181347	181369
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	38.3317				369	1	-153324	306654	306654	306671
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	Computationally intractable									
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	9737.73	9512.07					-12187.7	24379.3	24379.3	24390.6
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	Computationally intractable									
$w(n;\phi) = \frac{-1}{ln(1-p)}\frac{p^n}{n}$	238.734					0.618332	-90625.6	181255	181255	181266
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	169.421	41.7711	0		0		-71060.1	142128	142128	142151
$w(n;\phi) = Boole(n \geq a)$	146.61	2					-91412.1	182828	182828	182839
$w(n;\phi) = Boole(n \leq b)$	146.61		631				-91412.1	182828	182828	182839
$w(n;\phi) = Boole(n \geq a)\, Boole(n \leq b)$	146.61	2	631				-91412.1	182830	182830	182847
$w(n;\phi) = \binom{n}{a}$	144.61	2					-93010.6	186025	186025	186036
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	146.61					0	-91412.1	182828	182828	182839
$w(n;\phi) = (a)_n$	0.990283	1.43859					-12007	24017.9	24017.9	24029.1
$w(n;\phi) = (n)_a$	146.61	0					-91412.1	182828	182828	182839
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	91.2411				2	0.618122	-90648.5	181303	181303	181320
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	147.61	712.629	0	20867			-90625.6	181259	181259	181282
$w(n;\phi) = \left(\frac{-1}{ln(1-p)}\frac{p^n}{n}\right)^{-1}$	144.121					0.989772	-92198.7	184401	184401	184413
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	145.61	0					-92198.7	184401	184401	184413
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	5434.28	508.108	32193.7		6508.9		-12633.4	25274.9	25274.9	25297.3
$w(n;\phi) = \binom{n}{a}^{-1}$	146.61	0					-91412.1	182828	182828	182839
$w(n;\phi) = (a)_n^{-1}$	1	0					-2.53181×10^6	5.06363×10^6	5.06363×10^6	5.06364×10^6
$w(n;\phi) = (n)_a^{-1}$	9737.73	951307					-12187.7	24379.3	24379.3	24390.6

Table 10.20: Canada accidents - Incidents per hour

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	306.945						-235594	471190	471190	471196
$w(n) = n$	305.945						-236591	473185	473185	473190
$w(n;\phi) = n^{-a}$	338.521	31					-216226	432455	432455	432467
$w(n;\phi) = n + p$	306.945					4.4638×10^7	-235594	471192	471192	471204
$w(n;\phi) = an^3 + bn^2 + cn$	304.961	0.00190825	54.8669	0.942041			-237589	475185	475185	475207
$w(n;\phi) = (n+a)(n-b)^2$	305.726	1.37838×10^8	305.616				-227098	454201	454201	454218
$w(n;\phi) = (n+a)(n^2 - bn + c)$	305.357	641366	1.03581	22967.8			-236042	472092	472092	472115
$w(n;\phi) = a + \frac{b-ac}{n+c}$	226.519	17.6518	0.0345541	2			-261500	523007	523007	523030
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	80.3893				363	1	-415310	830625	830625	830642
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	Computationally intractable									
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	79952.4	79386.3					-14099.5	28203.1	28203.1	28214.3
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	Computationally intractable									
$w(n;\phi) = \frac{-1}{ln(1-p)}\frac{p^n}{n}$	321.901					0.956656	-234597	469198	469198	469209
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	850.649	710.397	0		97.303		-132491	264991	264991	264013
$w(n;\phi) = Boole(n \geq a)$	306.945	2					-235594	471192	471192	471204
$w(n;\phi) = Boole(n \leq b)$	306.945		4591				-235594	471192	471192	471204
$w(n;\phi) = Boole(n \geq a)\, Boole(n \leq b)$	306.945	2	4591				-235594	471194	471194	471211
$w(n;\phi) = \binom{n}{a}$	304945	2					-237607	475218	475218	475229
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	306.945					0	-235594	471192	471192	471204
$w(n;\phi) = (a)_n$	0.996256	1.15352					-13552	27108	27108	27119.2
$w(n;\phi) = (n)_a$	306.945	0					-235594	471192	471192	471204
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	24.7612				2	0.0804078	-234614	469233	469233	469250
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	307.948	50272.4	0.00646671	549978			-234597	469202	469202	496224
$w(n;\phi) = \left(\frac{-1}{ln(1-p)}\frac{p^n}{n}\right)^{-1}$	07988					0.00261092	-236591	473187	473187	473198
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	305.945	0					-236591	473187	473187	473198
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	198014	169617	800252		319489		-13595	27197.9	18198	27220.4
$w(n;\phi) = \binom{n}{a}^{-1}$	80570.8	80004.6					-14099.6	28203.1	28203.1	28214.3
$w(n;\phi) = (a)_n^{-1}$	1	0					-6.31134×10^6	1.26227×10^7	1.26227×10^7	1.26227×10^7
$w(n;\phi) = (n)_a^{-1}$	80570.8	80004.6					-14099.6	28203.1	28203.1	28214.3

Table 10.21: Canada accidents - Vehicles per hour

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	0.490491						-218606	437213	437213	437224
$w(n;\phi) = n + p$	0.303192					1.31557	-218443	436891	436891	436912
$w(n;\phi) = (n+a)(n-b)^2$	0.0789986	3.22185×10^7	0.274108				-221709	443425	443425	443456
$w(n;\phi) = (n+a)(n^2 - bn + c)$	0.279406	1.03389	1.03389	40.8218			-218272	436553	436553	436594
$w(n;\phi) = a + \frac{b-ac}{n+c}$	Convergence									
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	0.274879				2	1	-218489	436962	436962	436993
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	0.157002				61	0.0627272	-221709	443423	443423	443455
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	6.93968	123.321	3.1556		3.07126		-217765	435538	435538	435579
$w(n;\phi) = Boole(n \leq b)$	0.490491		53				-218606	437215	437215	437236
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	Convergence									
$w(n;\phi) = (a)_n$	0.0200181	24.0119					-218549	437102	437102	437123
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	0.351081	0	0.657169		114.61		-281257	436523	436523	436564
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	0.394499	1.47576	0.101053	0.0478654			-217944	435896	435896	435938
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	0.351081	0	0.657169		114.61		-218257	436523	436523	436564
$w(n;\phi) = (a)_n^{-1}$	1	1.74005					-220077	440158	440158	440178

Table 10.22: US gun violence - Injured per incident

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	0.251258						-148452	296905	296905	296915
$w(n;\phi) = n + p$	0.251258					1.86×10^7	-148452	296907	296907	296928
$w(n;\phi) = (n+a)(n-b)^2$	0.695869	7.13765×10^6	2.69528				-148961	297928	297928	297960
$w(n;\phi) = (n+a)(n^2 - bn + c)$	0.233164	6.59339	2.00731	28.0876			-148454	296917	296917	296958
$w(n;\phi) = a + \frac{b-ac}{n+c}$	Convergence									
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	0.251258				1	1	-148452	296909	296909	296940
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	0.581466				54	0.0090356	-149964	299934	299934	299966
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	78.1346	1892.53	2.47006		2.47007		-148240	296488	296488	296529
$w(n;\phi) = Boole(n \leq b)$	0.251258		50				-148452	296907	296907	296928
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	0.251258					0	-148452	296907	296907	296928
$w(n;\phi) = (a)_n$	0.0238964	10.2632					-148376	296756	296756	296777
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	0.465925				2	0.998996	-148743	297493	297493	297524
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	0.237411	0.137768	0.216916	1.33844			-148447	296902	296902	296943
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	0.132467	17.6383	1.03718		19.2308		-148484	296977	296977	297018
$w(n;\phi) = (a)_n^{-1}$	1	3.80208					-148792	297592	297592	297613

Table 10.23: US gun violence - Casualties per incident

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	38.8572						-5986.62	11975.2	11975.3	11980.6
$w(n) = n$	37.8572						-6012.05	12026.1	12026.1	12031.4
$w(n;\phi) = n^{-a}$	51.9448	12					-11377.8	22757.6	22757.6	22762.9
$w(n;\phi) = n + p$	38.8572					2.80937×10^7	-5986.62	11977.2	11977.3	11987.9
$w(n;\phi) = an^3 + bn^2 + cn$	37.401	0.0553273	0	269.192			-6006.21	12020.4	12020.4	12041.8
$w(n;\phi) = (n+a)(n-b)^2$	45.0877	2.58267×10^7	52.5344				-6341.24	12688.5	12688.5	12704.5
$w(n;\phi) = (n+a)(n^2 - bn + c)$	38.3367	686.626	0.00246164	5015.97			-5980.99	11970	11970	11991.4
$w(n;\phi) = a + \frac{b-ac}{n+c}$	38.9608	1	1.0178	1			-5986.81	11981.6	11981.7	12003
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	30.3448				12	0.999049	-6211.98	12430	12430	12446
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	4.41476				119	0.811073	-7907.33	15820.7	15820.7	15836.7
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	82.8808	41.8087					-5663.38	11330.8	11330.8	11341.5
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	40.9244	1	15				-5936.93	11879.9	11879.9	11895.9
$w(n;\phi) = \frac{-1}{ln(1-p)}\frac{p^n}{n}$	62.6276					0.636851	-5961.76	11927.5	11927.5	11938.2
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	87.1396	48.2218	0		3.08924		-5660.48	11329	11329	11350.3
$w(n;\phi) = Boole(n \geq a)$	38.8572	15					-5986.62	11977.2	11977.3	11987.9
$w(n;\phi) = Boole(n \leq b)$	38.8572		93				-5986.62	11977.2	11977.3	11987.9
$w(n;\phi) = Boole(n \geq a)\, Boole(n \leq b)$	38.8572	15	93				-5986.62	11979.20	11979.3	11995.3
$w(n;\phi) = \binom{n}{a}$	29.857	9					-6308.69	12621.4	12621.4	12632.1
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	38.8572					0	-5986.62	11977.2	11977.3	11987.9
$w(n;\phi) = (a)_n$	0.557367	30.8585					-5635.72	11275.4	11275.5	11286.1
$w(n;\phi) = (n)_a$	38.8572	0					-5986.62	11977.2	11977.3	11987.9
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	51.8572				14	1	-5779.52	11565	11565	11581.1
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	39.8844	1.00025×10^7	13741.1	3.59431×10^6			-5961.76	11931.5	11931.6	11952.9
$w(n;\phi) = \left(\frac{-1}{ln(1-p)}\frac{p^n}{n}\right)^{-1}$	0.786974					0.0007879	-6012.05	12028.1	12028.1	12038.8
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	37.8573	0					-6012.05	12028.1	12028.1	12038.8
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	80.3881	72.5824	1465.69		49.0207		-5660.92	11329.8	11329.9	11351.2
$w(n;\phi) = \binom{n}{a}^{-1}$	38.8572	0					-5986.62	11977.2	11977.3	11987.9
$w(n;\phi) = (a)_n^{-1}$	1	0					-327069	654142	654142	654153
$w(n;\phi) = (n)_a^{-1}$	82.8808	42.8087					-5663.38	11330.8	11330.8	11341.5

Table 10.24: US gun violence - Fatalities per day

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	75.8547						-9603.64	19209.3	19209.3	19214.6
$w(n) = n$	74.8547						-9659.06	19320.1	19320.1	19325.5
$w(n;\phi) = n^{-a}$	99.3405	21					-15175.8	30355.6	30355.6	30366.3
$w(n;\phi) = n + p$	75.8546					1.5739×10^7	-9603.64	19211.3	19211.3	19222
$w(n;\phi) = an^3 + bn^2 + cn$	74.3854	0.0382021	0	699.054			-9646.73	19301.5	19301.5	19322.8
$w(n;\phi) = (n+a)(n-b)^2$	84.564	3.1444×10^7	94.3097				-9395.99	18798	18798	18814
$w(n;\phi) = (n+a)(n^2 - bn + c)$	75.3398	44471.6	1.35605	15834.7			-9591.25	19190.5	19190.5	19211.9
$w(n;\phi) = a + \frac{b-ac}{n+c}$	76.8412	0	275.884	2			-9551.42	19110.8	19110.9	19132.2
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	29.6486				120	0.999028	-12170.5	24346.9	24346.9	24363
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	Computationally intractable									
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	430.585	348.465					-6990.39	13984.8	13984.8	13995.5
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	60.0199	1					-12066.7	24137.5	24137.5	24148.1
$w(n;\phi) = \frac{-1}{ln(1-p)}\frac{p^n}{n}$	119.904					0.641081	-9548.37	19100.7	19100.7	19111.4
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	97.6517	30.8128	39.274		0.0849167		-8996.44	18000.9	18000.9	18022.3
$w(n;\phi) = Boole(n \geq a)$	75.8547	17					-9603.64	19211.3	19211.3	19222
$w(n;\phi) = Boole(n \leq b)$	75.8547		179				-9603.64	19211.3	19211.3	19222
$w(n;\phi) = Boole(n \geq a)\, Boole(n \leq b)$	75.8547	17	179				-9603.64	19213.3	19213.3	19229.3
$w(n;\phi) = \binom{n}{a}$	58.8547	17					-10873.1	21750.1	21750.1	21760.8
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	75.8547					0	-9603.64	19211.3	19211.3	19222
$w(n;\phi) = (a)_n$	0.841052	14.3355					-6933.12	13870.2	13870.2	13880.9
$w(n;\phi) = (n)_a$	75.8547	0					-9603.64	19211.3	19211.3	19222
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	82.771				8	0.998991	-9259.34	18524.7	18524.7	18540.7
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	76.8635	1748.86	0.000107766	997.53			-9548.69	19105.4	19105.4	19126.8
$w(n;\phi) = \left(\frac{-1}{ln(1-p)}\frac{p^n}{n}\right)^{-1}$	0.724281					0.00967583	-9659.06	19322.1	19322.1	19332.8
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	74.8546	0					-9659.06	19322.1	19322.1	19332.8
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	400.028	6.54918	1505.67		437.535		-6980	13968	13968	13989.4
$w(n;\phi) = \binom{n}{a}^{-1}$	75.8547	0					-9603.64	19211.3	19211.3	19222
$w(n;\phi) = (a)_n^{-1}$	1	0					-795058	1.59012×10^6	1.59012×10^6	1.59013×10^6
$w(n;\phi) = (n)_a^{-1}$	430.585	349465					-6990.39	13984.8	13984.8	13995.5

Table 10.25: US gun violence - Injuries per day

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n)=1$	1.53145						-8907.89	17817.8	17817.8	17824.4
$w(n;\phi)=n+p$	1.53145					3.99959×10^{7}	-8907.89	17819.8	17819.8	17833.1
$w(n;\phi)=(n+a)(n-b)^{2}$	0.431881	6.01615×10^{7}	0.378461				-9284.08	18574.2	18574.2	18594.1
$w(n;\phi)=(n+a)(n^{2}-bn+c)$	1.45965	4.59652	7.07963	34.0248			-8896.86	17801.7	17801.1	17828.3
$w(n;\phi)=a+\frac{b-ac}{n+c}$	1.66616	2.91371	4.60381	1			-8899.63	17807.3	17807.3	18833.8
$w(n;\phi)=\binom{n+r-1}{n}p^{n}(1-p)^{r}$	1.53299				1	0.999001	-8907.89	17821.8	17821.8	17841.7
$w(n;\phi)=\binom{r}{n}p^{n}(1-p)^{r-n}$	0.471547				20	0.270272	-9613.4	19232.8	19232.8	19252.7
$w(n;\phi)=\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	8.87021	57.2139	10.0619		1.08593		-8896.41	17800.8	17800.8	17827.4
$w(n;\phi)=Boole(n\leq b)$	1.5315	9					-8907.87	17819.7	17819.7	17833
$w(n;\phi)=p\frac{n!}{e^{-\lambda}\lambda^{n}}Boole(n=0)+(1-p)$	1.53145					0	-8907.89	17819.8	17819.8	17833.1
$w(n;\phi)=(a)_{n}$	0.084907	16.5054					-8896.4	17796.8	17796.8	17810.1
$w(n;\phi)=\left(\binom{n+r-1}{n}p^{n}(1-p)^{r}\right)^{-1}$	2.26763				2	0.998999	-8924.23	17854.5	17854.5	17874.4
$w(n;\phi)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$	1.94461	71.9657	176.371	824.332			-8898.06	17804.1	17804.1	17830.7
$w(n;\phi)=\left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	1.47692	2.94988	18.953		1.33312		-8899.65	17807.3	17807.3	17833.9
$w(n;\phi)=(a)_{n}^{-1}$	1	0.279681					-11103.5	22211	22211	22224.3

Table 10.26: EPL games - Home team goals per game

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n)=1$	1.12972						-7923.45	15848.9	15848.9	15855.5
$w(n;\phi)=n+p$	1.12972					3.3176×10^{7}	7923.45	15850.9	15850.9	15864.2
$w(n;\phi)=(n+a)(n-b)^{2}$	0.291115	4.50759×10^{7}	0.364134				-8156.11	16318.2	16318.2	16338.1
$w(n;\phi)=(n+a)(n^{2}-bn+c)$	0.820105	2.34584	2.55243	10.7358			-7906.58	15821.2	15821.2	15847.7
$w(n;\phi)=a+\frac{b-ac}{n+c}$	1.33036	1.20031	2.55621	1			-7907.6	15823.2	15823.2	15849.7
$w(n;\phi)=\binom{n+r-1}{n}p^{n}(1-p)^{r}$	1.12972				1	1	-7923.45	15852.9	15852.9	15872.8
$w(n;\phi)=\binom{r}{n}p^{n}(1-p)^{r-n}$	0.416032				17	0.230515	-8498.16	17002.3	17002.3	17022.2
$w(n;\phi)=\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	1.68843	0.41548	1.35939		1.35939		-7906.92	15821.8	15821.8	15848.4
$w(n;\phi)=Boole(n\leq b)$	1.13078	6					-7922.46	15848.9	15848.9	15862.2
$w(n;\phi)=p\frac{n!}{e^{-\lambda}\lambda^{n}}Boole(n=0)+(1-p)$	1.12972					0	-7923.45	15850.9	15850.9	15864.2
$w(n;\phi)=(a)_{n}$	0.101221	10.0312					-7907.71	15819.4	15819.4	15832.7
$w(n;\phi)=\left(\binom{n+r-1}{n}p^{n}(1-p)^{r}\right)^{-1}$	1.76522				2	1	-7917.19	15840.4	15840.4	15860.3
$w(n;\phi)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$	1.56653	10.2208	17.3917	292.49			-7907	15822	15822	15848.5
$w(n;\phi)=\left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	1.39057	0	54.6125		1.45495		-7907.33	15822.7	15822.7	15849.2
$w(n;\phi)=(a)_{n}^{-1}$	1	0.531647					-8905.6	17815.2	17815.2	17828.5

Table 10.27: EPL games - Away team goals per game

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n)=1$	6.62786						-15249.1	30500.1	30500.1	30506.7
$w(n;\phi)=n+p$	6.62786					7.677×10^{7}	-15249.1	30502.1	30502.1	30515.4
$w(n;\phi)=(n+a)(n-b)^{2}$	3.9235	0.115532	4.48468				-17391.3	34788.6	34788.6	34808.5
$w(n;\phi)=(n+a)(n^{2}-bn+c)$	5.72877	6.6459	9.6459	66.1057			-15019	30046	30046	30072.6
$w(n;\phi)=a+\frac{b-ac}{n+c}$	7.62413	0	75039	1			-15008.2	30024.4	30024.4	30051
$w(n;\phi)=\binom{n+r-1}{n}p^{n}(1-p)^{r}$	6.62786				1	1	-15249.1	30504.1	30504.1	30524
$w(n;\phi)=\binom{r}{n}p^{n}(1-p)^{r-n}$	Computationally intractable									
$w(n;\phi)=\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	54.9179	87.5138	1.64204		4.55045		-14684.6	29377.3	29377.3	29403.8
$w(n;\phi)=Boole(n\leq b)$	6.62786		24				-15249.1	30502.1	30502.1	30515.4
$w(n;\phi)=p\frac{n!}{e^{-\lambda}\lambda^{n}}Boole(n=0)+(1-p)$	6.62786					0	-15249.1	30502.1	30502.1	50515.4
$w(n;\phi)=(a)_{n}$	0.440041	8.43405					-14685.5	29375.1	29375.1	29388.3
$w(n;\phi)=\left(\binom{n+r-1}{n}p^{n}(1-p)^{r}\right)^{-1}$	11.5442				6	1	-14760.6	29527.1	29527.1	29547
$w(n;\phi)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$	7.81974	8484.8	2241.44	28706.6			-14957.3	29922.7	29922.7	29949.2
$w(n;\phi)=\left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	11.6702	0	591.779		6.25691		-14760.2	29528.3	29528.3	29554.9
$w(n;\phi)=(a)_{n}^{-1}$	1	0.00241219					-90217.6	180439	180439	180452

Table 10.28: EPL games - Home team shots on target per game

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	5.07762						-14115.1	28232.2	28232.2	28238.9
$w(n;\phi) = n + p$	5.07762					7.326×10^7	-14115.1	28234.2	28234.2	28247.5
$w(n;\phi) = (n+a)(n-b)^2$	2.26153	0.235644	2.49518				-16295.7	32597.5	32597.5	32617.4
$w(n;\phi) = (n+a)(n^2 - bn + c)$	4.45409	31.205	2.69845	54.7052			-14015.6	28039.3	28039.3	28065.8
$w(n;\phi) = a + \frac{b-ac}{n+c}$	6.06348	0	82621.6	1			-13889.3	27786.6	27786.6	27813.1
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	5.07762				1	1	-14115.1	28236.2	28236.2	28256.1
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	0.900728				36	0.514681	-16518.1	33042.2	33042.2	33062.1
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	48.3317	84.2263	3.68949		1.79638		-13680	27368	27368	27394.5
$w(n;\phi) = Boole(n \leq b)$	5.07762		20				-14115.1	28234.2	28234.2	28247.5
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	5.07762					0	-14115.1	28234.2	28234.2	28247.5
$w(n;\phi) = (a)_n$	0.401701	7.56268					-13681.7	27367.4	27367.4	27380.7
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	7.02665				3	0.998999	-13794.3	27594.7	27594.7	27614.6
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	6.26389	156.375	65.2648	8586.57			-13854.4	27716.9	27716.9	27743.4
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	8.56664	0	537.852		4.68777		-13753.7	27515.5	27515.5	27542
$w(n;\phi) = (a)_n^{-1}$	1	0.0089961					-59095	118194	188194	118207

Table 10.29: EPL games - Away team shots on target per game

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	16.0717						-2.37739×10^7	4.75478×10^7	4.75478×10^7	4.75478×10^7
$w(n) = n$	15.0717						-2.41985×10^7	4.8397×10^7	4.8397×10^7	4.8397×10^7
$w(n;\phi) = n^{-a}$	24.5102	7					-3.72876×10^7	7.45753×10^7	7.45753×10^7	7.45753×10^7
$w(n;\phi) = n + p$	16.0717					8.78091×10^8	-2.37739×10^7	4.75478×10^7	4.75478×10^7	4.75478×10^7
$w(n;\phi) = an^3 + bn^2 + cn$	14.7735	0.788063	0	1063.35			-2.41218×10^7	4.82436×10^7	4.82436×10^7	4.82436×10^7
$w(n;\phi) = (n+a)(n-b)^2$	19.0517	1.39386×10^8	21.5172				-2.20494×10^7	4.40989×10^7	4.40989×10^7	4.40989×10^7
$w(n;\phi) = (n+a)(n^2 - bn + c)$	15.7465	527.513	3.52896	1336.59			-2.36873×10^7	4.73746×10^7	4.73746×10^7	4.73747×10^7
$w(n;\phi) = a + \frac{b-ac}{n+c}$	17.0717	0.001	86236.4	1			-2.3408×10^7	4.68161×10^7	4.68161×10^7	4.68161×10^7
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	16.0908				1	0.998808	-2.37739×10^7	4.75478×10^7	4.75478×10^7	4.75478×10^7
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	Computationally intractable									
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	53.2362	33.3994					-2.05635×10^7	4.11269×10^7	4.11269×10^7	4.11269×10^7
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	22.5027	4					-2.20171×10^7	4.40342×10^7	4.40342×10^7	4.40342×10^7
$w(n;\phi) = \frac{-1}{ln(1-p)} \frac{p^n}{n}$	24.6641					0.695107	-2.33636×10^7	4.67272×10^7	4.67272×10^7	4.67272×10^7
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	1884.32	2564.32	3.27899		3.34795		-1.92024×10^7	3.84048×10^7	3.84048×10^7	3.84049×10^7
$w(n;\phi) = Boole(n \geq a)$	16.0717	1					-2.37739×10^7	4.75478×10^7	4.75478×10^7	4.75478×10^7
$w(n;\phi) = Boole(n \leq b)$	16.0717		225				-2.37739×10^7	4.75478×10^7	4.75478×10^7	4.75478×10^7
$w(n;\phi) = Boole(n \geq a)\, Boole(n \leq b)$	16.0717	1	225				-2.37739×10^7	4.75478×10^7	4.75478×10^7	4.75478×10^7
$w(n;\phi) = \binom{n}{a}$	15.0717	1					-2.41985×10^7	4.8397×10^7	4.8397×10^7	4.8397×10^7
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	16.0717					0	-2.37739×10^7	4.75478×10^7	4.75478×10^7	4.75478×10^7
$w(n;\phi) = (a)_n$	0.713479	6.45412					-1.92327×10^7	3.84655×10^7	3.84655×10^7	3.84655×10^7
$w(n;\phi) = (n)_a$	16.0717	0					-2.37739×10^7	4.75478×10^7	4.75478×10^7	4.75478×10^7
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	18.0487				3	0.998728	-2.30899×10^7	4.61797×10^7	4.61797×10^7	4.61798×10^7
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	17.1445	1408.35	11.5715	93017.2			-2.33644×10^7	4.67288×10^7	4.67288×10^7	4.67288×10^7
$w(n;\phi) = \left(\frac{-1}{ln(1-p)} \frac{p^n}{n}\right)^{-1}$	14.9238					0.990187	-2.41985×10^7	4.8397×10^7	4.8397×10^7	4.8397×10^7
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	15.0717	0					-2.41985×10^7	4.8397×10^7	4.8397×10^7	4.8397×10^7
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	64.5809	361.343	2244.12		62.2009		-2.02037×10^7	4.04073×10^7	4.04073×10^7	4.04074×10^7
$w(n;\phi) = \binom{n}{a}^{-1}$	16.0717	0					-2.37739×10^7	4.75478×10^7	4.75478×10^7	4.75478×10^7
$w(n;\phi) = (a)_n^{-1}$	1	0.001					-3.62224×10^8	4.24447×10^8	4.24447×10^8	4.24447×10^8
$w(n;\phi) = (n)_a^{-1}$	53.2362	34.3994					-2.05635×10^7	4.11269×10^7	4.11269×10^7	4.11269×10^7

Table 10.30: US flights - Departure taxi time

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	7.43497						-1.8776×10^7	3.75521×10^7	3.75521×10^7	3.75521×10^7
$w(n) = n$	6.43497						-1.93262×10^7	3.86523×10^7	3.86523×10^7	3.86523×10^7
$w(n;\phi) = n^{-a}$	11.8344	4					-2.84427×10^7	5.68853×10^7	5.68853×10^7	5.68854×10^7
$w(n;\phi) = n+p$	7.43497					5.88967×10^8	-1.8776×10^7	3.75521×10^7	3.75521×10^7	3.75521×10^7
$w(n;\phi) = an^3+bn^2+cn$	6.2197	0.210416	0	81.5885			-1.92361×10^7	3.84723×10^7	3.84723×10^7	3.84723×10^7
$w(n;\phi) = (n+a)(n-b)^2$	9.55213	8.23308×10^7	11.5555				-1.74343×10^7	3.48687×10^7	3.48687×10^7	3.48687×10^7
$w(n;\phi) = (n+a)(n^2-bn+c)$	7.10275	104.646	3.77721	293.825			-1.8665×10^7	3.7330×10^7	3.7330×10^7	3.7331×10^7
$w(n;\phi) = a+\frac{b-ac}{n+c}$	8.43314	0	668021	1			-1.83903×10^7	3.67807×10^7	3.67807×10^7	3.67807×10^7
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	7.42497				1	1	-1.8776×10^7	3.75521×10^7	3.75521×10^7	3.75521×10^7
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	Computationally intractable									
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	23.4717	13.2667					-1.72623×10^7	3.45246×10^7	3.45246×10^7	3.45246×10^7
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	10.0541	1	1				-1.79858×10^7	3.59716×10^7	3.59716×10^7	3.59716×10^7
$w(n;\phi) = \frac{-1}{ln(1-p)}\frac{p^n}{n}$	10.5532					0.816398	-1.83072×10^7	3.66144×10^7	3.66144×10^7	3.66144×10^7
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	1019.78	1601.01	2.47317		2.51019		-1.58951×10^7	3.17902×10^7	3.17902×10^7	3.17902×10^7
$w(n;\phi) = Boole(n \geq a)$	7.43056	1					-1.87726×10^7	3.75453×10^7	3.75453×10^7	3.75453×10^7
$w(n;\phi) = Boole(n \leq b)$	7.43497		248				-1.8776×10^7	3.75521×10^7	3.75521×10^7	3.75521×10^7
$w(n;\phi) = Boole(n \geq a)\,Boole(n \leq b)$	7.43056	1	248				-1.87726×10^7	3.75453×10^7	3.75453×10^7	3.75453×10^7
$w(n;\phi) = \binom{n}{a}$	7.43497	0					-1.8776×10^7	3.75521×10^7	3.75521×10^7	3.75521×10^7
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n}Boole(n=0)+(1-p)$	7.43497					0	-1.8776×10^7	3.75521×10^7	3.75521×10^7	3.75521×10^7
$w(n;\phi) = (a)_n$	0.620404	4.54911					-1.59678×10^7	3.19356×10^7	3.19356×10^7	3.19356×10^7
$w(n;\phi) = (n)_a$	7.43156	0					-1.87726×10^7	3.75453×10^7	3.75453×10^7	3.75453×10^7
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	9.41572				3	0.998718	-1.81077×10^7	3.62155×10^7	3.62155×10^7	3.62155×10^7
$w(n;\phi) = \left(a+\frac{b-ac}{n+c}\right)^{-1}$	8.57604	970.704	307.543	78764.4			-1.83567×10^7	3.67135×10^7	3.67135×10^7	3.67136×10^7
$w(n;\phi) = \left(\frac{-1}{ln(1-p)}\frac{p^n}{n}\right)^{-1}$	10.5532					0.816398	-1.83072×10^7	3.66144×10^7	3.66144×10^7	3.66144×10^7
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	6.43497	0					-1.9326×10^7	3.86523×10^7	3.86523×10^7	3.86523×10^7
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	31.6508	0.111737	105.704		37.7525		-1.70191×10^7	3.40382×10^7	3.40382×10^7	3.40382×10^7
$w(n;\phi) = \binom{n}{a}^{-1}$	8.61557	1					-1.83072×10^7	3.66144×10^7	3.66144×10^7	3.66144×10^7
$w(n;\phi) = (a)_n^{-1}$	1	0					-1.15371×10^7	2.30742×10^8	2.30742×10^8	2.30742×10^8
$w(n;\phi) = (n)_a^{-1}$	23.4717	14.2667					-1.72623×10^7	3.42546×10^7	3.42546×10^7	3.42546×10^7

Table 10.31: US flights - Arrival taxi time

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	32.6729						-5.68381×10^7	1.13676×10^8	1.13676×10^8	1.13676×10^8
$w(n) = n$	31.6729						-5.86609×10^7	1.17322×10^8	1.17322×10^8	1.17322×10^8
$w(n;\phi) = n^{-a}$	65.3631	16					-3.94468×10^7	7.88936×10^7	7.88936×10^7	7.88936×10^7
$w(n;\phi) = n+p$	32.6729					4.21475×10^9	-5.68381×10^7	1.13676×10^8	1.13676×10^8	1.13676×10^8
$w(n;\phi) = an^3+bn^2+cn$	31.3892	0.00218118	0	13.5185			-5.85037×10^7	1.17007×10^8	1.17007×10^8	1.17007×10^8
$w(n;\phi) = (n+a)(n-b)^2$	31.7376	132161	31.7703				-5.14456×10^7	1.02891×10^8	1.02891×10^8	1.02891×10^8
$w(n;\phi) = (n+a)(n^2-bn+c)$	31.4181	248.26	4.88903	969.517			-5.69675×10^7	1.13935×10^8	1.13935×10^8	1.13935×10^8
$w(n;\phi) = a+\frac{b-ac}{n+c}$	33.6729	0	76064.3	1			-5.52614×10^7	1.10523×10^8	1.10523×10^8	1.10523×10^8
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	43.4				1	0.752832	-5.68381×10^7	1.13676×10^8	1.13676×10^8	1.13676×10^8
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	Computationally intractable									
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	22896.9	22903					-9.41825×10^6	1.88365×10^7	1.88365×10^7	1.88365×10^7
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	34.7723	1					-5.3196×10^7	1.06392×10^8	1.06392×10^8	1.06392×10^8
$w(n;\phi) = \frac{-1}{ln(1-p)}\frac{p^n}{n}$	63.0721					0.534398	-5.50164×10^7	1.10033×10^8	1.10033×10^8	1.10033×10^8
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	112.619	78.2942	0.0110069		4.44404		-3.01042×10^7	6.02085×10^7	6.02085×10^7	6.02086×10^7
$w(n;\phi) = Boole(n \geq a)$	32.6729	1					-5.68381×10^7	1.13676×10^8	1.13676×10^8	1.13676×10^8
$w(n;\phi) = Boole(n \leq b)$	32.6729		1988				-5.68381×10^7	1.13676×10^8	1.13676×10^8	1.13676×10^8
$w(n;\phi) = Boole(n \geq a)\,Boole(n \leq b)$	32.6729	1	1988				-5.68381×10^7	1.13676×10^8	1.13676×10^8	1.13676×10^8
$w(n;\phi) = \binom{n}{a}$	316729	1					-5.86609×10^7	1.17322×10^8	1.17322×10^8	1.17322×10^8
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n}Boole(n=0)+(1-p)$	32.6729					0	-5.68381×10^7	1.13676×10^8	1.13676×10^8	1.13676×10^8
$w(n;\phi) = (a)_n$	0.978876	0.70507					-9.48171×10^6	1.89634×10^7	1.89634×10^7	1.89634×10^7
$w(n;\phi) = (n)_a$	32.6728	0					-5.68381×10^7	1.13676×10^8	1.13676×10^8	1.13676×10^8
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	28.9754				14	0.634411	-4.36879×10^7	8.73758×10^7	8.73758×10^7	8.73758×10^7
$w(n;\phi) = \left(a+\frac{b-ac}{n+c}\right)^{-1}$	33.6837	24089.3	14794.4	22454.2			-5.51802×10^7	1.1036×10^8	1.1036×10^8	1.1036×10^8
$w(n;\phi) = \left(\frac{-1}{ln(1-p)}\frac{p^n}{n}\right)^{-1}$	1.30411					0.0411744	-5.86609×10^7	1.17322×10^8	1.17322×10^8	1.17322×10^8
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	31.6729	0					-5.86609×10^7	1.17322×10^8	1.17322×10^8	1.17322×10^8
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	Convergence									
$w(n;\phi) = \binom{n}{a}^{-1}$	33.7056	1					-5.50164×10^7	1.10033×10^8	1.10033×10^8	1.10033×10^8
$w(n;\phi) = (a)_n^{-1}$	1	0					-4.54496×10^8	9.08992×10^8	9.08992×10^8	9.08992×10^8
$w(n;\phi) = (n)_a^{-1}$	23168.8	23177.2					-9.41826×10^8	1.88365×10^7	1.88365×10^7	1.88365×10^7

Table 10.32: US flights - Departure delay

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	33.113						-5.46812×10^7	1.09362×10^8	1.09362×10^8	1.09362×10^8
$w(n) = n$	32.113						-5.63536×10^7	1.12707×10^8	1.12707×10^8	1.12707×10^8
$w(n;\phi) = n^{-a}$	65.4194	16					-3.93039×10^7	7.86077×10^7	7.86077×10^7	7.86077×10^7
$w(n;\phi) = n + p$	33.133					7.1043×10^9	-5.46812×10^7	1.09362×10^8	1.09362×10^8	1.09362×10^8
$w(n;\phi) = an^3 + bn^2 + cn$	31.8311	0.0303346	0	194.668			-5.62038×10^7	1.12408×10^8	1.12408×10^8	1.12408×10^8
$w(n;\phi) = (n+a)(n-b)^2$	34.312	189612	35.4338				-4.92968×10^7	9.85935×10^7	9.85935×10^7	9.85935×10^7
$w(n;\phi) = (n+a)(n^2 - bn + c)$	25.9007	0.273149	0.395746	5.49594			-5.99639×10^7	1.19928×10^8	1.19928×10^8	1.19928×10^8
$w(n;\phi) = a + \frac{b-ac}{n+c}$	34.113	0	13330	1			-5.32196×10^7	1.06439×10^8	1.06439×10^8	1.06439×10^8
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	33.113				1	1	-5.46812×10^7	1.09362×10^8	1.09362×10^8	1.09362×10^8
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	Computationally intractable									
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	20006.7	19995.6					-9.39228×10^6	1.87846×10^7	1.87846×10^7	1.87846×10^7
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	35.211	1					-5.13396×10^7	1.02679×10^8	1.02679×10^8	1.02679×10^8
$w(n;\phi) = \frac{-1}{ln(1-p)}\frac{p^n}{n}$	58.7597					0.581101	-5.30098×10^7	1.0602×10^8	1.0602×10^8	1.0602×10^8
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	122.467	86.3643	0.0101213		1.88191		-2.36087×10^7	4.72174×10^7	4.72174×10^7	4.72174×10^7
$w(n;\phi) = Boole(n \geq a)$	33.113	1					-5.46812×10^7	1.09362×10^8	1.09362×10^8	1.09362×10^8
$w(n;\phi) = Boole(n \leq b)$	33.113		1971				-5.46812×10^7	1.09362×10^8	1.09362×10^8	1.09362×10^8
$w(n;\phi) = Boole(n \geq a)\,Boole(n \leq b)$	33.113	1	1971				-5.46812×10^7	1.09362×10^8	1.09362×10^8	1.09362×10^8
$w(n;\phi) = \binom{n}{a}$	32.113	1					-5.63536×10^7	1.12707×10^8	1.12707×10^8	1.12707×10^8
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n}Boole(n=0) + (1-p)$	33.113					0	-5.46812×10^7	1.09362×10^8	1.09362×10^8	1.09362×10^8
$w(n;\phi) = (a)_n$	0.977883	0.748927					-9.3647×10^6	1.87294×10^7	1.87294×10^7	1.87294×10^7
$w(n;\phi) = (n)_a$	33.113	0					-5.468122×10^7	1.09362×10^8	1.09362×10^8	1.09362×10^8
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	27.8055				12	0.630326	-4.36262×10^7	8.72524×10^7	8.72524×10^7	8.72524×10^7
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	34.1452	36903.1	0.0228884	2.2621×10^6			-5.30098×10^7	1.0602×10^8	1.0602×10^8	1.0602×10^8
$w(n;\phi) = \left(\frac{-1}{ln(1-p)}\frac{p^n}{n}\right)^{-1}$	1.14546					0.0356697	-5.63536×10^7	1.12707×10^8	1.12707×10^8	1.12707×10^8
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	32.113	0					-5.63536×10^7	1.12707×10^8	1.12707×10^8	1.12707×10^8
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	Convergence									
$w(n;\phi) = \binom{n}{a}^{-1}$	34.1452	1					-5.30098×10^7	1.0602×10^8	1.0602×10^8	1.0602×10^8
$w(n;\phi) = (a)_n^{-1}$	1	0					-4.50245×10^8	9.0049×10^8	9.0049×10^8	9.0049×10^8
$w(n;\phi) = (n)_a^{-1}$	20531.3	20523.6					-9.39229×10^6	1.87846×10^7	1.87846×10^7	1.87846×10^7

Table 10.33: US flights - Arrival delay

10.4.2 Previous method comparisons

Note that for the Bosch and Ryan [16] models, the variable they referred to at θ has been replaced with a to enable the tables to fit onto a page.

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	1.18						-113.836	229.673	229.673	232.224
$w(n;\phi) = n + p$	0.204567					0.00515226	-64.0713	132.143	132.266	137.353
$w(n;\phi) = (n+a)(n-b)^2$	0.102357	3.07979×10^6	0.047958				-64.6117	135.223	135.473	143.039
$w(n;\phi) = (n+a)(n^2 - bn + c)$	0.182852	0.00274333	3.00274	5.05765			-63.6656	135.331	135.752	145.752
$w(n;\phi) = a + \frac{b-ac}{n+c}$	0.269135	4.35299	0.0196282	1			-63.9696	135.939	136.36	146.36
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	0.135101				15	1	-95.1111	196.222	196.472	204.038
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	1.15154				4	0.380495	-87.5313	181.063	181.313	188.878
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	0.0291491	1007.63	395.359		335.85		-93.6167	195.233	195.655	205.654
$w(n;\phi) = Boole(n \leq b)$	1.21277		4				-113.069	230.139	230.263	235.349
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n}Boole(n=0) + (1-p)$	1.18					0	-113.866	231.673	231.796	236.883
$w(n;\phi) = (a)_n$	0.000231007	5105.82					-113.845	231.691	231.814	236.901
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	1.83046				2	1	-124.761	255.523	255.773	263.338
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	0.366978	0.0932228	0.0112236	0.001			-63.7608	135.522	135.943	145.942
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	0.202347	0.01	0.01		0.020542		-64.0799	136.16	136.581	146.581
$w(n;\phi) = (a)_n^{-1}$	0.44172	0.011195					-64.8869	133.774	133.898	138.984
CBRD1	0.46	1.47					-99.95	203.9	204.024	209.11
CBRD2	-3.5	3.6					-67.75	139.5	139.624	144.71

Table 10.34: Sea-urchin egg fertilization

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	1.63133						-35943.9	71889.9	71889.9	71897.4
$w(n;\phi) = n + p$	1.63133					1.3969×10^8	-35943.9	71891.9	71891.9	71906.9
$w(n;\phi) = (n+a)(n-b)^2$	1.90467	2.1831×10^8	2.57342				-28429.8	56865.6	56865.6	56888.1
$w(n;\phi) = (n+a)(n^2 - bn + c)$	1.1885	3.20848	2.94118	8.25925			-33781.7	67571.5	67571.5	67601.5
$w(n;\phi) = a + \frac{b-ac}{n+c}$	2.3903	0.00100002	176741	1			-32448.1	64904.2	64904.2	64934.2
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	1.63133				1	1	-35943.9	71893.9	71893.9	71916.4
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	0.405822				49	0.163602	-52937.8	105882	105882	105904
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	296.025	329.994	0.541878		0.541878		-21814.3	43636.6	43636.6	43666.7
$w(n;\phi) = Boole(n \leq b)$	1.63133		20				-35943.9	71891.9	71891.9	71906.9
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	1.63133					0	-35943.9	71891.9	71891.9	71906.9
$w(n;\phi) = (a)_n$	0.842871	0.304116					-21821.8	43647.5	43647.5	43662.6
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	4.9171				10	0.636688	-26734.8	53475.6	53475.6	53498.1
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	4.88589	10188.3	223.071	3443.11			-25070.1	50148.3	50148.3	50178.3
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	Convergence									
$w(n;\phi) = (a)_n^{-1}$	1	0.775677					-60280.8	120566	120566	120581
CBRD1	1.89	0.56					-21635	43274	43274	43289
CBRD2	1	0.97					-23551.5	47107	47107	47112

Table 10.35: Canadian doctor visits

Note that for the Consul and Jain [29] models, the variable they referred to at λ_1 has been replaced with λ, and λ_2 has been replaced with a to enable the tables to fit onto a page.

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	0.465224						-617.184	1236.37	1236.37	1240.84
$w(n;\phi) = n + p$	0.465224					9.59731×10^6	-617.184	1238.37	1238.39	1247.31
$w(n;\phi) = (n+a)(n-b)^2$	0.165211	0.0518097	1.42071				-592.188	1190.38	1190.41	1203.79
$w(n;\phi) = (n+a)(n^2 - bn + c)$	0.566645	0.705085	0.971961	0.317144			-671.182	1350.36	1350.43	1368.25
$w(n;\phi) = a + \frac{b-ac}{n+c}$	0.819698	0	7345.4	1			-604.039	1216.08	1216.14	1233.97
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	0.4656589				1	0.999001	-617.184	1240.37	1240.41	1453.79
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	0.378581				15	0.0959253	-674.718	1355.44	1355.47	1368.85
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	2.16807	1.66555	0.624742		0.624742		-591.642	1191.28	1191.35	1209.17
$w(n;\phi) = Boole(n \leq b)$	0.465277		5				-617.178	1238.36	1238.38	1247.3
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	0.465224					0	-617.184	1238.37	1238.39	1247.31
$w(n;\phi) = (a)_n$	0.349703	0.865116					-592.267	1188.53	1188.55	1197.48
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	3.02373				8	0.973263	-594.008	1194.02	1194.05	1207.43
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	1.39995	15.2451	3.85134	113.613			-591.888	1191.78	1191.84	1209.67
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	224952	1.4961×10^6	3.29855×10^6		1.46374×10^6		-592.48	1192.96	1193.02	1210.85
$w(n;\phi) = (a)_n^{-1}$	1	2.05042					-647.086	1298.17	1298.19	1307.12
Generalised Poisson	0.371	0.2					-592.61	1189.22	1189.24	1198.16

Table 10.36: Accidents to women working on shells

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	0.61						-206.107	414.213	414.234	417.512
$w(n;\phi) = n + p$	0.573565					15.1684	-206.106	416.213	416.274	422.809
$w(n;\phi) = (n+a)(n-b)^2$	0.128252	0	0.319025				-206.856	419.712	419.835	429.607
$w(n;\phi) = (n+a)(n^2 - bn + c)$	0.254283	1.37211	0	2.76619			-206.075	420.15	420.355	433.343
$w(n;\phi) = a + \frac{b-ac}{n+c}$	0.614758	1.24228	2.56433	2			-206.106	420.213	420.418	433.406
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	0.61				1	1	-206.107	418.213	418.336	428.108
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	0.194581				13	0.254212	-212.102	430.205	430.327	440.1
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	0.62615	0.01	0.985023		65.8905		-206.106	420.212	420.418	433.406
$w(n;\phi) = Boole(n \leq b)$	0.61194		4				-206.021	416.042	416.103	422.639
$w(n;\phi) = p \frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	0.61					0	-206.107	416.213	416.274	422.81
$w(n;\phi) = (a)_n$	0.0120028	50.1931					-206.118	416.236	416.297	422.832
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	1.04217				2	1	-206.943	419.886	420.009	429.781
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	0.617962	2.7687	0.0737525	0.0272801			-206.103	420.206	420.411	433.339
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	0.232044	9.32095	2.03606		2.64239		-206.178	420.155	420.36	433.348
$w(n;\phi) = (a)_n^{-1}$	1	1.32888					-209.564	423.128	423.188	429.724
Generalised Poisson	0.611	0					-206.107	416.214	416.275	422.811

Table 10.37: Deaths due to horse kicks

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	1.03783						-584.389	1170.78	1170.79	1174.83
$w(n;\phi) = n + p$	1.03782					6.51046×10^6	-584.389	1172.78	1172.81	1180.87
$w(n;\phi) = (n+a)(n-b)^2$	0.28184	1.59856×10^7	0.38857				-599.089	1204.18	1204.23	1216.32
$w(n;\phi) = (n+a)(n^2 - bn + c)$	0.802309	6.29444	1.90788	6.62244			-579.768	1167.54	1167.63	1183.72
$w(n;\phi) = a + \frac{b-ac}{n+c}$	1.40823	0.001	4120.26	1			-581.313	1170.63	1170.72	1186.82
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	1.03886				1	0.999001	-584.389	1174.78	1174.84	1186.92
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	1.69985				17	0.0602107	-637.14	1280.28	1280.34	1292.42
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	7.44123	77.5626	16.7709		0.605035		-579.568	1167.14	1167.23	1183.33
$w(n;\phi) = Boole(n \leq b)$	1.03792		7				-584.383	1172.77	1172.8	1180.86
$w(n;\phi) = p \frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	1.03783					0	-584.389	1172.78	1172.81	1180.87
$w(n;\phi) = (a)_n$	0.190002	4.42434					-579.864	1163.73	1163.76	1171.82
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	1.64253				2	0.998999	-579.667	1165.33	1165.39	1177.48
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	1.62603	16.0382	13.904	19.8735			-597.597	1167.19	1167.29	1183.38
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	1.693	0.25025	44.579		2.24363		-579.596	1167.19	1167.29	1183.38
$w(n;\phi) = (a)_n^{-1}$	1	0.651758					-657.679	1319.36	1319.39	1327.45
Generalised Poisson	0.931	0.1					-579.913	1163.83	1163.85	1171.92

Table 10.38: Lost items in the telephone building

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	5						-199.144	400.227	400.268	402.833
$w(n) = n$	4						-194.898	391.795	391.836	394.4
$w(n;\phi) = n^{-a}$	4.96511	1					-352.843	709.686	709.81	714.897
$w(n;\phi) = n + p$	4					0	-194.898	393.795	393.919	399.005
$w(n;\phi) = an^3 + bn^2 + cn$	2.72816	3.42902	0	5.17431			-192.327	392.655	393.076	403.075
$w(n;\phi) = (n+a)(n-b)^2$	2.10291	0	1.51033				-193.191	392.381	392.631	400.197
$w(n;\phi) = (n+a)(n^2 - bn + c)$	2.329	2.32802	1.2081	0.671049			-195.683	399.367	399.788	409.788
$w(n;\phi) = a + \frac{b-ac}{n+c}$	4.18324	565.636	0.001	17			-195.030	398.061	398.482	408.481
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	1.68746				17	0.754036	-191.56	389.121	389.371	396.936
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	3.62875				24	0.283745	-191.636	389.272	389.552	397.088
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	6.19462	0.00100014					-204.706	413.412	413.536	418.623
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	7.75959	1	1				-215.02	436.04	436.29	443.856
$w(n;\phi) = \frac{-1}{ln(1-p)} \frac{p^n}{n}$	7.2486					0.85443	-204.7	413.4	413.524	418.611
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	0.441622	127.681	174.66		157.161		-191.239	390.477	390.899	400.898
$w(n;\phi) = Boole(n \geq a)$	4.96511	1					-198.426	400.852	400.975	406.062
$w(n;\phi) = Boole(n \leq b)$	5.23715		9				-195.445	394.89	395.013	400.1
$w(n;\phi) = Boole(n \geq a)\, Boole(n \leq b)$	5.20015	1	9				-194.879	395.758	396.008	403.573
$w(n;\phi) = \binom{n}{a}$	4	1					-194.898	393.795	393.919	399.005
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	5					0	-199.114	402.227	402.351	407.438
$w(n;\phi) = (a)_n$	0.00529091	939.341					-199.239	402.479	402.603	407.689
$w(n;\phi) = (n)_a$	1.123831	13.3313					-191.856	387.711	387.835	392.922
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	5.9849				2	1	-203.538	413.076	413.326	420.891
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	4.0004	0	2.5893	0.001			-194.899	397.798	398.219	408.209
$w(n;\phi) = \left(\frac{-1}{ln(1-p)} \frac{p^n}{n}\right)^{-1}$	1.43543					0.358858	-194.898	393.795	393.919	399.005
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	1.23831	12.3313					-191.856	387.711	387.835	392.922
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	0.125224	238.582	1.28712		1.29898		-191.133	390.266	390.687	400.687
$w(n;\phi) = \binom{n}{a}^{-1}$	5	0					-199.144	402.277	402.351	407.438
$w(n;\phi) = (a)_n^{-1}$	1	0					-899.884	1803.77	1803.89	1808.98
$w(n;\phi) = (n)_a^{-1}$	4.96631	0					-198.432	400.864	400.988	406.074
Generalised Poisson	6.901	-0.38					-190.448	384.896	385.02	390.107

Table 10.39: Rifle shots at targets

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	0.540984						-120.241	242.482	242.516	245.286
$w(n;\phi) = n + p$	0.540983					1.5338×10^6	-120.241	244.482	244.583	250.09
$w(n;\phi) = (n+a)(n-b)^2$	0.12478	4.91676×10^6	0.339667				-121.622	249.324	249.528	257.736
$w(n;\phi) = (n+a)(n^2 - bn + c)$	0.495974	1.64572	4.12282	8.26209			-119.771	247.542	247.884	258.758
$w(n;\phi) = a + \frac{b-ac}{n+c}$	0.820968	0.16096	1.04241				-119.895	245.791	245.994	254.203
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	0.540984				1	1	-120.241	246.482	246.686	254.895
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	0.377661				10	0.163963	-126.122	258.245	258.448	266.657
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	55.0716	378.351	1.83511		1.83795		-199.821	247.642	247.984	258.859
$w(n;\phi) = Boole(n \leq b)$	0.542118		4				-120.211	244.422	244.523	250.03
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	0.540984					0	-120.241	244.482	244.583	250.09
$w(n;\phi) = (a)_n$	0.109111	4.41713					-119.836	243.671	243.772	249.279
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	0.937565				2	1	-119.897	245.794	245.997	254.206
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	0.889296	3.2606	3.82269	698.832			-119.884	247.768	248.11	258.984
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	0.854577	0	1.82234	105.648			-119.891	247.781	248.123	258.997
$w(n;\phi) = (a)_n^{-1}$	1	1.59599					-123.27	250.54	250.641	256.148
Generalised Poisson	0.511	-0.06					-119.835	243.67	243.771	249.278

Table 10.40: Home accidents to 122 men

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n)=1$	0.99359						-191.936	385.872	385.898	388.922
$w(n;\phi)=n+p$	0.466504					0.418559	-187.598	379.196	379.275	385.296
$w(n;\phi)=(n+a)(n-b)^2$	0.172976	6.5707×10^6	0.248971				-189.981	385.962	386.12	395.111
$w(n;\phi)=(n+a)(n^2-bn+c)$	0.113551	0.00523427	2.74439	1.88739			-186.533	381.065	381.33	393.264
$w(n;\phi)=a+\frac{b-ac}{n+c}$	0.636592	2.14173	0.531118				-187.435	380.87	381.027	390.019
$w(n;\phi)=\binom{n+r-1}{n}p^n(1-p)^r$	0.24597				6	1	-187.913	381.827	381.985	390.977
$w(n;\phi)=\binom{r}{n}p^n(1-p)^{r-n}$	0.729007				10	0.196983	-189.141	384.282	384.44	393.432
$w(n;\phi)=\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	0.628316	31.8071	11.2521		11.367		-187.901	383.801	384.066	396.001
$w(n;\phi)=Boole(n\leq b)$	1.00957		4				-191.361	386.722	386.8	392.821
$w(n;\phi)=p\frac{n!}{e^{-\lambda}\lambda^n}Boole(n=0)+(1-p)$	0.99359					0	-191.936	387.872	387.951	393.972
$w(n;\phi)=(a)_n$	0.00235533	420.485					-191.984	387.968	388.046	394.068
$w(n;\phi)=\left(\binom{n+r-1}{n}p^n(1-p)^r\right)^{-1}$	1.58343				2	0.998999	-197.311	400.623	400.781	409.772
$w(n;\phi)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$	0.729674	1.09927	0.00353164	0.00144138			-187.291	382.582	382.847	394.781
$w(n;\phi)=\left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	0.511201	0	0.329303		122.553		-187.53	383.061	383.326	395.26
$w(n;\phi)=(a)_n^{-1}$	1	0.529493					-190.525	385.05	385.129	391.15
Generalised Poisson	1.131	-0.14					-188.827	381.654	381.732	387.754

Table 10.41: Number of coal strike outbreaks in the UK

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n)=1$	1.43987						-1438.79	2879.58	2879.59	2884.03
$w(n;\phi)=n+p$	1.43987					3.37274×10^8	-1438.79	2881.58	2881.6	2890.48
$w(n;\phi)=(n+a)(n-b)^2$	1.7965	2.2778×10^7	2.52791				-1123.23	2252.45	2252.49	2265.8
$w(n;\phi)=(n+a)(n^2-bn+c)$	0.94613	0.830066	3.83007	5.17921			-1267.74	2543.48	2543.54	2561.27
$w(n;\phi)=a+\frac{b-ac}{n+c}$	2.1579	0	11035.4	1			-1309.11	2626.22	2626.28	2644.01
$w(n;\phi)=\binom{n+r-1}{n}p^n(1-p)^r$	1.44131				1	0.999001	-1438.79	2883.58	2883.62	2896.93
$w(n;\phi)=\binom{r}{n}p^n(1-p)^{r-n}$	0.449691				32	0.175704	-2022.7	4051.41	4051.45	4064.75
$w(n;\phi)=\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	22.5577	18.6515	0.449024		0.449024		-975.509	1959.02	1959.08	1976.81
$w(n;\phi)=Boole(n\leq b)$	1.4399		9				-1438.79	2881.58	2881.6	2890.47
$w(n;\phi)=p\frac{n!}{e^{-\lambda}\lambda^n}Boole(n=0)+(1-p)$	1.43987					0	-1438.79	2881.58	2881.6	2890.47
$w(n;\phi)=(a)_n$	0.81113	0.335273					-981.678	1967.36	1967.37	1976.25
$w(n;\phi)=\left(\binom{n+r-1}{n}p^n(1-p)^r\right)^{-1}$	24.5155				39	0.998988	-1065.13	2136.26	2136.3	2149.61
$w(n;\phi)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$	4.43721	319.809	9.59517	1200.52			-1034.31	2076.62	2076.69	2094.42
$w(n;\phi)=\left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	Convergence									
$w(n;\phi)=(a)_n^{-1}$	1	0.815425					-2235.63	4475.26	4475.28	4484.16
Generalised Poisson	0.571	0.6					-987.685	1979.37	1979.39	1988.27

Table 10.42: Number of statistics journal authors

Note that for the Shmueli et al. [126] models, the variable they referred to at λ_1 has been replaced with λ, and γ has been replaces with a to enable the tables to fit onto a page.

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	3.30448						-97715	195432	195432	195441
$w(n) = n$	2.30448						-91428.5	182859	182859	182868
$w(n;\phi) = n^{-a}$	3.16498	1					-160643	321289	321289	321307
$w(n;\phi) = n + p$	2.30448					0	-91428.5	182861	182861	182879
$w(n;\phi) = an^3 + bn^2 + cn$	1.2238	3.52834	0	0			-87223	174454	17444	174490
$w(n;\phi) = (n+a)(n-b)^2$	1.02905	0	0.567443				-86402.6	172811	172811	172838
$w(n;\phi) = (n+a)(n^2 - bn + c)$	0.884193	0	2.15404	1.2518			-86192.1	172392	172392	172428
$w(n;\phi) = a + \frac{b-ac}{n+c}$	2.49434	2.4251×10^8	0	7			-91626.6	183261	183261	183297
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	0.507641				22	1	-89138.9	178284	178284	178311
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	1.34038				9	0.61674	-86851.4	173709	173709	173736
$w(n;\phi) = a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	4.33553	0					-101179	202362	202362	202380
$w(n;\phi) = \frac{ab^a}{n^{a+1}}$	5.89614	1	1				-109240	218486	218486	218512
$w(n;\phi) = \frac{-1}{ln(1-p)} \frac{p^n}{n}$	5.01724					0.863904	-101174	202353	202353	202371
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	0.0623577	126.859	455.044		463.294		-88463.8	176936	176936	176972
$w(n;\phi) = Boole(n \geq a)$	3.16498	1					-95410.8	190826	190826	190844
$w(n;\phi) = Boole(n \leq b)$	3.32075		9				-97585.3	195175	195175	195193
$w(n;\phi) = Boole(n \geq a)\, Boole(n \leq b)$	3.17876	1	9				-95311.6	190629	190629	190656
$w(n;\phi) = \binom{n}{a}$	2.30448	1					-91428.5	182861	182861	182879
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	3.30448					0	-97715	195434	195434	195452
$w(n;\phi) = (a)_n$	0.020986	154.087					-98101.8	196208	196208	196226
$w(n;\phi) = (n)_a$	0.0325661	276.284					-86837.9	173680	173680	173698
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	4.24262				2	1	-102244	204493	204493	204520
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	2.30505	0	1.65847	0.001			-91433.6	182875	182875	182911
$w(n;\phi) = \left(\frac{-1}{ln(1-p)} \frac{p^n}{n}\right)^{-1}$	2.17504					0.943832	-91428.5	182861	182861	182879
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	0.0154007	586.727					-86823.3	173651	173651	173669
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	0.0378432	172.867	0.01		0.01		-86257	172522	172522	172558
$w(n;\phi) = \binom{n}{a}^{-1}$	3.30448	0					-97715	195434	195434	195452
$w(n;\phi) = (a)_n^{-1}$	1	0					-231124	462251	462251	462269
$w(n;\phi) = (n)_a^{-1}$	3.16622	0					-95418	190840	190840	190858
$w(n;\phi) = (n!)^{1-\alpha}$	52.1802	3.05325					-86327.9	172660	172660	172678

Table 10.43: Hungarian word length

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	3.61732						-9130.96	18263.9	18263.9	18270
$w(n;\phi) = n + p$	3.61732					5.67322×10^7	-9130.96	18265.9	18265.9	18278.1
$w(n;\phi) = (n+a)(n-b)^2$	2.5686	8.35738×10^7	2.54412				-9135.52	18707	18707.1	18725.3
$w(n;\phi) = (n+a)(n^2 - bn + c)$	2.79606	2.95678	5.95678	19.6129			-8673.47	17354.9	17355	17379.3
$w(n;\phi) = a + \frac{b-ac}{n+c}$	4.56947	0.001	67454.5	1			-8620.95	17249.9	17249.9	17274.2
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	3.61732				1	1	-9130.96	18267.9	18267.9	18286.2
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	0.564059				54	0.3396	-12033.5	24073	24073	24091.2
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	1152.17	1755.11	0.571388		2.6095		-7690.77	15389.5	15389.5	15413.8
$w(n;\phi) = Boole(n \leq b)$	3.61732		30				-9130.96	18265.9	18265.9	18278.1
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	3.61732					0	-9130.96	18265.9	18265.9	18278.1
$w(n;\phi) = (a)_n$	0.692012	1.60993					-7695.01	15394	15394	15406.2
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	4.06259				4	0.636041	-8187.85	16381.7	16381.7	16399.9
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	5.44784	8336.51	804.955	1683.06			-8145.52	16299	16299.1	16323.4
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	29.2262	1.05636	173.441		40.1638		-7699.21	15406.4	15406.4	15430.7
$w(n;\phi) = (a)_n^{-1}$	1	0.106285					-24876.1	49756.3	49756.3	49758.4
$w(n;\phi) = (n!)^{1-\alpha}$	0.991097	0.135148					-7691.54	15387.1	15387.1	15399.2

Table 10.44: Clothing sales

Weight function	$\hat{\lambda}$	$\hat{a}$	Max log-likelihood	AIC	AICc	BIC
$w(n;\phi) = (a)_n$	0.148743	34.9643	-1416.38	2836.76	2836.78	2845.53
$w(n;\phi) = (n!)^{1-\alpha}$	4.49253	0.837851	-1416.21	2836.41	2836.43	2845.39

Table 10.45: Airplane Crashes – Number incidents per month

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{r}$	Max log-likelihood	AIC	AICc	BIC
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	6.93968	123.321	3.1556	3.07126	-217765	435538	435538	435579
$w(n;\phi) = (n!)^{1-\alpha}$	0.59374	1.12583			-218534	437071	437071	437092

Table 10.46: US gun violence – Injured per incident

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	Max log-likelihood	AIC	AICc	BIC
$w(n;\phi) = (n+a)(n-b)^2$	0.340524	0	0.324713	-403086	806179	806179	806211
$w(n;\phi) = (n!)^{1-\alpha}$	5.45389	1.97762		-444474	888953	888953	888974

Table 10.47: Airplane Crashes – Number incidents per month

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	0.214354						-5490.78	10983.	10983.6	10990.7
$w(n;\phi) = n+p$	0.214354					2.01633×10^7	-5490.78	10985.6	10985.6	10999.9
$w(n;\phi) = (n+a)(n-b)^2$	0.616986	71.70982×10^7	2.59526				-5560.80	11127.6	11127.6	11149.1
$w(n;\phi) = (n+a)(n^2-bn+c)$	0.0566645	0.705085	0.971961	0.317144			-5464.29	10936.6	10936.6	10965.2
$w(n;\phi) = a + \frac{b-ac}{n+c}$	0.151408	1.62859	1.29233	1.51643			-5556.07	11120.1	11120.1	11148.8
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	0.214354				1	1	-5490.78	10987.6	10987.6	11009
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	0.780194				14	0.0215932	-5755.85	11517.7	11517.7	11539.2
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	211.309	865.385	0.812762		0.824369		-5348.01	10704	10704	10732.6
$w(n;\phi) = Boole(n \leq b)$	0.214354		7				-5490.78	10983.6	10983.6	10990.7
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	0.214354					0	-5490.78	10983.6	10983.6	10990.7
$w(n;\phi) = (a)_n$	0.234045	0.701512					-5348.04	10700.1	10700.1	10714.4
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	3.59363				20	0.997925	-5361.63	10729.3	10729.3	10750.7
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	0.787352	76.3369	19.8757	1014			-5363.03	10734.1	10734.1	10762.7
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	333387	6.12324×10^6	5.21406×10^6		6.43914×10^6		-5354.68	10715.4	10715.4	10736.8
$w(n;\phi) = (a)_n^{-1}$	1	4.63489					-5557.63	11119.3	11119.3	11133.6
$w(n;\phi) = e^{r\left(-\sqrt{n^2+a}\right)}$	0.115553	32			-4.3		-5342.15	10690.3	10690.3	10711.8
$w(n;\phi) = e^{r\left(-e^{-ak}\right)}$	10.1691	0.2			-22.64		-5342.32	10690.6	10690.6	10712.1
$w(n;\phi) = e^{r\left(\frac{k+1}{k+a}\right)}$	54.8966	14.25			-94.72		-5342.78	10691.6	10691.6	10713
$w(n;\phi) = e^{r\ln(k+a)}$	99.9148	7			-47.85		-5342.55	10691.1	10691.1	10712.6

Table 10.48: Castillo and Perez - Car accidents

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n;\phi) = (a)_n$	0.349703	0.865116					-592.267	1188.53	1188.55	1197.48
$w(n;\phi) = (n+a)^r$	2.14237	0.783974			-2.42914		-591.63	1189.26	1189.3	1202.68

Table 10.49: Accidents to women working on shells

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	0.99359						-191.936	385.872	385.898	388.922
$w(n;\phi) = n + p$	0.466504					0.418559	-187.598	379.196	379.275	385.296
$w(n;\phi) = (n+a)(n-b)^2$	0.172976	6.5707×10^6	0.248971				-189.981	385.962	386.12	395.111
$w(n;\phi) = (n+a)(n^2 - bn + c)$	0.113551	0.00523427	2.74439	1.88739			-186.533	381.065	381.33	393.264
$w(n;\phi) = a + \frac{b-ac}{n+c}$	0.636592	2.14173	0.531118				-187.435	380.87	381.027	390.019
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	0.24597				6	1	-187.913	381.827	381.985	390.977
$w(n;\phi) = \binom{r}{n} p^n (1-p)^{r-n}$	0.729007				10	0.196983	-189.141	384.282	384.44	393.432
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	0.628316	31.8071	11.2521		11.367		-187.901	383.801	384.066	396.001
$w(n;\phi) = Boole(n \leq b)$	1.00957		4				-191.361	386.722	386.8	392.821
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	0.99359					0	-191.936	387.872	387.951	393.972
$w(n;\phi) = (a)_n$	0.00235533	420.485					-191.984	387.968	388.046	394.068
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	1.58343				2	0.998999	-197.311	400.623	400.781	409.772
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	0.729674	1.09927	0.00353164	0.00144138			-187.291	382.582	382.847	394.781
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	0.511201	0	0.329303		122.553		-187.53	383.061	383.326	395.26
$w(n;\phi) = (a)_n^{-1}$	1	0.529493					-190.525	385.05	385.129	391.15
$w(n;\phi) = (n+a)^r$	0.0591327	5.02521			17.6826		-188.209	382.418	382.576	391.567

Table 10.50: Number of strike outbreaks

10.4.3 Novel process data fits

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	2219.92						-264.096	530.193	530.273	532.144
$w(n) = n$	2218.92						-264.1	530.2	530.28	532.151
$w(n;\phi) = n^{-a}$	334.292	2					-61798.8	123602	123602	123606
$w(n;\phi) = n + p$	2219.92					447518	-264.096	532.193	532.438	536.095
$w(n;\phi) = an^3 + bn^2 + cn$	2217.92	0	406.852	14.8862			-264.103	536.206	537.087	544.011
$w(n;\phi) = (n+a)(n-b)^2$	2219.92	180315	299010				-264.096	534.193	534.693	540.047
$w(n;\phi) = (n+a)(n^2 - bn + c)$	2217.9	86040.1	0.348517	126.878			-364.103	536.206	537.057	544.011
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	1931.39				332	1	-265.048	536.096	536.596	541.95
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	316.896	0.0153326	58.9356		1.35271		-64411.6	128831	128832	128839
$w(n;\phi) = Boole(n \geq a)$	2219.92	42					-264.096	532.193	532.438	536.095
$w(n;\phi) = Boole(n \leq b)$	2283.98		2220				-289.592	583.184	583.428	587.086
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	2219.92					0	-264.096	532.193	532.438	536.095
$w(n;\phi) = (a)_n$	0.70881	912.948					-261.929	527.858	528.103	531.76
$w(n;\phi) = (n)_a$	1128.79	1925.27					-311.639	627.279	627.523	631.181
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	15.4094				2	1	-234445	468895	468896	468901
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	472.142	1.03425	89.509	4.88025			-45003.8	90015.6	90016.5	90023.4
$w(n;\phi) = \left(\frac{-1}{ln(1-p)} \frac{p^n}{n}\right)^{-1}$	1.03402					0.00046599	-264.1	532.2	532.444	536.102
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	1128.79	1925.27					-311.21574	626.4315	626.67638	630.3339
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	994.804	0.004569	182.013		6.00783		-14931	29870.1	29870.9	29877.9
$w(n;\phi) = \binom{n}{a}^{-1}$	2219.92	0					-264.096	532.193	532.438	536.095
$w(n;\phi) = (a)_n^{-1}$	1	0					-760228	1520461	1520461	1520461

Table 10.51: Item sales - Item 409

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n)=1$	27.5484						-71.6355	145.271	145.414	146.672
$w(n)=n$	26.5504						-71.4551	144.91	145.053	146.311
$w(n;\phi)=n^{-a}$	28.587	1					-71.829	147.658	148.103	150.46
$w(n;\phi)=n+p$	26.5504					0	-71.4551	146.91	147.355	149.713
$w(n;\phi)=an^3+bn^2+cn$	24.6687	1248.43	0.00592	0.0795168			-71.1284	150.257	151.857	155.862
$w(n;\phi)=(n+a)(n-b)^2$	31.6161	0	40.431				-69.00054	144.001	144.924	148.205
$w(n;\phi)=(n+a)(n^2-bn+c)$	24.5577	0.154709	3.15471	2.48806			-71.09	148.18	149.103	152.384
$w(n;\phi)=\binom{n+r-1}{n}p^n(1-p)^r$	9.05278				84	0.774777	-69.396488	144.793	145.716	148.997
$w(n;\phi)=a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}$	28.5881	0					-71.8292	147.658	148.103	150.461
$w(n;\phi)=\frac{-1}{ln(1-p)}\frac{p^n}{n}$	48.1208					0.594068	-71.829	147.658	148.103	150.46
$w(n;\phi)=\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	2.39206	1.21919	640.546		591.03		-69.0788	146.158	147.758	151.762
$w(n;\phi)=Boole(n\geq a)$	27.5484	2					-71.6355	147.271	147.715	150.073
$w(n;\phi)=Boole(n\leq b)$	37.4904		31				-65.602	135.204	135.648	138.006
$w(n;\phi)=Boole(n\geq a)\,Boole(n\leq b)$	37.4904	2	31				-65.602	137.204	138.127	141.408
$w(n;\phi)=\binom{n}{a}$	25.5334	2					-71.2732	146.546	146.991	149.349
$w(n;\phi)=p\frac{n!}{e^{-\lambda}\lambda^n}Boole(n=0)+(1-p)$	27.5484					0	-71.6355	147.271	147.715	150.073
$w(n;\phi)=(a)_n$	0.0325355	818.433					-71.8035	147.607	148.051	150.409
$w(n;\phi)=(n)_a$	0.470287	1582.61					-68.9551	141.91	142.355	144.713
$w(n;\phi)=\left(\binom{n+r-1}{n}p^n(1-p)^r\right)^{-1}$	14.054				2	0.492305	-71.8144	149.629	150.552	153.832
$w(n;\phi)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$	26.5508	0	6.67601	0			-71.4551	150.91	152.51	156.515
$w(n;\phi)=\left(\frac{-1}{ln(1-p)}\frac{p^n}{n}\right)^{-1}$	0.928012					0.0349529	-71.4551	146.91	147.355	149.713
$w(n;\phi)=\left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	0.422858	1762.3					-68.9526	141.905	142.35	144.708
$w(n;\phi)=\left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	0.754396	943.086	0.0514531		0.01		-68.9253	145.851	147.451	151.455
$w(n;\phi)=\binom{n}{a}^{-1}$	27.5484						-71.6355	147.271	147.715	150.073
$w(n;\phi)=(a)_n^{-1}$	1.99999	0					-1850.09	3704.17	3704.62	3706.98
$w(n;\phi)=(n)_a^{-1}$	27.5494	0					-71.6357	147.271	147.716	150.074

Table 10.52: Mass shooting incidents - June 2013

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n)=1$	184.929						-1442.87	2887.75	2887.76	2891.65
$w(n)=n$	183.955						-1443.23	2888.46	2888.47	2892.36
$w(n;\phi)=n^{-a}$	194.962	10					-1439.24	2882.48	2882.51	2890.28
$w(n;\phi)=n+p$	184.929					358088000	-1442.87	2889.75	2889.78	2897.55
$w(n;\phi)=an^3+bn^2+cn$	183.424	0	0	739.083			-1443.14	2894.28	2894.4	2909.88
$w(n;\phi)=(n+a)(n-b)^2$	184.928	216447	322937000				-1442.87	2891.75	2891.82	2903.45
$w(n;\phi)=(n+a)(n^2-bn+c)$	184.373	389572000	0.665283	90020.9			-1442.79	2893.57	2893.68	2909.17
$w(n;\phi)=\binom{n+r-1}{n}p^n(1-p)^r$	110.122				131	0.99899	-1446.02	2938.04	2938.11	2949.74
$w(n;\phi)=\frac{-1}{ln(1-p)}\frac{p^n}{n}$	228.293					0.81434	-1442.52	2889.04	2889.07	2896.84
$w(n;\phi)=\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}$	236.188	57.0505	16.3393		0.74365		-1429.15	2866.31	2866.42	2881.91
$w(n;\phi)=Boole(n\geq a)$	184.929	1					-1442.87	2889.74	2889.77	2897.54
$w(n;\phi)=Boole(n\leq b)$	22905900		189				-1684.95	3373.9	3373.93	3381.7
$w(n;\phi)=Boole(n\geq a)\,Boole(n\leq b)$	265333	1	189				-1684.97	3375.94	3376.01	3387.64
$w(n;\phi)=\binom{n}{a}$	183.955	1					-1443.23	2890.46	2890.49	2898.26
$w(n;\phi)=p\frac{n!}{e^{-\lambda}\lambda^n}Boole(n=0)+(1-p)$	184.929					0	-1442.87	2889.74	2889.77	2897.54
$w(n;\phi)=(a)_n$	0.666966	90.0725					-1404.44	2812.87	2812.9	2820.67
$w(n;\phi)=(n)_a$	187.853	0					-1447.06	2898.12	2898.16	2905.92
$w(n;\phi)=\left(a+\frac{b-ac}{n+c}\right)^{-1}$	185.899	305.608	2.20237	21077.1			-1442.52	2893.04	2893.15	2908.63
$w(n;\phi)=\left(\frac{-1}{ln(1-p)}\frac{p^n}{n}\right)^{-1}$	0.591063					0.003213	-1443.23	2890.46	2890.49	2898.26
$w(n;\phi)=\left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	183.955	0					-1443.23	2890.46	2890.49	2898.26
$w(n;\phi)=\left(\frac{\Gamma(r+n)}{\Gamma(r)n!}\frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	474.181	0.017301	411.724		1492.15		-1400.42	2804.84	2804.88	2812.64
$w(n;\phi)=\binom{n}{a}^{-1}$	184.929	0					-1442.87	2889.75	2889.78	2897.55
$w(n;\phi)=(a)_n^{-1}$	1	0					-271803	543640	543640	543618

Table 10.53: US gun violence – Rhode Island 2014

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	321.312						-110.048	222.096	222.234	223.53
$w(n) = n$	320.305						-110.027	222.054	222.192	223.488
$w(n;\phi) = n^{-a}$	331.566	10					-110.261	224.522	224.951	227.39
$w(n;\phi) = n + p$	320.305					0	-110.027	224.054	224.483	226.922
$w(n;\phi) = an^3 + bn^2 + cn$	318.299	766.169	0.572295	50.7653			-109.986	227.972	229.51	233.708
$w(n;\phi) = (n+a)(n-b)^2$	337.207	0	360.009				-108.058	222.116	223.005	226.418
$w(n;\phi) = (n+a)(n^2 - bn + c)$	318.289	2.71599	5.71599	17.5258			-109.985	227.971	229.509	233.707
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	72.947				1088	0.999221	-106.623	219.245	220.134	223.547
$w(n;\phi) = \frac{-1}{ln(1-p)} \frac{p^n}{n}$	327.871					0.983079	-110.069	224.138	224.566	227.006
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	323.849	14.6766758671			51.2828		-109.917	227.834	229.372	233.57
$w(n;\phi) = Boole(n \le b)$	213986		316				-95.2035	194.407	194.836	197.275
$w(n;\phi) = \binom{n}{a}$	310.226	11					-109.817	223.634	224.062	226.502
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	321.312					0	-110.048	224.096	224.525	226.964
$w(n;\phi) = (a)_n$	0.280466	824.898					-112.338	228.676	229.104	231.544
$w(n;\phi) = (n)_a$	20.9069	4565.11					-106.116	216.312	216.740	219.180
$w(n;\phi) = \left(\binom{n+r-1}{n} p^n (1-p)^r\right)^{-1}$	28.2663	3302.08	0.285396		1.35243		-106.222	220.444	221.983	226.18
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	320.305	0	58.4612	0			-110.027	228.054	229.593	233.79
$w(n;\phi) = \left(\frac{-1}{ln(1-p)} \frac{p^n}{n}\right)^{-1}$	0.694326					0.00216771	-110.027	224.054	224.483	226.922
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	11.1034	8882.71					-106.075	216.149	216.578	219.017
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	28.2663	3302.08	0.285312		1.35245		-106.222	220.444	221.982	226.18
$w(n;\phi) = \binom{n}{a}^{-1}$	321.312	0					-110.048	224.096	224.525	226.964
$w(n;\phi) = (a)_n^{-1}$	1	0					-46574.3	93152.6	93153.1	93155.5

Table 10.54: Global terrorism - Successful bombings January 2017

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	549.543						-435.190	872.381	872.426	874.892
$w(n) = n$	548.561						-435.292	872.584	872.629	875.094
$w(n;\phi) = n^{-a}$	550.528	1					-435.089	874.178	874.314	879.2
$w(n;\phi) = n + p$	549.543					100000000	-435.19	874.381	874.517	879.402
$w(n;\phi) = an^3 + bn^2 + cn$	547.58	0	107.027	0.0200431			-435.393	878.786	879.252	888.83
$w(n;\phi) = (n+a)(n-b)^2$	549.543	662198	23133700				-435.19	876.381	876.657	883.913
$w(n;\phi) = (n+a)(n^2 - bn + c)$	547.74	725646	1.00551	3999.87			-435.387	878.773	879.817	888.817
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	333.181				366	0.999968	-456.688	919.376	919.652	926.908
$w(n;\phi) = Boole(n \ge a)$	549.543	4					-435.19	874.381	874.517	879.402
$w(n;\phi) = Boole(n \le b)$	549.558		633				-435.359	874.718	874.855	879.74
$w(n;\phi) = Boole(n \ge a)\, Boole(n \le b)$	549.558	4	633				-435.359	876.718	876.994	884.251
$w(n;\phi) = \binom{n}{a}$	545.613	4					-435.598	875.196	875.333	880.218
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	549.543					0	-435.19	874.381	874.517	579.402
$w(n;\phi) = (a)_n$	0.733446	196.481					-399.913	803.825	803.962	808.847
$w(n;\phi) = (n)_a$	337.032	347.423					-456.522	917.045	917.181	922.066
$w(n;\phi) = \left(\frac{-1}{ln(1-p)} \frac{p^n}{n}\right)^{-1}$	0.865993					0.00157866	-435.292	874.584	874.72	879.605
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	548.567	0					-435.292	874.584	874.72	879.605
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	553.372	3.21783	190.296		29.7279		-403.957	815.913	816.378	825.956
$w(n;\phi) = \binom{n}{a}^{-1}$	549.543	0					-435.19	874.381	874.517	879.402
$w(n;\phi) = (a)_n^{-1}$	1	0					-257038	514080	514080	514085

Table 10.55: New York cyclist injuries - January 1 to March 31 2016

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	93.9474						-152.071	306.141	306.255	307.752
$w(n) = n$	92.9082						-152.268	306.535	306.65	308.146
$w(n;\phi) = n^{-a}$	117.236	20					-148.026	300.052	300.405	303.274
$w(n;\phi) = n + p$	93.9472					452622	-152.071	308.141	308.494	311.363
$w(n;\phi) = an^3 + bn^2 + cn$	92.4272	0.0386308	0	1036.87			-152.22	312.439	313.689	318.883
$w(n;\phi) = (n+a)(n-b)^2$	97.1057	1084130	97.9639				-138.303	282.606	283.3333	287.439
$w(n;\phi) = (n+a)(n^2 - bn + c)$	93.4867	716561	0.923608	28939.4			-152.022	312.044	313.294	318.488
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	49.6997				81	0.999015	-159.812	325.624	326.351	330.457
$w(n;\phi) = \frac{-1}{ln(1-p)} \frac{p^n}{n}$	185.806					0.511275	-151.873	307.746	308.099	310.968
$w(n;\phi) = \frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}$	320.934	215.939	0.0100002		0.0100039		-139.295	286.59	287.84	293.033
$w(n;\phi) = Boole(n \geq a)$	93.9474	2					-152.071	308.141	308.494	311.363
$w(n;\phi) = Boole(n \leq b)$	3850030		71				-337.013	678.025	678.378	681.247
$w(n;\phi) = Boole(n \geq a)\, Boole(n \leq b)$	983164	2	71				-337.014	680.027	680.754	684.86
$w(n;\phi) = \binom{n}{a}$	91.8685	2					-152.469	308.937	309.29	312.159
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	93.9474					0	-152.071	308.141	308.494	311.363
$w(n;\phi) = (a)_n$	0.790089	26.2131					-137.371	278.742	279.095	281.964
$w(n;\phi) = (n)_a$	93.9474	0					-152.071	308.141	308.494	311.363
$w(n;\phi) = \left(a + \frac{b-ac}{n+c}\right)^{-1}$	94.9053	667.595	910.724	912.805			-151.88	311.761	313.011	318.204
$w(n;\phi) = \left(\frac{-1}{ln(1-p)} \frac{p^n}{n}\right)^{-1}$	0.692236					0.00745076	-152.268	308.535	308.888	311.757
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	92.9082	0					-152.268	308.535	308.888	311.757
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	435.257	0.168068	350.743		11666.5		-138.42	284.841	286.091	291.284
$w(n;\phi) = \binom{n}{a}^{-1}$	93.9474	0					-152.071	308.141	308.494	311.363
$w(n;\phi) = (a)_n^{-1}$	1	0					-11835.8	23675.6	23676	23678.9

Table 10.56: Boston crime – Monthly bomb hoaxes June 2015 to June 2018

Weight function	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{r}$	$\hat{p}$	Max log-likelihood	AIC	AICc	BIC
$w(n) = 1$	599.849						-335.645	673.291	673.371	675.242
$w(n) = n$	598.872						-335.688	673.375	673.455	675.326
$w(n;\phi) = n + p$	599.848					669196	-335.645	675.291	675.536	679.193
$w(n;\phi) = an^3 + bn^2 + cn$	598.261	0.364801	362.72	442486			-335.689	679.378	680.29	687.183
$w(n;\phi) = (n+a)(n-b)^2$	599.855	1046790	182444				-335.645	677.291	677.791	683.145
$w(n;\phi) = (n+a)(n^2 - bn + c)$	597.905	95840.2	1.00547	523.993			-335.729	679.458	680.31	687.263
$w(n;\phi) = \binom{n+r-1}{n} p^n (1-p)^r$	340.516				470	0.9999	-344.728	695.456	695.956	701.31
$w(n;\phi) = \frac{-1}{ln(1-p)} \frac{p^n}{n}$	605.935					0.991572	-335.603	675.207	675.452	679.109
$w(n;\phi) = Boole(n \geq a)$	599.849	5					-335.645	675.29	675.535	679.192
$w(n;\phi) = Boole(n \leq b)$	599.852		704				-335.645	675.29	675.535	679.192
$w(n;\phi) = Boole(n \geq a)\, Boole(n \leq b)$	599.852	5	704				-335.645	677.289	677.789	683.143
$w(n;\phi) = \binom{n}{a}$	594.964	5					-335.856	675.712	675.957	679.615
$w(n;\phi) = p\frac{n!}{e^{-\lambda}\lambda^n} Boole(n=0) + (1-p)$	599.849	0					-335.645	675.29	675.535	679.192
$w(n;\phi) = (a)_n$	0.958732	23.8434					-278.183	560.366	560.611	564.269
$w(n;\phi) = (n)_a$	600.013	0.000001					-335.646	675.291	675.536	679.193
$w(n;\phi) = \left(\frac{-1}{ln(1-p)} \frac{p^n}{n}\right)^{-1}$	0.895279					0.00149	-335.688	675.375	675.62	679.278
$w(n;\phi) = \left(a\frac{\Gamma(n)\Gamma(a+1)}{\Gamma(n+a+1)}\right)^{-1}$	598.875	0.000059					-335.688	675.375	675.62	679.278
$w(n;\phi) = \left(\frac{\Gamma(r+n)}{\Gamma(r)n!} \frac{Beta(a+r,b+n)}{Beta(a,b)}\right)^{-1}$	605.106	0.277841	162.454		24.4205		-329.245	666.491	667.342	674.296
$w(n;\phi) = \binom{n}{a}^{-1}$	559.849	0					-335.645	675.291	675.536	679.193
$w(n;\phi) = (a)_n^{-1}$	1	0.0000001					-165043	330090	330090	330094

Table 10.57: Missouri Barry County – Weekly unemployment claims 2019

Bibliography

[1] *Application of Katz family of distributions for detecting and testing overdispersion in Poisson regression models*, volume 44, 2008.

[2] J. S. Allison, S. Betsch, B. Ebner, and I. J. H. Visagie. New weighted L2-type tests for the inverse Gaussian distribution. *arXiv preprint arXiv:1910.14119*, 2019.

[3] J. Anderson and M. Siddiqui. The sampling distribution of the index of dispersion. *Communications in Statistics - Theory and Methods*, 23(3):897–911, 1994.

[4] L. J. Bain and M. Engelhardt. *Introduction to Probability and Mathematical Statistics*. Brooks/Cole, 1987.

[5] N. Balakrishnan, L. Bordes, C. Paroissin, and J. Turlot. Single change-point detection methods for small lifetime samples. *Metrika*, 79(5):531–551, 2016.

[6] N. Balakrishnan and T. J. Kozubowski. A class of weighted Poisson processes. *Statistics & Probability Letters*, 78(15):2346–2352, 2008.

[7] L. Baringhaus, N. Gürtler, and N. Henze. Theory & methods: weighted integral test statistics and components of smooth tests of fit. *Australian & New Zealand Journal of Statistics*, 42(2):179–192, 2000.

[8] L. Baringhaus and N. Henze. A goodness of fit test for the Poisson distribution based on the empirical generating function. *Statistics & Probability Letters*, 13(4):269–274, 1992.

[9] G. A. Barnard. Time intervals between accidents, a note on Maguire, Pearson and Wynn's paper. *Biometrika*, 40(1-2):212–213, 1953.

[10] J. J. Bartko, S. W. Greenhouse, and C. S. Patlak. On expectations of some functions of Poisson variates. *Biometrics*, pages 97–102, 1968.

[11] G. I. Bateman. The power of the χ^2 index of dispersion test when Neyman's contagious distribution is the alternate hypothesis. *Biometrika*, 37:59–63, 1950.

[12] J. I. Beltran-Beltran and F. J. O Reilly. On goodness of fit tests for the Poisson, negative binomial and binomial distributions. *Statistical Papers*, pages 1–18, 2016.

[13] D. J. Best and J. C. W. Rayner. Goodness of fit for the Poisson distribution. *Statistics & Probability Letters*, 44(3):259–265, 1999.

[14] D. Böhning, E. Dietz, R. Schaub, P. Schlattmann, and B. G. Lindsay. The distribution of the likelihood ratio for mixtures of densities from the one-parameter exponential family. *Annals of the Institute of Statistical Mathematics*, 46(2):373–388, 1994.

[15] D. D. Boos and C. Brownie. Comparing variances and other measures of dispersion. *Statistical Science*, pages 571–578, 2004.

[16] R. J. Bosch and L. M. Ryan. Generalized Poisson models arising from Markov processes. *Statistics & Probability Letters*, 39(3):205–212, 1998.

[17] S. K. Byung and C. Park. Some remarks on testing goodness of fit for the Poisson assumption. *Communications in Statistics - Theory and Methods*, 21(4):979–995, 1992.

[18] A. C. Cameron and P. K. Trivedi. Econometric models based on count data. Comparisons and applications of some estimators and tests. *Journal of Applied Econometrics*, 1(1):29–53, 1986.

[19] D. B. Campbell and C. A. Oprian. On the Kolmogorov-Smirnov test for the Poisson distribution with unknown mean. *Biometrical Journal*, 21(1):17–24, 1979.

[20] C. Chatfield and G. J. Goodhardt. A consumer purchasing model with Erlang interpurchase times. *Journal of the American Statistical Association*, 68(344):828–835, 1973.

[21] V. Choulakian, R. A. Lockhart, and M. A. Stephens. Cramér-von Mises statistics for discrete distributions. *Canadian Journal of Statistics*, 22(1):125–137, 1994.

[22] M. W. Coffey. On some series representations of the Hurwitz zeta function. *Journal of Computational and Applied Mathematics*, 216(1):297–305, 2008.

[23] A. C. Cohen. Estimating the parameter in a conditional Poisson distribution. *Biometrics*, 16(2):203–211, 1960.

[24] B. J. Collings and B. H. Margolin. Testing goodness of fit for the Poisson assumption when observations are not identically distributed. *Journal of the American Statistical Association*, 80(390):411–418, 1985.

[25] W. J. Conover. A Kolmogorov goodness-of-fit test for discontinuous distributions. *Journal of the American Statistical Association*, 67(339):591–596, 1972.

[26] P. C. Consul. On some models leading to the generalized Poisson distribution. *Communications in Statistics - Theory and Methods*, 17(2):423–442, 1988.

[27] P. C. Consul. Generalized Poisson distribution: properties and applications. *Decker, New York*, 1989.

[28] P. C. Consul and G. C. Jain. On the generalization of Poisson distribution. *Annals of Mathematical Statistics*, 41(4):1387, 1970.

[29] P. C. Consul and G. C. Jain. A generalization of the Poisson distribution. *Technometrics*, 15(4):791–799, 1973.

[30] P. C. Consul and G. C. Jain. On some interesting properties of the generalized Poisson distribution. *Biometrische Zeitschrift*, 15(7):495–500, 1973.

[31] P. C. Consul and M. M. Shoukri. The generalized Poisson distribution when the sample mean is larger than the sample variance. *Communications in Statistics - Simulation and Computation*, 14(3):667–681, 1985.

[32] R. W. Conway and W. L. Maxwell. A queuing model with state dependent service rates. *Journal of Industrial Engineering*, 12(2):132–136, 1962.

[33] D. R. Cox. Some remarks on overdispersion. *Biometrika*, 70(1):269–274, 1983.

[34] H. Cramér. On the composition of elementary errors: First paper: Mathematical deductions. *Scandinavian Actuarial Journal*, 1928(1):13–74, 1928.

[35] R. C. Dahiya and J. Gurland. Functions of the sample mean and sample variance of a Poisson variate. *Biometrics*, pages 171–173, 1969.

[36] J. H. Darwin. The power of the Poisson index of dispersion. *Biometrika*, 44:286–289, 1957.

[37] T. De Oliveira. Some elementary tests for mixtures of discrete distributions. Technical report, Columbia University New York, 1963.

[38] J. Del Castillo and M. Pérez-Casany. Weighted Poisson distributions for overdispersion and underdispersion situations. *Annals of the Institute of Statistical Mathematics*, 50(3):567–585, 1998.

[39] J. Del Castillo and M. Pérez-Casany. Overdispersed and underdispersed Poisson generalizations. *Journal of Statistical Planning and Inference*, 134(2):486–500, 2005.

[40] E. Dietz and D. Böhning. On estimation of the Poisson parameter in zero-modified Poisson models. *Computational Statistics & Data Analysis*, 34(4):441–459, 2000.

[41] J. Durbin. Some methods of constructing exact tests. *Biometrika*, 48(1-2):41–65, 1961.

[42] B. Efron. Double exponential families and their use in generalized linear regression. *Journal of the American Statistical Association*, 81(395):709–721, 1986.

[43] B. Efron and D. V. Hinkley. Assessing the accuracy of the maximum likelihood estimator: Observed versus expected Fisher information. *Biometrika*, 65(3):457–483, 1978.

[44] B. Efron and R. Tibshirani. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, pages 54–75, 1986.

[45] B. Epstein. Tests for the validity of the assumption that the underlying distribution of life is exponential part i. *Technometrics*, 2(1):83–101, 1960.

[46] A. Erdélyi. Chap. VIII. In *Higher Trascendental Functions*, volume 2, pages 115–132. McGraw-Hill New York, 1953.

[47] W. Feller. On a generealized class of contagious distrubtions. *Annals of Mathematical Statistics*, (14):389–400, 1943.

[48] W. Feller. *An introduction to probability theory and its applications, vol 2*. John Wiley & Sons, 2008.

[49] R. Fierro and A. Tapia. Testing homogeneity for poisson processes. *Revista Colombiana de Estadística*, 34(3):421–432, 2011.

[50] R. A. Fisher. The effect of methods of ascertainment upon the estimation of frequencies. *Annals of Human Genetics*, 6(1):13–25, 1934.

[51] R. A. Fisher. The significance of deviations from expectation in a Poisson series. *Biometrics*, 6(1):17–24, 1950.

[52] M. F. Freeman and J. W. Tukey. Transformations related to the angular and the square root. *The Annals of Mathematical Statistics*, pages 607–611, 1950.

[53] J. Frey. An exact Kolmogorov-Smirnov test for the Poisson distribution with unknown mean. *Journal of Statistical Computation and Simulation*, 82(7):1023–1033, 2012.

[54] A. E. Gelfand and S. R. Dalal. A note on overdispersed exponential families. *Biometrika*, 77(1):55–64, 1990.

[55] R. Giacomini, D. N. Politis, and H. White. A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators. *Econometric Theory*, 29(3):567–589, 2013.

[56] J. Gill and G. King. What to do when your Hessian is not invertible: Alternatives to model respecification in nonlinear estimation. *Sociological Methods & Research*, 33(1):54–87, 2004.

[57] D. E. Goldberg and J. H. Holland. Genetic Algorithms and Machine Learning. 1988.

[58] J. M. Gonzalez-Barrios, F. O Reilly, and R. Rueda. Goodness of fit for discrete random variables using the conditional density. *Metrika*, 64(1):77–94, 2006.

[59] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series and Products*. Academic Press, 2007.

[60] M. Greenwood and G. U. Yule. An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *Journal of the Royal Statistical Society*, 83(2):255–279, 1920.

[61] A. K. Gupta, T. F. Móri, and G. J. Székely. Testing for Poissonity-normality vs. other infinite divisibility. *Statistics & Probability Letters*, 19(3):245–248, 1994.

[62] N. Gürtler and N. Henze. Recent and classical goodness-of-fit tests for the Poisson distribution. *Journal of Statistical Planning and Inference*, 90(2):207–225, 2000.

[63] F. A. Haight. *Handbook of the Poisson distribution*. Wiley, 1967.

[64] J. A. Hausman, B. H. Hall, and Z. Griliches. Econometric models for count data with an application to the patents-R&D relationship, 1984.

[65] N. Henze. Empirical-distribution-function goodness-of-fit tests for discrete models. *Canadian Journal of Statistics*, 24(1):81–93, 1996.

[66] N. Henze and B. Klar. Properly rescaled components of smooth tests of fit are diagnostic. *Australian Journal of Statistics*, 38(1):61–74, 1996.

[67] N. Henze and B. Klar. Goodness-of-fit tests for the inverse Gaussian distribution based on the empirical Laplace transform. *Annals of the Institute of Statistical Mathematics*, 54(2):425–444, 2002.

[68] M. R. Hestenes. Multiplier and gradient methods. *Journal of Optimization Theory and Applications*, 4(5):303–320, 1969.

[69] J. M. Hilbe. *Negative Binomial Regression*. Cambridge University Press, 2011.

[70] J. Hinde and C. G. B. Demétrio. Overdispersion: Models and estimation. *Computational Statistics & Data Analysis*, 27(2):151–170, 1998.

[71] D. B. Hning. A note on a test for Poisson overdispersion. *Biometrika*, 81(2):418–419, 1994.

[72] J. H. Holland. Genetic algorithms. *Scientific American*, 267(1):44–50, 1992.

[73] S. D. Horn. Goodness-of-fit tests for discrete data: A review and an application to a health impairment scale. *Biometrics*, pages 237–247, 1977.

[74] N. L. Johnson, A. W. Kemp, and S. Kotz. *Univariate discrete distributions*, volume 444. John Wiley & Sons, 2005.

[75] D. Karlis and E. Xekalaki. Minimum Hellinger distance estimation for Poisson mixtures. *Computational Statistics & Data Analysis*, 1998.

[76] D. Karlis and E. Xekalaki. A simulation comparison of several procedures for testing the Poisson assumption. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 49(3):355–382, 2000.

[77] D. Karlis and E. Xekalaki. Mixed Poisson distributions. *International Statistical Review*, 73(1):35–58, 2005.

[78] N. Karmarkar. A new polynomial-time algorithm for linear programming. In *Proceedings of the Sixteenth Annual ACM Symposium on Theory of Computing*, pages 302–311. ACM, 1984.

[79] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

[80] B. Klar. Goodness-of-fit tests for discrete models based on the integrated distribution function. *Metrika*, 49(1):53–69, 1999.

[81] S. Kocherlakota and K. Kocherlakota. Goodness of fit tests for discrete distributions. *Communications in Statistics - Theory and Methods*, 15(3):815–829, 1986.

[82] C. C. Kokonendji. Over- and underdispersion models. *Methods and applications of statistics in clinical Trials: Planning, analysis, and inferential methods*, 2:506–526, 2014.

[83] C. C. Kokonendji, D. Mizere, and N. Balakrishnan. Connections of the Poisson weight function to overdispersion and underdispersion. *Journal of Statistical Planning and Inference*, 138(5):1287–1296, 2008.

[84] A. Kolmogorov. Sulla determinazione empirica di una lgge di distribuzione. *Inst. Ital. Attuari, Giorn.*, 4:83–91, 1933.

[85] D. G. Konstantinides. *Risk Theory: A Heavy Tail Approach*. World Scientific Publishing, 2017.

[86] J. F. Lawless. Negative binomial and mixed Poisson regression. *Canadian Journal of Statistics*, 15(3):209–225, 1987.

[87] T. Ledwina and G. Wyłupek. On Charlier polynomials in testing Poissonity. *Communications in Statistics - Simulation and Computation*, 46(3):1918–1932, 2017.

[88] S. Lee. Detecting departures from a Poisson model. *Communications in Statistics - Theory and Methods*, 27(5):1201–1210, 1998.

[89] E. L. Lehmann. *Elements of Large-Sample Theory*. Springer Science & Business Media, 2004.

[90] P. Lewis. Some results on tests for Poisson processes. *Biometrika*, 52:67–77, 1965.

[91] B. G. Lindsay. Exponential family mixture models (with least-squares estimators). *The Annals of Statistics*, pages 124–137, 1986.

[92] B. G. Lindsay and K. Roeder. Residual diagnostics for mixture models. *Journal of the American Statistical Association*, 87(419):785–794, 1992.

[93] P. McCullagh and J. A. Nelder. *Generalized Linear Models*, volume 37. CRC Press, 1989.

[94] S. G. Meintanis and Y. Y. Nikitin. A class of count models and a new consistent test for the Poisson distribution. *Journal of Statistical Planning and Inference*, 138(12):3722–3732, 2008.

[95] P. A. Mijburgh, A. Bekker, and S. W. Human. Distributional properties of ratios of gamma random variables in the context of quality control. Master's thesis, University of Pretoria, 2016.

[96] S. Munuswamy. On certain properties of a class of compound Poisson distributions. *Sankhyā: The Indian Journal of Statistics, Series B*, pages 353–362, 1970.

[97] S. Munuswamy. The discrete Poisson-Lindley distribution. *Biometrics*, pages 145–149, 1970.

[98] M. Nakamura and V. Pérez-Abreu. Use of an empirical probability generating function for testing a Poisson model. *Canadian Journal of Statistics*, 21(2):149–156, 1993.

[99] C. A. G. Nass. The χ^2 test for small expectations in contingency tables, with special reference to accidents and absenteeism. *Biometrika*, 46:365–385, 1959.

[100] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, 1965.

[101] I. Newton. De analysi per aequationes numero terminorum infinitas (1669). *London, UK: W. Jones*, 1711.

[102] J. Neyman and E. S. Pearson. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London*, 231(694-706):289–337, 1933.

[103] J. Nocedal and S. Wright. *Numerical Optimization*. Springer Science & Business Media, 2006.

[104] I. Olkin, A. J. Petkau, and J. V. Zidek. A comparison of n estimators for the binomial distribution. *Journal of the American Statistical Association*, 76(375):637–642, 1981.

[105] A. B. Oosthuysen, I. J. H. Visagie, and C. Pretorius. New L^p-type tests for exponentiality based on the interarrival times of the Poisson process. In *Annual Proceedings of the South African Statistical Association Conference*, number Congress 1, pages 33–40, 2019.

[106] G. Patil and S. Joshi. Dictionary and Bibliography of Discrete Distributions. 1968.

[107] G. P. Patil and C. R. Rao. Weighted distributions and size-biased sampling with applications to wildlife populations and human families. *Biometrics*, pages 179–189, 1978.

[108] J. N. Perry and R. Mead. On the power of the index of dispersion test to detect spatial pattern. *Biometrics*, pages 613–622, 1979.

[109] H. M. Pettigrew and W. C. Mohler. A rapid test for the Poisson distribution using the range. *Biometrics*, pages 685–692, 1967.

[110] S. Poisson. *Mémoire sur l'équilibre et mouvement des corps élastiques*. L'Académie des sciences, 1828.

[111] R. F. Potthoff and M. Whittinghill. Testing for homogeneity: II. The Poisson distribution. *Biometrika*, 53:183–190, 1966.

[112] M. J. D. Powell. A method for nonlinear constraints in minimization problems. *Optimization*, pages 283–298, 1969.

[113] C. R. Rao. On discrete distributions arising out of methods of ascertainment. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 311–324, 1965.

[114] J. C. W. Rayner and D. J. Best. Smooth tests of goodness of fit for regular distributions. *Communications in Statistics - Theory and Methods*, 17(10):3235–3267, 1988.

[115] J. C. W. Rayner and R. I. McIntyre. Use of the score statistic for testing goodness-of-fit of some generalised distributions. *Biometrical Journal*, 27(2):159–166, 1985.

[116] T. R. C. Read and N. A. C. Cressie. *Goodness-of-fit Statistics for Discrete Multivariate Data*. Springer Science & Business Media, 2012.

[117] M. S. Ridout and P. Besbeas. An empirical model for underdispersed count data. *Statistical Modelling*, 4(1):77–89, 2004.

[118] T. Rolski, H. Schmidli, V. Schmidt, and J. L. Teugels. *Stochastic processes for insurance and finance*, volume 505. John Wiley & Sons, 2009.

[119] S. M. Ross et al. *Stochastic processes, (Volume 2).* John Wiley & Sons New York, 1996.

[120] R. Rueda and F. O Reilly. Tests of fit for discrete distributions based on the probability generating function. *Communications in Statistics - Simulation and Computation*, 28(1):259–274, 1999.

[121] R. Rueda, F. O Reilly, and V. Perez-Abreu. Goodness of fit for the Poisson distribution based on the probability generating function. *Communications in Statistics - Theory and Methods*, 20(10):3093–3110, 1991.

[122] B. Selby. The index of dispersion as a test statistic. *Biometrika*, 52:627–629, 1965.

[123] K. F. Sellers, S. Borle, and G. Shmueli. The COM-Poisson model for count data: A survey of methods and applications. *Applied Stochastic Models in Business and Industry*, 28(2):104–116, 2012.

[124] K. F. Sellers and D. S. Morris. Underdispersion models: Models that are under the radar. *Communications in Statistics - Theory and Methods*, 46(24):12075–12086, 2017.

[125] M. Shared. On mixtures from exponential families. *Journal of the Royal Statistical Society: Series B*, 42(2):192–198, 1980.

[126] G. Shmueli, T. P. Minka, J. B. Kadane, S. Borle, and P. Boatwright. A useful distribution for fitting discrete data: Revival of the Conway-Maxwell-Poisson distribution. *Journal of the Royal Statistical Society: Series C*, 54(1):127–142, 2005.

[127] H. S. Sichel. On a distribution representing sentence-length in written prose. *Journal of the Royal Statistical Society: Series A*, 137(1):25–34, 1974.

[128] H. S. Sichel. On a distribution law for word frequencies. *Journal of the American Statistical Association*, 70:542–547, 1975.

[129] J. Singh. A characterization of positive Poisson distribution and its statistical application. *SIAM Journal on Applied Mathematics*, 34(3):545–548, 1978.

[130] L. J. Slater. Generalized Hypergeometric Functions. 1966.

[131] J. J. Spinelli and M. A. Stephens. Cramér-von Mises tests of fit for the Poisson distribution. *Canadian Journal of Statistics*, 25(2):257–268, 1997.

[132] R. Storn. On the usage of differential evolution for function optimization. In *Proceedings of North American Fuzzy Information Processing*, pages 519–523. IEEE, 1996.

[133] R. Storn and K. Price. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4):341–359, 1997.

[134] P. Sur, G. Shmueli, S. Bose, and P. Dubey. Modeling bimodal discrete data using Conway-Maxwell-Poisson mixture models. *Journal of Business & Economic Statistics*, 33(3):352–365, 2015.

[135] G. J. Székely and M. L. Rizzo. Mean distance test of Poisson distribution. *Statistics & Probability Letters*, 67(3):241–247, 2004.

[136] S. C. Tan and J. P. San Lau. Time series clustering: A superior alternative for market basket analysis. In *Proceedings of the First International Conference on Advanced Data and Information Engineering*, pages 241–248. Springer, Singapore, 2014.

[137] D. M. Titterington, A. F. M. Smith, and U. E. Makov. *Statistical Analysis of Finite Mixture Distributions*. Wiley, 1985.

[138] B. Treutler. *Tests for the Poisson distribution*. PhD thesis, Diploma Thesis, University of Karlsruhe (in German), 1995.

[139] J. M. Ver Hoef and P. L. Boveng. Quasi-Poisson vs. neagtive binomial regression: How should we model overdispersed count data? *Ecology*, 88(11):2766–2772, 2007.

[140] R. Von Mises. Statistik Und Wahrheit. *Springer*, 20, 1928.

[141] G. E. Willmot. On recursive evaluation of mixed Poisson probabilities and related quantities. *Scandinavian Actuarial Journal*, 1993(2):114–133, 1993.

[142] G. E. Willmot and X. S. Lin. *Lundberg approximations for compound distributions with insurance applications*, volume 156. Springer Science & Business Media, 2001.

[143] G. Wimmer, R. Köhler, R. Grotjahn, and G. Altmann. Towards a theory of word length distribution. *Journal of Quantitative Linguistics*, 1(1):98–106, 1994.

[144] R. Winkelmann. *Econometric Analysis of Count Data*. Springer Science & Business Media, 2008.

[145] D. Zelterman and C. Chen. Homogeneity tests against central-mixture alternatives. *Journal of the American Statistical Association*, 83(401):179–182, 1988.

[146] H. Zhang and B. Li. Characterizations of discrete compound Poisson distributions. *Communications in Statistics - Theory and Methods*, 45(22):6789–6802, 2016.

[147] H. Zhang, Y. Liu, and B. Li. Notes on discrete compound Poisson model with applications to risk theory. *Insurance: Mathematics and Economics*, 59:325–336, 2014.

Data Sources:

`https://archive.ics.uci.edu/ml/datasets/Sales_Transactions_Dataset_Weekly`

`https://www.kaggle.com/crawford/weekly-sales-transactions`

`https://opendata.socrata.com/Government/Airplane-Crashes-and-Fatalities-Since-1908/q2te-8cvq`

`https://www.kaggle.com/saurograndi/airplane-crashes-since-1908`

`https://www.kaggle.com/zusmani/us-mass-shootings-last-50-years`

`https://www.shootingtracker.com`

https://www.kaggle.com/silicon99/dft-accident-data

https://www.kaggle.com/tbsteal/canadian-car-accidents-19942014

https://www.kaggle.com/jameslko/gun-violence-data

https://www.kaggle.com/usdot/flight-delays#flights.csv

https://www.kaggle.com/serifkaya/global-terrorism-database

https://www.kaggle.com/new-york-city/nypd-motor-vehicle-collisions

https://www.kaggle.com/AnalyzeBoston/crimes-in-boston

https://catalog.data.gov/dataset/missouri-weekly-report-of-initial-unemployment-claims-2c10b