

Preliminary test and Bayes Estimation of a Location Parameter Under Blinex Loss

J. COETSEE, A. BEKKER, AND S. MILLARD

Department of Statistics, University of Pretoria, Pretoria, South Africa

In this article, the preliminary test estimator is considered under the BLINEX loss function. The problem under consideration is the estimation of the location parameter from a normal distribution. The risk under the null hypothesis for the preliminary test estimator, the exact risk function for restricted maximum likelihood and approximated risk function for the unrestricted maximum likelihood estimator, are derived under BLINEX loss and the different risk structures are compared to one another both analytically and computationally. As a motivation on the use of BLINEX rather than LINEX, the risk for the preliminary test estimator under BLINEX loss is compared to the risk of the pre-liminary test estimator under LINEX loss and it is shown that the LINEX expected loss is higher than BLINEX expected loss. Furthermore, two feasible Bayes estimators are derived under BLINEX loss, and a feasible Bayes preliminary test estimator is defined and compared to the classical preliminary test estimator.

Keywords Preliminary test estimator; LINEX loss function; BLINEX loss function; Feasible Bayes estimator; Feasible Bayes preliminary test estimator.

Mathematics Subject Classification 62F15; 62F10.

1. Introduction

In this article, the performance of the preliminary test estimator (PTE), restricted maximum likelihood estimator (RMLE), and unrestricted maximum likelihood estimator (UMLE) are all considered under the BLINEX loss function and the risk functions are compared to one another in order to determine which estimator performs better compared to the rest. At the end of this article, all three estimators mentioned above are considered under both the BLINEX and LINEX loss function in order to determine which of the two loss functions performs the best in the sense of having smaller risk. A number of articles have already been published on the preliminary test estimators using different proposed loss structures by Bancroft (1944, 1964), Giles and Giles (1996), Giles (2002), Ohtani et al. (1997), Kibria and Saleh (1993, 2006), Saleh (2006), and Arashi et al. (2008), to name a few. Furthermore, two feasible Bayes estimators are derived under BLINEX loss and a feasible

Address correspondence to J. Coetsee, Department of Statistics, Faculty of Natural and Agricultural Sciences, University of Pretoria, Pretoria, South Africa; E-mail: judy.coetsee@up.ac.za

Bayes preliminary test estimator (FBPTE) is defined and its performance is compared to the performance of the classical preliminary test estimator.

In an article by Wen and Levy (2001), a new parametric family of bounded and asymmetric loss functions, called the BLINEX loss function, was developed and the mathematical properties of the BLINEX loss function were discussed. In Parsian and Kirmani (2002), Porosiński and Kamińska (2009), Arashi (2010), and Arashi and Tabatabaey (2010), the properties of the LINEX function, which was first proposed by Varian (1975), was presented together with results of estimation under LINEX loss for a number of probability distributions. The BLINEX function is the bounded alternative to the LINEX loss function, therefore it doesn't have the same limitations as the LINEX loss function, for example, the LINEX loss function exhibits a huge increase in expected loss which limits the application of the loss function in practice. The BLINEX function is therefore both bounded and asymmetric, which allowed the same flexibility as exhibited by the LINEX loss function but it also has the added advantage of being bounded. The relationship between the BLINEX and LINEX loss functions will be considered next.

Assume that τ is an estimator of the unknown parameter θ then the LINEX loss function is defined as follows:

$$\mathcal{L}^*(\tau, \theta) = d \left(e^{a(\tau-\theta)} - a(\tau - \theta) - 1 \right),$$

where a is the shape parameter and d is the scale parameter.

The BLINEX loss function is derived from the LINEX loss function; let

$$\begin{aligned} \mathcal{L}(\tau, \theta) &= \frac{d \left(e^{a(\tau-\theta)} - a(\tau - \theta) - 1 \right)}{1 + \lambda d \left(e^{a(\tau-\theta)} - a(\tau - \theta) - 1 \right)} \\ &= \frac{1}{\lambda} \left[1 - \frac{1}{1 + b \left(e^{a(\tau-\theta)} - a(\tau - \theta) - 1 \right)} \right] \end{aligned} \quad (1)$$

with $b = \lambda d$.

The BLINEX loss function depends on the selected values of three parameters, namely a , b , and λ , where each of these values plays a significant role in the shape of the loss function. The bounding parameter, λ , specifies the range of the loss which is bounded between 0 and $\frac{1}{\lambda}$. The sign of the constant a will determine the direction of error penalisation, where negative values will penalise negative errors and a positive value will penalise positive errors. This constant also influences the flatness of the curve, for small positive values of a , the curve will be flatter and large positive values will result in a steeper curve. The constant b is an asymmetry parameter and it can be seen that for smaller values of b the loss function is more asymmetric and for larger values of b the function is more symmetric. The influence of the different parameters as discussed above is depicted in Figure 1 for different parameter values.

The problem under study is to consider the estimation of a location parameter μ from a normal population with unknown variance in order to evaluate the finite-sample properties of the preliminary test estimator in terms of its risk. Suppose that X_1, X_2, \dots, X_n is a random sample from $N(\mu, \sigma^2)$ where both $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$ are unknown.

The outcome of a preliminary test on μ will determine which of the estimators will be used to estimate the population mean μ . The unrestricted maximum likelihood estimator, $\bar{X} = \frac{\sum X_i}{n}$, will be regarded as the estimator of μ when nothing is a priori known about the parameter μ , whereas the restricted maximum likelihood estimator will be chosen when

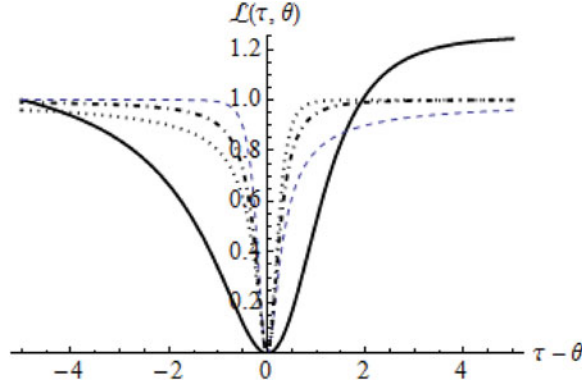


Figure 1. BLINEX loss function 1 with (i) — $\lambda = 0.8, a = b = 1$, (ii) - - - $\lambda = b = 1, a = -5$, (iii) . . . $\lambda = b = 1, a = 5$, and (iv) .-. $\lambda = a = 1, b = 25$.

some constraints are imposed to the model as a null hypothesis. The standard t - test for the population mean will be the appropriate test since the scale parameter σ is unknown.

Now consider the case where we are dealing with the following hypotheses:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_A : \mu \neq \mu_0. \end{cases}$$

Then it is well known that the test statistic is given as $t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} \sim t(n - 1)$ with $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$.

Alternatively, we can use the following statistic $t^2 = \frac{(\bar{X} - \mu_0)^2}{\frac{s^2}{n}} = F \sim F(1, n - 1)$.

In more general terms, the sampling distribution of the test statistic will be the non central F variate, $F \sim F(1, n - 1, \lambda^{**})$ with $\lambda^{**} = \frac{n\delta^2}{2\sigma^2}$ known as the noncentrality parameter and the estimation error is defined as $\delta = \mu_0 - \mu$. It can easily be seen that when $\delta = 0$, the sampling distribution under the null hypothesis will be the central $F(1, n - 1, 0)$ distribution. Based on this proposal the unrestricted maximum likelihood estimator (UMLE) is given by $\mu_1 = \bar{X}$ and the restricted maximum likelihood estimator (RMLE) is equal to μ_0 .

Then the preliminary test estimator of μ is defined as

$$\mu_p = [I_R(F) \times \mu_1] + [I_{NR}(F) \times \mu_0]$$

where

$$\begin{aligned} I_R(F) &= 1 && \text{when the null hypothesis is rejected} \\ &= 0 && \text{otherwise} \end{aligned}$$

and

$$\begin{aligned} I_{NR}(F) &= 1 && \text{when the null hypothesis is not rejected} \\ &= 0 && \text{otherwise.} \end{aligned}$$

The null hypothesis will not be rejected when $F \leq c_\alpha$, where c_α is a critical value defined as $F_{\alpha, 1, n-1, \lambda^{**}=0}$, and α is the level of significance.

The outline of this article is as follows. In Sec. 2, the risk function of the RMLE, UMLE and PTE are derived under BLINEX loss. These different risk functions will be computationally compared. In Sec. 3, two feasible Bayes estimators are derived under BLINEX loss. A feasible Bayes preliminary test estimator (FBPTE) is then defined and the risk function of the PTE is compared to that of the FBPTE. Section 4 is a discussion of the results obtained in this article and illustrations are provided for the motivation to use BLINEX loss rather than LINEX loss functions.

2. Risk Functions

2.1 Risk Function of RMLE

The risk function of RMLE is defined and followed by an illustration of how the risk function behaves for different choices of the parameters in the function.

The risk function of the RMLE is defined as follows:

$$\begin{aligned}\mathfrak{R}(\mu_0) &= E \left[\frac{1}{\lambda} \left[1 - \frac{1}{1 + b(e^{a\delta} - a\delta - 1)} \right] \right] \\ &= \frac{1}{\lambda} \left[1 - \frac{1}{1 + b(e^{a\delta} - a\delta - 1)} \right]\end{aligned}$$

with $\delta = \mu_0 - \mu$.

In the case where $\delta = 0$, $\mathfrak{R}(\mu_0) = 0$.

By examining the risk function it can be seen that when both a and δ are negative values or when both a and δ are positive, with δ either tending to $-\infty$ or ∞ , the risk of RMLE, ($\mathfrak{R}(\mu_0)$) tends to $\frac{1}{\lambda}$. When δ tends to zero, the term $b(e^{a\delta} - a\delta - 1)$ tends to zero and therefore $\mathfrak{R}(\mu_0)$ tends to 0.

2.2 Risk Function of UMLE

By definition, the risk function of UMLE can be calculated as below:

$$\mathfrak{R}(\mu_1) = \int_{-\infty}^{\infty} \frac{1}{\lambda} \left[1 - \frac{1}{1 + b[e^{a(\mu_1 - \mu)} - a(\mu_1 - \mu) - 1]} \right] p(\mu_1) d\mu_1, \quad (2)$$

where $p(\mu_1)$ is the density of $N\left(\mu, \frac{\sigma^2}{n}\right)$.

Therefore, by applying the binomial expansion to $(1 - Z)^{-1}$ where $Z = b[a(\mu_1 - \mu) - e^{a(\mu_1 - \mu)} + 1]$

$$\mathfrak{R}(\mu_1) = \frac{1}{\lambda} - \frac{1}{\lambda} \left(\frac{n}{2\pi\sigma^2} \right)^{\frac{1}{2}} \int_{-\infty}^{\infty} \frac{e^{-\frac{(\mu_1 - \mu)^2}{\frac{n}{2\sigma^2}}}}{1 - z} d\mu_1.$$

The risk function of UMLE can be expanded to:

$$= \frac{b}{\lambda} \left[e^{\frac{a^2\sigma^2}{2n}} - 1 \right] + \frac{a^2b^2\sigma^2}{\lambda n} \left[2e^{\frac{a^2\sigma^2}{2n}} - 1 \right] + \frac{b^2}{\lambda} \left[2e^{\frac{a^2\sigma^2}{2n}} - e^{\frac{2a^2\sigma^2}{n}} - 1 \right]$$

$$\begin{aligned}
& + \frac{3a^2b^3\sigma^2}{\lambda n} \left[e^{\frac{a^2\sigma^2}{2n}} + e^{\frac{a^2\sigma^2}{2n}} \left(\frac{a^2\sigma^2}{n} \right) - 1 \right] + \frac{6a^2b^3\sigma^2}{\lambda n} \left[e^{\frac{a^2\sigma^2}{2n}} - e^{\frac{2a^2\sigma^2}{n}} \right] \\
& + \frac{3b^3}{\lambda} \left[e^{\frac{a^2\sigma^2}{2n}} - e^{\frac{2a^2\sigma^2}{n}} \right] + \frac{b^3}{\lambda} \left[e^{\frac{9a^2\sigma^2}{2n}} - 1 \right] + \dots
\end{aligned} \tag{3}$$

The risk function of the UMLE is not a function of the estimation error, δ , and is therefore also constant with regard to the non centrality parameter λ^{**} .

2.2.1 Monte Carlo Simulations of Risk of UMLE. The risk function of UMLE was also evaluated as a function of λ and it was found that the risk function is a decreasing function of both n and λ . Table 1 gives the simulation results where the function was evaluated for different values of a , b , σ , and n and using $p = 1000$ iterations.

The results in the table just confirm some of the expected characteristics of the risk function, namely that the risk of UMLE is a decreasing function of both λ and n , but the risk is an increasing function of the variance for a specific choice of parameter.

2.3 Risk of PTE

In this section the null risk (risk under the null hypothesis) of the preliminary test estimator is derived under BLINEX loss. The following lemma is due to Judge and Bock (1978) and work done by Clarke (1986).

Lemma 2.1. *Let w be a non central chi-square variate with g degrees of freedom and non centrality parameter θ , let $\phi(\cdot)$, be any real-valued function and let n be any real value such that $n > \frac{-g}{2}$, then*

$$E(w^n \phi(w)) = 2^n \sum_{m=0}^{\infty} \frac{e^{-\theta} \theta^m}{m!} \frac{[\Gamma(\frac{1}{2}(g+2n+2m))] E[\phi(\chi_{(g+2n+2m)}^2)]}{\Gamma[\frac{1}{2}(g+2m)]}$$

Theorem 2.1. *Ignoring the terms of order 3, the risk function of the preliminary test estimator using the BLINEX loss function under H_0 is given by*

$$\begin{aligned}
\mathfrak{R}(\mu_p) &= \frac{2ab}{\lambda} (1 + e^{a\delta}) \left(\frac{2\sigma^2}{n} \right)^{\frac{1}{2}} \sum_{m=0}^{\infty} \frac{e^{-\lambda^{**}} \lambda^{**m} \Gamma[\frac{1}{2}(2+2m)]}{m! \Gamma[\frac{1}{2}(1+2m)]} P[F_1 < c_{\alpha}^*] \\
&+ \frac{3b}{\lambda} \left[\sum_{r=1}^{\infty} \frac{a^r}{r!} \left(\frac{\left(\frac{2\sigma^2}{n} \right)^{\frac{r}{2}} \Gamma(\frac{r+1}{2})}{2\pi} + \delta^r P[F_2 < c_{\alpha}^{**}] - \sum_{j=0}^r (\delta^j) \binom{r}{j} \left[\frac{2\sigma^2}{n} \right]^{\frac{(r-j)}{2}} \right. \right. \\
&\times \left. \left. \sum_{i=0}^{\infty} \frac{e^{-\lambda^{**}} \lambda^{**i} \Gamma[\frac{1}{2}(1+r-j+2i)]}{i! \Gamma[\frac{1}{2}(1+2i)]} P[F_3 < c_{\alpha}^{***}] \right) \right] \\
&- \frac{a^2b}{\lambda} \left[\frac{\left(\frac{2\sigma^2}{n} \right) \Gamma(\frac{3}{2})}{2\pi} + \delta P[F_2 < c_{\alpha}^{**}] - \sum_{j=0}^1 \delta^j \left[\frac{2\sigma^2}{n} \right]^{\frac{(1-j)}{2}} \right]
\end{aligned}$$

Table 1
Risk of UMLE for specific choices of a, b, σ, n and $\lambda = 1$

<i>Sample size</i>	$a = 1, b = 0.1, \lambda = 1$			$a = 0.1, b = 1, \lambda = 1$			$a = 0.1, b = 0.1, \lambda = 1$		
	$\sigma = 1$	$\sigma = 2$	$\sigma = 5$	$\sigma = 1$	$\sigma = 2$	$\sigma = 5$	$\sigma = 1$	$\sigma = 2$	$\sigma = 5$
$n = 10$	0.00469	0.01903	0.10772	0.00046	0.00185	0.01132	0.00005	0.00019	0.00117
$n = 30$	0.00162	0.00647	0.04009	0.00016	0.00065	0.00401	0.00002	0.00007	0.00041
$n = 50$	0.00105	0.00421	0.02632	0.00011	0.00042	0.00261	0.00001	0.00004	0.00026
$n = 100$	0.00051	0.00203	0.01281	0.00005	0.00020	0.00130	0.00000	0.00002	0.00013
$n = 1000$	0.00005	0.00022	0.00135	0.00001	0.00002	0.00013	0.00000	0.00000	0.00001

$$\begin{aligned}
& \times \sum_{k=0}^{\infty} \frac{e^{-\lambda^{**}} \lambda^{**k}}{k!} \frac{\Gamma\left[\frac{1}{2}(2-j+2k)\right]}{\Gamma\left[\frac{1}{2}(1+2k)\right]} P\left[F_4 < c_{\alpha}^{***}\right] \\
& - \frac{b}{\lambda} \left[\sum_{r=1}^{\infty} \frac{(2a)^r}{r!} \left(\frac{\left(\frac{2\sigma^2}{n}\right)^{\frac{r}{2}} \Gamma\left(\frac{r+1}{2}\right)}{2\pi} + \delta^r P[F_2 < c_{\alpha}^{**}] \right. \right. \\
& - \sum_{j=0}^r (\delta^j) \binom{r}{j} \left[\frac{2\sigma^2}{n} \right]^{\frac{r-j}{2}} \sum_{i=0}^{\infty} \frac{e^{-\lambda^{**}} \lambda^{**i}}{i!} \frac{\Gamma\left[\frac{1}{2}(1+r-j+2i)\right]}{\Gamma\left[\frac{1}{2}(1+2i)\right]} \\
& \left. \left. \times P\left[F_3 < c_{\alpha}^{***}\right] \right) \right] + O(Z^3)
\end{aligned}$$

$$\text{with } F_1 = \frac{(n-1)\chi_{(2+2m, \lambda^{**})}^2}{(2+2m)\chi_{(n-1)}^2}, F_2 = \frac{(n-1)\chi_{(1, \lambda^{**})}^2}{\chi_{(n-1)}^2}, F_3 = \frac{(n-1)\chi_{(1+r-j+2i, \lambda^{**})}^2}{(1+r-j+2i)\chi_{(n-1)}^2} \text{ and } F_4 = \frac{(n-1)\chi_{(2-j+2k, \lambda^{**})}^2}{(2-j+2k)\chi_{(n-1)}^2}.$$

Proof. By applying the binomial expansion for $Z = b[a(\mu_p - \mu) - e^{a(\mu_p - \mu)} + 1]$ the risk function is given by

$$\begin{aligned}
\mathfrak{R}(\mu_p) &= E[\mathcal{L}(\mu_p, \mu)] \\
&\approx -\frac{1}{\lambda} [E(Z) + E(Z^2) + O(Z^3)]
\end{aligned} \tag{4}$$

where

$$E(Z) = -b \sum_{r=2}^{\infty} \frac{a^r E(\mu_p - \mu)^r}{r!}.$$

Using the fact that

$$\begin{aligned}
(\mu_p - \mu)^r &= (\mu_1 - \mu)^r I_R(F) - \delta^r I_{NR}(F) \\
&= (\mu_1 - \mu)^r + [\delta^r - (\mu_1 - \mu)^r] I_{NR}(F) \quad \text{for } r = 1, 2, 3, \dots \\
&= (\mu_1 - \mu)^r + \left[\delta^r - \sum_{j=0}^r (\delta^j) \binom{r}{j} (\mu_1 - \mu_0)^{r-j} \right] I_{NR}(F) \\
&= (\mu_1 - \mu)^r + \left[\delta^r - \sum_{j=0}^r (\delta^j) \binom{r}{j} \left[\frac{\sigma^2 \chi_{(1, \lambda^{**})}^2}{n} \right]^{\frac{r-j}{2}} \right] I_{NR}(F),
\end{aligned}$$

where $\chi_{(1, \lambda^{**})}^2$ is a non central Chi-squared variate with 1 degree of freedom and a non centrality parameter $\lambda^{**} = \frac{n\delta^2}{2\sigma^2}$.

Thus,

$$\begin{aligned}
E(\mu_p - \mu)^r &= E[(\mu_1 - \mu)^r] + \delta^r E[I_{NR}(F)] - \sum_{j=0}^r (\delta^j) \binom{r}{j} \left[\frac{\sigma^2}{n} \right]^{\frac{r-j}{2}} \\
&\quad \times E\left[I_{NR}(F) \left(\chi_{(1, \lambda^{**})}^2 \right)^{\frac{r-j}{2}} \right].
\end{aligned}$$

Using Lemma 2.1, we have:

$$E \left[I_{NR}(F) (\chi_{(1, \lambda^{**})}^2)^{\frac{(r-j)}{2}} \right] = \sum_{j=0}^r (\delta^j) \binom{r}{j} \left[\frac{2\sigma^2}{n} \right]^{\frac{(r-j)}{2}} \sum_{i=0}^{\infty} \frac{e^{-\lambda^{**}} \lambda^{**i}}{i!} \\ \times \frac{\Gamma \left[\frac{1}{2} (1 + r - j + 2i) \right]}{\Gamma \left[\frac{1}{2} (1 + 2i) \right]} P [F_3 < c_{\alpha}^{***}]$$

$$E(Z) = -b \left[\sum_{r=2}^{\infty} \frac{a^r}{r!} \left(E[(\mu_1 - \mu)^r] + \delta^r P[F_2 < c_{\alpha}^{**}] - \sum_{j=0}^r (\delta^j) \binom{r}{j} \left[\frac{2\sigma^2}{n} \right]^{\frac{(r-j)}{2}} \right. \right. \\ \left. \left. \times \sum_{i=0}^{\infty} \frac{e^{-\lambda^{**}} \lambda^{**i}}{i!} \frac{\Gamma \left[\frac{1}{2} (1 + r - j + 2i) \right]}{\Gamma \left[\frac{1}{2} (1 + 2i) \right]} P [F_3 < c_{\alpha}^{***}] \right) \right].$$

Consider the following expressions (see Zellner, 1971, pp. 364-365):

$$E(\mu_1 - \mu)^{2r} = \frac{\left(\frac{2\sigma^2}{n} \right)^r \Gamma \left(r + \frac{1}{2} \right)}{2\pi} \quad \text{for } r = 1, 2, 3, \dots \\ E(\mu_1 - \mu)^{2r-1} = 0 \quad \text{for } r = 1, 2, 3, \dots$$

We finally get

$$E(Z) = -b \left[\sum_{r=2}^{\infty} \frac{a^r}{r!} \left(\frac{\left(\frac{2\sigma^2}{n} \right)^{\frac{r}{2}} \Gamma \left(\frac{r+1}{2} \right)}{2\pi} + \delta^r P[F_2 < c_{\alpha}^{**}] - \sum_{j=0}^r (\delta^j) \binom{r}{j} \left[\frac{2\sigma^2}{n} \right]^{\frac{(r-j)}{2}} \right. \right. \\ \left. \left. \times \sum_{i=0}^{\infty} \frac{e^{-\lambda^{**}} \lambda^{**i}}{i!} \frac{\Gamma \left[\frac{1}{2} (1 + r - j + 2i) \right]}{\Gamma \left[\frac{1}{2} (1 + 2i) \right]} P [F_3 < c_{\alpha}^{***}] \right) \right]. \quad (5)$$

Furthermore, we have

$$E(Z^2) = a^2 b E(\mu_p - \mu)^2 + 2ab E(\mu_p - \mu) - 2ab E \left[(\mu_p - \mu) e^{a(\mu_p - \mu)} \right] \\ - 2E \left[e^{a(\mu_p - \mu)} \right] + b E \left[e^{2a(\mu_p - \mu)} \right] + b.$$

Since

$$(\mu_p - \mu) = (\mu_1 - \mu) I_R(F) + (\mu_0 - \mu) I_{NR}(F) \\ = (\mu_1 - \mu) + [\delta - (\mu_1 - \mu)] I_{NR}(F) \\ = (\mu_1 - \mu) + (\mu_1 - \mu_0) I_{NR}(F) \\ = (\mu_1 - \mu) - \left(\frac{\sigma^2}{n} \right)^{\frac{1}{2}} (\chi_{(1, \lambda^{**})}^2)^{\frac{1}{2}} I_{NR}(F)$$

and under H_0 we find

$$\begin{aligned}
E \left[(\mu_p - \mu) e^{a(\mu_p - \mu)} \right] &= E(\mu_1 - \mu) - \left(\frac{\sigma^2}{n} \right)^{\frac{1}{2}} E \left[(\chi_{1, \lambda^{**}}^2)^{\frac{1}{2}} I_{NR}(F) e^{a(\mu_p - \mu)} \right] \\
&= - \left(\frac{\sigma^2}{n} \right)^{\frac{1}{2}} E \left[(\chi_{1, \lambda^{**}}^2)^{\frac{1}{2}} I_{NR}(F) e^{a(\mu_1 - \mu) I_R(F) + a(\mu_0 - \mu) I_{NR}(F)} \right] \\
&= - \left(\frac{\sigma^2}{n} \right)^{\frac{1}{2}} e^{a(\mu_0 - \mu)} E \left[(\chi_{1, \lambda^{**}}^2)^{\frac{1}{2}} I_{NR}(F) \right] \\
&= - \left(\frac{2\sigma^2}{n} \right)^{\frac{1}{2}} e^{a\delta} \sum_{m=0}^{\infty} \frac{e^{-\lambda^{**}} \lambda^{**m}}{m!} \frac{\Gamma \left[\frac{1}{2} (2 + 2m) \right]}{\Gamma \left[\frac{1}{2} (1 + 2m) \right]} P \left[F_1 < c_{\alpha}^* \right],
\end{aligned}$$

otherwise

$$E \left[(\mu_p - \mu) e^{a(\mu_p - \mu)} \right] = 0.$$

We then find $E(Z^2)$ under H_0 :

$$\begin{aligned}
E(Z^2) &= a^2 b \left[\frac{\left(\frac{2\sigma^2}{n} \right) \Gamma \left(\frac{3}{2} \right)}{2\pi} + \delta P(F_2 < c_{\alpha}^{**}) \right. \\
&\quad - \sum_{j=0}^1 \delta^j \left(\frac{2\sigma^2}{n} \right)^{\frac{1-j}{2}} \sum_{k=0}^{\infty} \frac{e^{-\lambda^{**}} \lambda^{**k}}{k!} \frac{\Gamma \left[\frac{1}{2} (2 - j + 2k) \right]}{\Gamma \left[\frac{1}{2} (1 + 2k) \right]} P \left[F_4 < c_{\alpha}^{****} \right] \\
&\quad - 2ab \left(\frac{2\sigma^2}{n} \right)^{\frac{1}{2}} \sum_{m=0}^{\infty} \frac{e^{-\lambda^{**}} \lambda^{**m}}{m!} \frac{\Gamma \left[\frac{1}{2} (2 + 2m) \right]}{\Gamma \left[\frac{1}{2} (1 + 2m) \right]} P \left[F_1 < c_{\alpha}^* \right] \\
&\quad + 2ab (e^{a\delta}) \left(\frac{2\sigma^2}{n} \right)^{\frac{1}{2}} \sum_{m=0}^{\infty} \frac{e^{-\lambda^{**}} \lambda^{**m}}{m!} \frac{\Gamma \left[\frac{1}{2} (2 + 2m) \right]}{\Gamma \left[\frac{1}{2} (1 + 2m) \right]} P \left[F_1 < c_{\alpha}^* \right] \\
&\quad - 2b \left[\sum_{r=1}^{\infty} \frac{a^r}{r!} \left(\frac{2\sigma^2}{n} \right)^{\frac{r}{2}} \frac{\Gamma \left(\frac{r+1}{2} \right)}{2\pi} + \delta^r P \left[F_2 < c_{\alpha}^{**} \right] - \sum_{j=0}^r (\delta^j) \binom{r}{j} \right. \\
&\quad \times \left. \left[\frac{2\sigma^2}{n} \right]^{\frac{(r-j)}{2}} \sum_{i=0}^{\infty} \frac{e^{-\lambda^{**}} \lambda^{**i}}{i!} \frac{\Gamma \left[\frac{1}{2} (1 + r - j + 2i) \right]}{\Gamma \left[\frac{1}{2} (1 + 2i) \right]} P \left[F_3 < c_{\alpha}^{****} \right] \right] \\
&\quad + b \left[\sum_{r=1}^{\infty} \frac{(2a)^r}{r!} \left(\frac{2\sigma^2}{n} \right)^{\frac{r}{2}} \frac{\Gamma \left(\frac{r+1}{2} \right)}{2\pi} + \delta^r P \left[F_2 < c_{\alpha}^{**} \right] - \sum_{j=0}^r (\delta^j) \binom{r}{j} \right. \\
&\quad \times \left. \left[\frac{2\sigma^2}{n} \right]^{\frac{(r-j)}{2}} \sum_{i=0}^{\infty} \frac{e^{-\lambda^{**}} \lambda^{**i}}{i!} \frac{\Gamma \left[\frac{1}{2} (1 + r - j + 2i) \right]}{\Gamma \left[\frac{1}{2} (1 + 2i) \right]} P \left[F_3 < c_{\alpha}^{****} \right] \right]. \quad (6)
\end{aligned}$$

Substituting (5) and (6) into (4) the result follows. \square

2.4 Comparison Results

The risk of UMLE and PTE are analytically compared to one another and result is given in the following theorem.

Theorem 2.2. *The PTE performs better than the UMLE in the sense of having smaller risk if the following inequality holds:*

$$\delta \leq \frac{1}{a} \ln \left[\left(\frac{\frac{1}{b} E \left[1 - \frac{1}{x_{\mu_1}} \right] + a E (\mu_p - \mu) + 1}{E \left[e^{a(\mu_1 - \mu) I_R[F]} \right]} - P(F > c) \right) / P(F < c) \right].$$

Proof. It is enough to show that $\mathfrak{R}(\mu_p) \leq \mathfrak{R}(\mu_1)$. Therefore, we should have $E[1 - \frac{1}{x_{\mu_p}}] \leq E[1 - \frac{1}{x_{\mu_1}}]$, where $x_{\mu_p} = 1 + b[e^{a(\mu_p - \mu)} - a(\mu_p - \mu) - 1]$ and $x_{\mu_1} = 1 + b[e^{a(\mu_1 - \mu)} - a(\mu_1 - \mu) - 1]$, knowing that $E[1 - \frac{1}{x_{\mu_1}}]$ is not a function of δ .

Applying the inequality, $1 - \frac{1}{x} \leq x - 1, \forall x > 0$ to the LHS of the above expression it is sufficient to prove that $E[x_{\mu_p} - 1] \leq E[1 - \frac{1}{x_{\mu_1}}]$.

Simplifying the relevant term it yields

$$E \left[e^{a(\mu_p - \mu)} \right] \leq \frac{1}{b} E \left[1 - \frac{1}{x_{\mu_1}} \right] + a E (\mu_p - \mu) + 1.$$

Setting $\frac{1}{b} E[1 - \frac{1}{x_{\mu_1}}] + a E(\mu_p - \mu) + 1 = Q$ we then have

$$E \left[e^{a(\mu_p - \mu)} \right] \leq Q. \quad (7)$$

The LHS of the above inequality can be written as

$$E \left[e^{a(\mu_p - \mu)} \right] = E \left[e^{a(\mu_1 - \mu) I_R[F]} e^{a\delta I_{NR}[F]} \right]. \quad (8)$$

By applying the Cauchy-Schwarz inequality and using (7) and (8), we obtain

$$E \left[e^{a(\mu_1 - \mu) I_R[F]} e^{a\delta I_{NR}[F]} \right] \leq E \left[e^{a(\mu_1 - \mu) I_R[F]} \right] E \left[e^{a\delta I_{NR}[F]} \right].$$

Then, it is enough to show that $E \left[e^{a(\mu_1 - \mu) I_R[F]} \right] E \left[e^{a\delta I_{NR}[F]} \right] \leq Q$.

From the fact that $E \left[e^{a(\mu_1 - \mu) I_R[F]} \right] > 0$, we have

$$E \left[e^{a\delta I_{NR}[F]} \right] \leq \frac{Q}{E \left[e^{a(\mu_1 - \mu) I_R[F]} \right]}. \quad (9)$$

However, the LHS of the above inequality can be written as

$$\begin{aligned} E \left[e^{a\delta I_{NR}[F]} \right] &= E \left[\sum_{k=0}^{\infty} \frac{(a\delta)^k}{k!} I_{NR}[F] \right] \\ &= P(F > c) + P(F < c) \left[\sum_{k=0}^{\infty} \frac{(a\delta)^k}{k!} \right] \\ &= P(F > c) + P(F < c) e^{a\delta} \end{aligned} \quad (10)$$

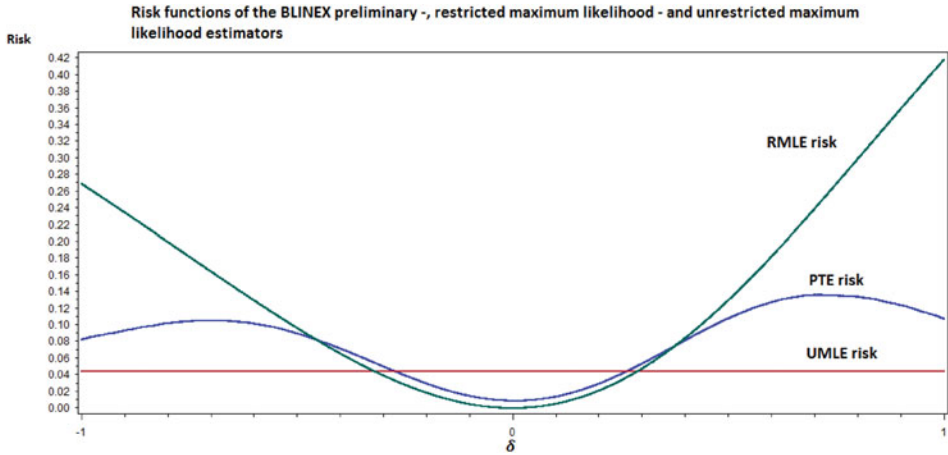


Figure 2. $a = b = \lambda = 1, \sigma = 1, n = 10$.

By making use of the expression in (9) and (10) we have

$$P(F > c) + P(F < c)e^{a\delta} \leq \frac{Q}{E[e^{a(\mu_1 - \mu)I_R[F]}]}.$$

By simplifying and taking logs on both sides of the inequality the result follows. \square

The risk of RMLE, UMLE, and PTE calculated under BLINEX loss are computationally compared to one another in order to determine the performance of the estimators relative to one another. The sketches (see Figs. 2-10) in this section will be discussed in two ways, first by keeping σ constant and allowing the sample size, n , to increase and secondly by keeping n constant and allowing σ to increase.

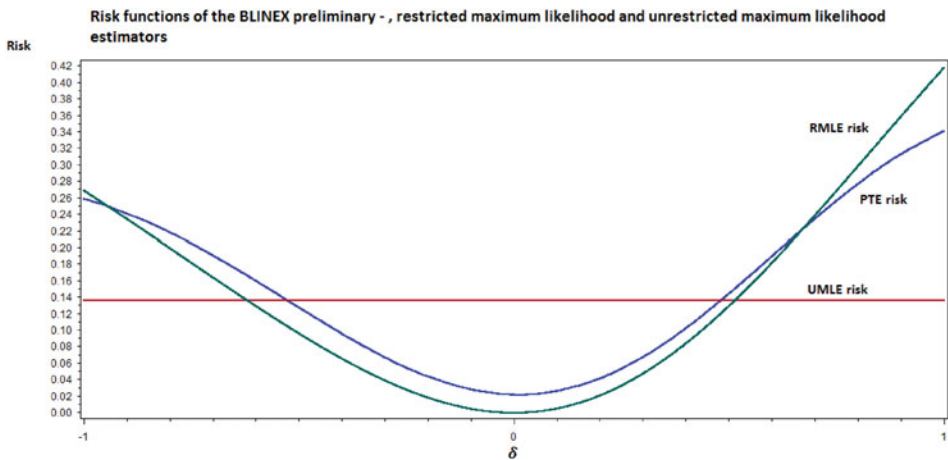


Figure 3. $a = b = \lambda = 1, \sigma = 2, n = 10$.

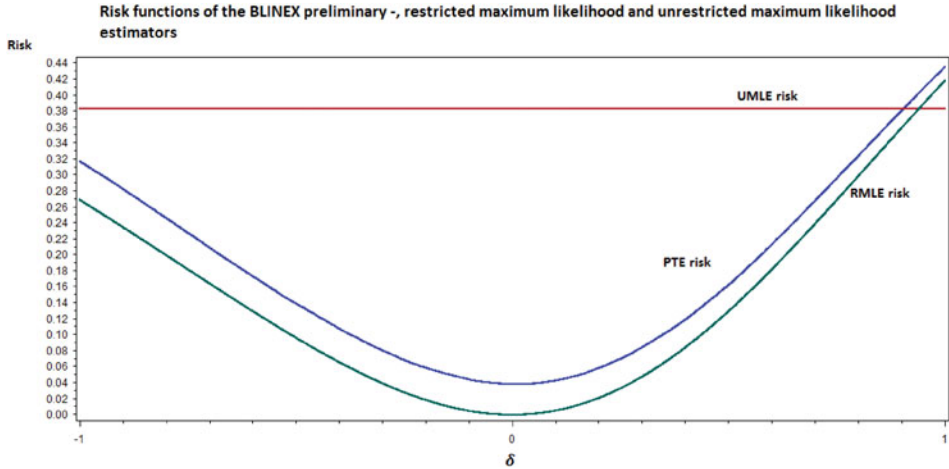


Figure 4. $a = b = \lambda = 1, \sigma = 5, n = 10$.

For the case where $\sigma = 1$, it can be seen that as the sample size increases for this choice of σ , the risk of PTE tends towards the risk of UMLE as $\delta \rightarrow \pm\infty$. It is well known that \bar{X} (UMLE) is the most efficient estimator as the sample size increases, since $Var(\bar{X}) = \frac{\sigma^2}{n}$. It is clear that as the sample size increases the $Var(\bar{X})$ will tend to zero. Therefore as the sample size increases, PTE becomes just as efficient as the UMLE. Similar results are obtained for different values of σ .

When σ increases for a specific sample size, it can be seen that in the interval near the origin, the domination interval of PTE relative to the UMLE becomes larger. Therefore, when working with a specific sample size and if the variance of the sample is relatively high, the PTE is a more efficient estimator compared to the UMLE.

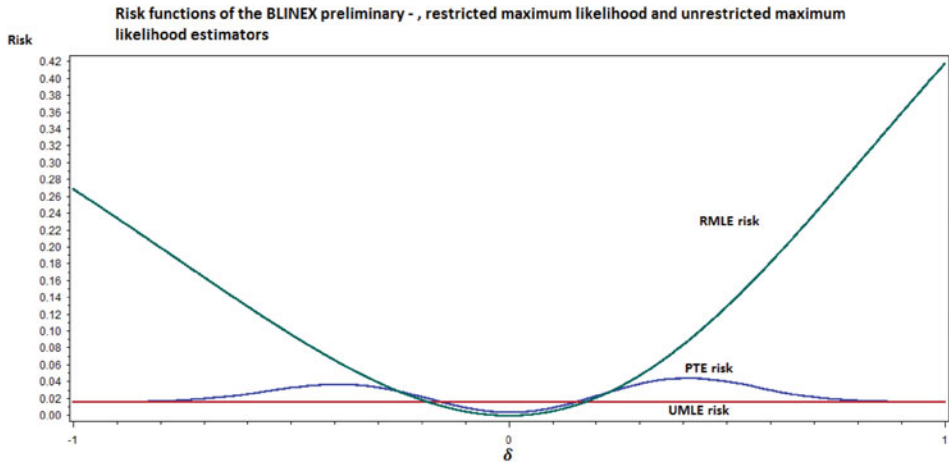


Figure 5. $a = b = \lambda = 1, \sigma = 1, n = 30$.

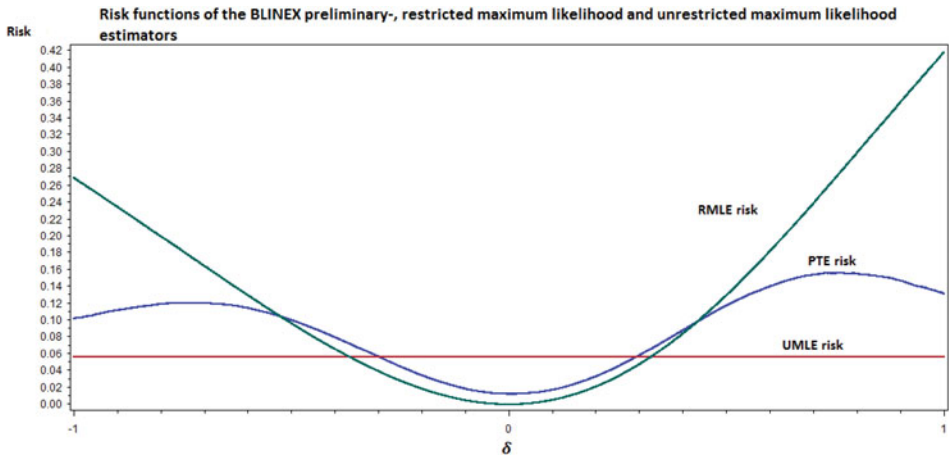


Figure 6. $a = b = \lambda = 1, \sigma = 2, n = 30$.

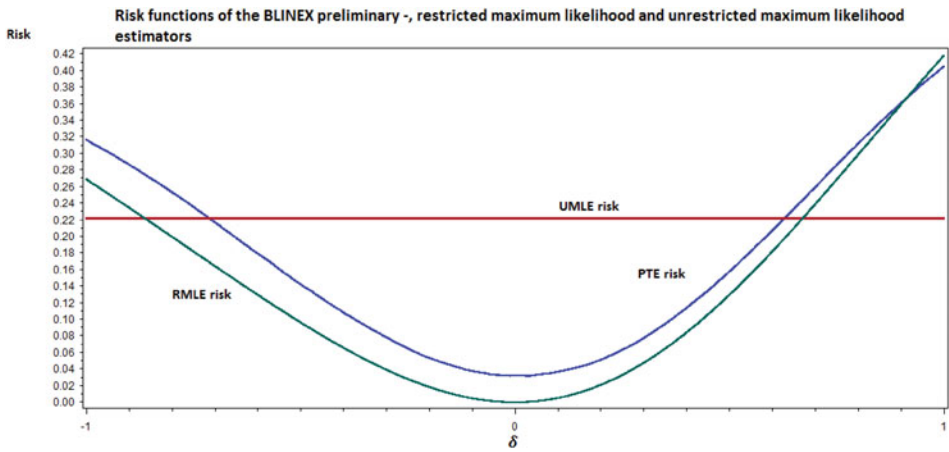


Figure 7. $a = b = \lambda = 1, \sigma = 5, n = 30$.

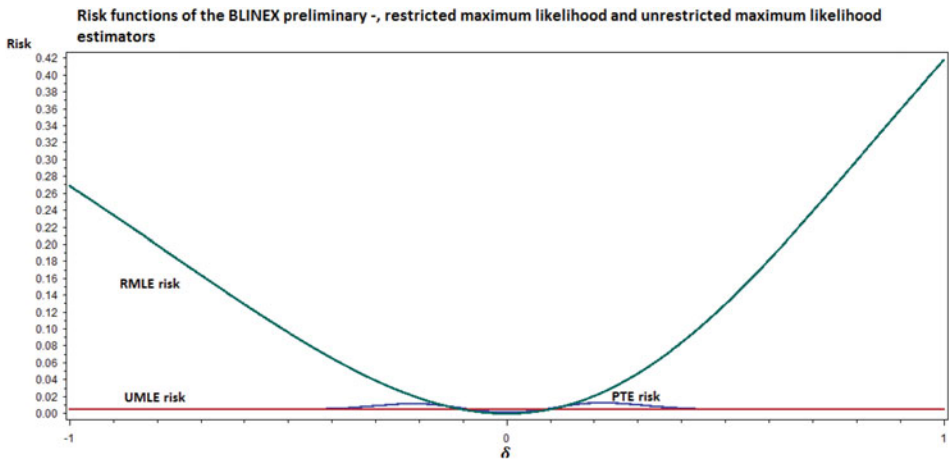


Figure 8. $a = b = \lambda = 1, \sigma = 1, n = 100$.

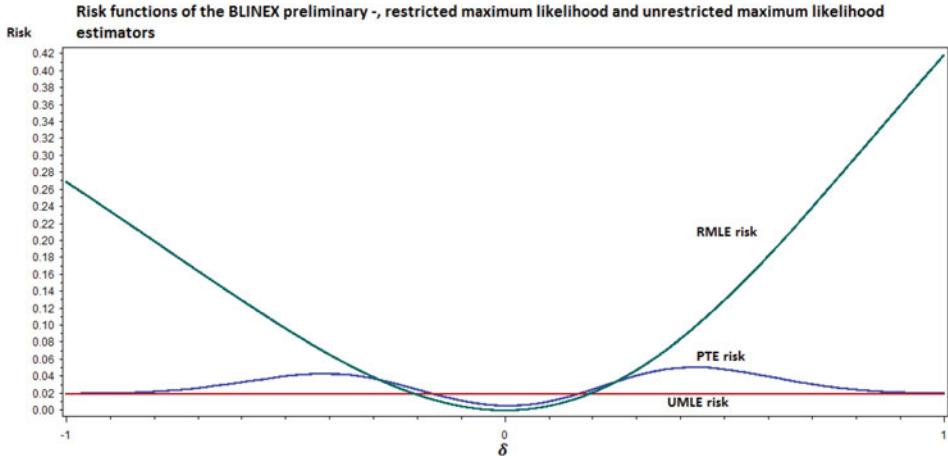


Figure 9. $a = b = \lambda = 1, \sigma = 2, n = 100$.

3. Bayesian Estimation under the BLINEX Loss Function

3.1 Feasible Bayes Estimator

The Bayes estimator under BLINEX loss is derived in this section and due to the complex nature of BLINEX loss function it was necessary to make use of alternative methods in order to propose feasible estimators.

Suppose that X_1, X_2, \dots, X_n is a sample from $N(\mu, \sigma^2)$ where $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$ is assumed to be known. For the case of σ unknown, one may look at the work of Arashi (2010), the results obtained in his paper are very complex. It is wellknown that the conjugate prior for the unknown parameter μ is $N(\mu', \sigma'^2)$ and the posterior of $\mu | \mathbf{x}$ is $N(\mu'', \sigma''^2)$ where $\mu'' = \frac{(n\mu_1\sigma'^2 + \mu'(\sigma^2))}{n\sigma'^2 + \sigma^2}$ and $\sigma''^2 = \frac{\sigma^2\sigma'^2}{n\sigma'^2 + \sigma^2}$.

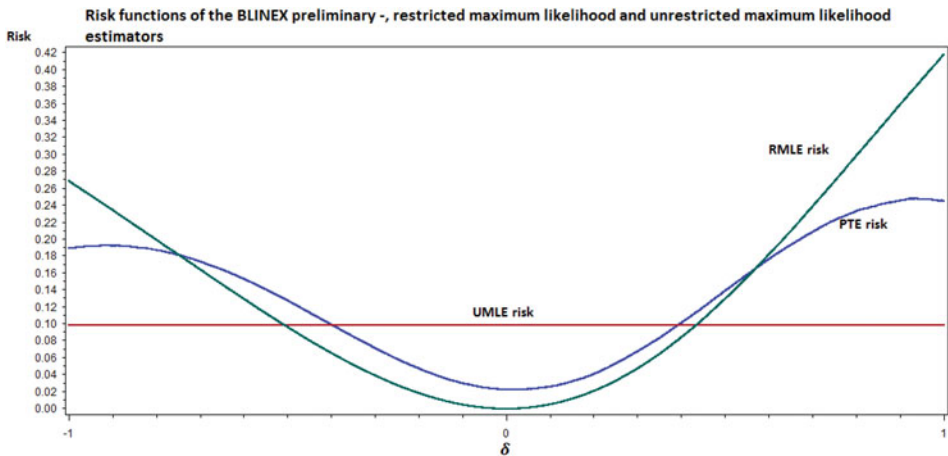


Figure 10. $a = b = \lambda = 1, \sigma = 5, n = 100$.

To derive the Bayes estimator, the expected loss is equal to

$$EL = \int_{-\infty}^{\infty} \mathcal{L}(\mu_B, \mu) p(\mu|\mathbf{x}) d\mu$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n)$

Then under the pre-specified assumptions we have

$$EL(\mu_B) = \frac{1}{\lambda} - \frac{1}{\lambda} E \left[\frac{1}{1 + b(e^{a(\mu_B - \mu)} - a(\mu_B - \mu) - 1)} \right]. \quad (11)$$

Then the Bayes estimator of μ denoted by $\hat{\mu}_B$ is the one that minimises the posterior expectation of the loss function, i.e.,

$$\hat{\mu}_B = \arg \min_{\mu_B} EL(\mu_B).$$

Note that

$$\frac{\partial EL(\mu_B)}{\partial \mu_B} = \frac{b}{\lambda} E \left[\frac{ae^{a(\mu_B - \mu)} - a}{[1 + b(e^{a(\mu_B - \mu)} - a(\mu_B - \mu) - 1)]^2} \right] = 0.$$

Take $f_a(\mu_B - \mu) = ae^{a(\mu_B - \mu)} - a$. It can be seen that $f_a(\mu_B - \mu) \geq 0$ when $\mu_B \geq \mu$ and $f_a(\mu_B - \mu) < 0$ when $\mu_B < \mu$. Also note that for a random variable $X \geq 0$, $E(X) = 0$ implies $X = 0$ a.e. Therefore, subject to the constraint $\mu_B \geq \mu$ the Bayes estimator is equal to

$$\hat{\mu}_B = \mu. \quad (12)$$

Now under the restriction $\mu_B < \mu$, consider the following optimisation problem

$$\arg \max_{\mu_B} EL(\mu_B)$$

$$\arg \max_{\mu_B} E \left[\frac{1}{1 + b(e^{a(\mu_B - \mu)} - a(\mu_B - \mu) - 1)} \right]. \quad (13)$$

Consider that $g(\mu_B - \mu) = \left[\frac{1}{1 + b(e^{a(\mu_B - \mu)} - a(\mu_B - \mu) - 1)} \right]$ is an increasing function since $g'(\mu_B - \mu) > 0$ when $\mu_B < \mu$. Also from the fact that $e^{a(\mu_B - \mu)} > a(\mu_B - \mu) + 1$ when $\mu_B < \mu$, it can be seen that $g(\mu_B - \mu) > 0$. There is no unique solution for (13), therefore (12) is a valid result.

The problem with the above solution (12) is that $\hat{\mu}_B$ is not a valid estimator, because μ is unknown. In this case a feasible estimator, denoted by $\hat{\mu}_{FB}$, can be obtained by substituting a proper estimator of μ into the solution to get

$$\hat{\mu}_{FB}^{(1)} = \mu'.$$

Another possible solution is to use two sets of available information contained in the prior distribution, for the unknown parameter μ and the density function $.X|\mu$, which contains the information of the original sample.

In this case, define

$$\lambda \equiv L(\mu, \sigma^2) \cdot p(\mu)$$

$$\propto e^{-\frac{1}{2\sigma^2} \sum (x_i - \mu)^2 - \frac{(\mu - \mu')^2}{2\sigma'^2}}.$$

Then taking the derivative with respect to μ , equating to zero and substituting the final result into (12) we obtain the feasible estimator

$$\widehat{\mu}_{FB}^{(2)} = \frac{\left(\sum_{i=1}^n x_i + \frac{\mu' \sigma^2}{\sigma'^2}\right)}{\left(\frac{\sigma^2}{\sigma'^2} + n\right)}. \quad (14)$$

Note that λ is just a way to produce an estimator of μ and it totally independent of the loss function. The result of λ was substituted into (11), that was initiated from BLINEX loss and the sketches were generated using simulation exercises based on (11). This feasible Bayes estimator is exactly the same as the Bayes estimator under squared error loss.

The properties of $\widehat{\mu}_{FB}^{(2)}$ are given below:

$$E\left[\widehat{\mu}_{FB}^{(2)}\right] = \frac{n\mu + \mu' \left(\frac{\sigma^2}{\sigma'^2}\right)}{\frac{\sigma^2}{\sigma'^2} + n};$$

if $\mu = \mu'$ then $E\left[\widehat{\mu}_{FB}^{(2)}\right] = \frac{\left(\frac{\sigma^2}{\sigma'^2} + n\right)\mu}{\frac{\sigma^2}{\sigma'^2} + n} = \mu$, therefore under these circumstances $\widehat{\mu}_{FB}^{(2)}$ is an unbiased estimator of μ ;

$$Var\left[\widehat{\mu}_{FB}^{(2)}\right] = \frac{n\sigma^2}{\left(\frac{\sigma^2}{\sigma'^2} + n\right)^2} = \frac{n\sigma^2\sigma'^2}{(\sigma^2 + n\sigma'^2)^2};$$

if $\sigma^2 = \sigma'^2$ then $Var\left[\widehat{\mu}_{FB}^{(2)}\right] = \left[\frac{n}{(n+1)^2}\right]\sigma^2$, it can therefore be seen that when $n \rightarrow \infty$, $Var\left[\widehat{\mu}_{FB}^{(2)}\right] \rightarrow 0$.

3.2 A Feasible Bayes Preliminary Test estimator

Facing the same problem as in the introduction and using $\widehat{\mu}_{FB}^{(2)}$ as the unrestricted estimator, the feasible Bayes preliminary test estimator is defined by

$$\mu_{FBPT} = \left[I_R(F) \times \widehat{\mu}_{FB}^{(2)}\right] + [I_{NR}(F) \times \mu_0].$$

The performance of FBPT will be explored under BLINEX loss in the following section.

3.2.1 Comparing PTE with FBPT under BLINEX loss function. The risk functions for the PTE and FBPT were calculated for specific choices of the parameters involved and are compared to one another. In all the sketches the risk function of the PTE is presented by the dashed line curve and the risk function of the FBPT is given by the solid line curve. Both risk functions of the estimators are functions of δ , where $-1 \leq \delta \leq 1$. In the calculation of the FBPT only a fixed value for the prior mean, $\mu' = 13$ and prior standard deviation,

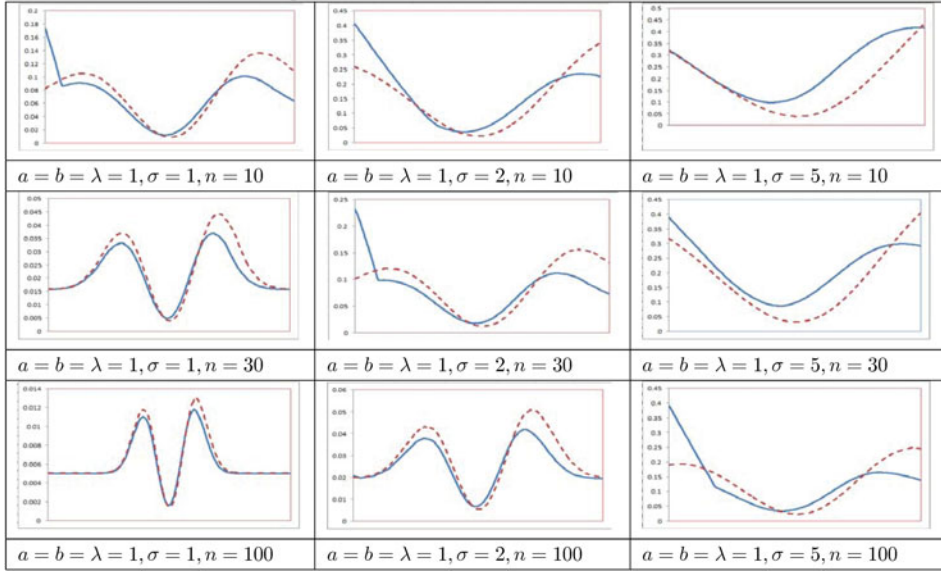


Figure 11. Comparison of PTE and FBPTE for different parameter values.

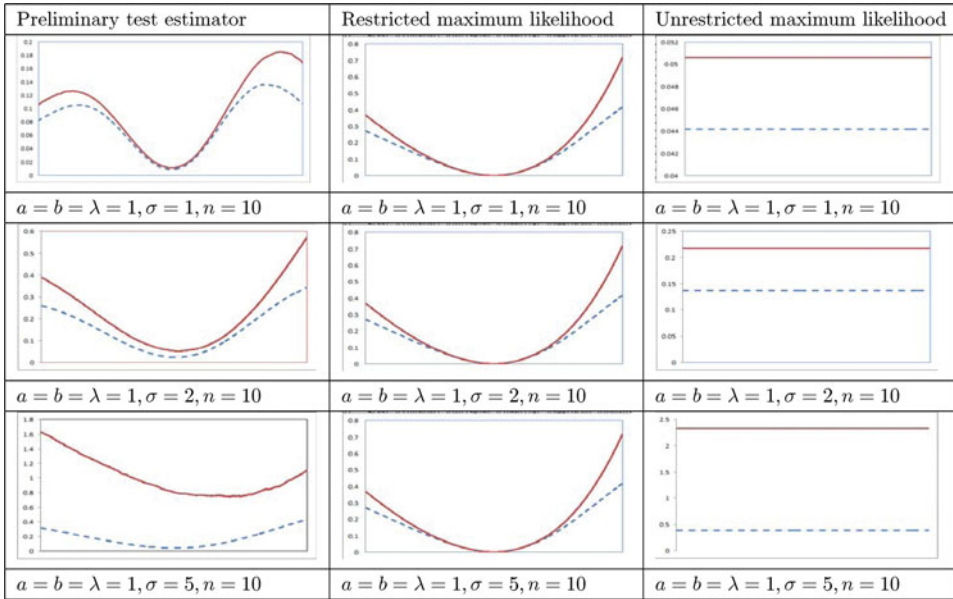
$\sigma' = 1$, were considered and the performance of the estimators must still be evaluated for different choices of parameters. For these choices it can be seen that depending on the value of δ , σ and n , the one estimator outperforms the other one.

4. Discussion

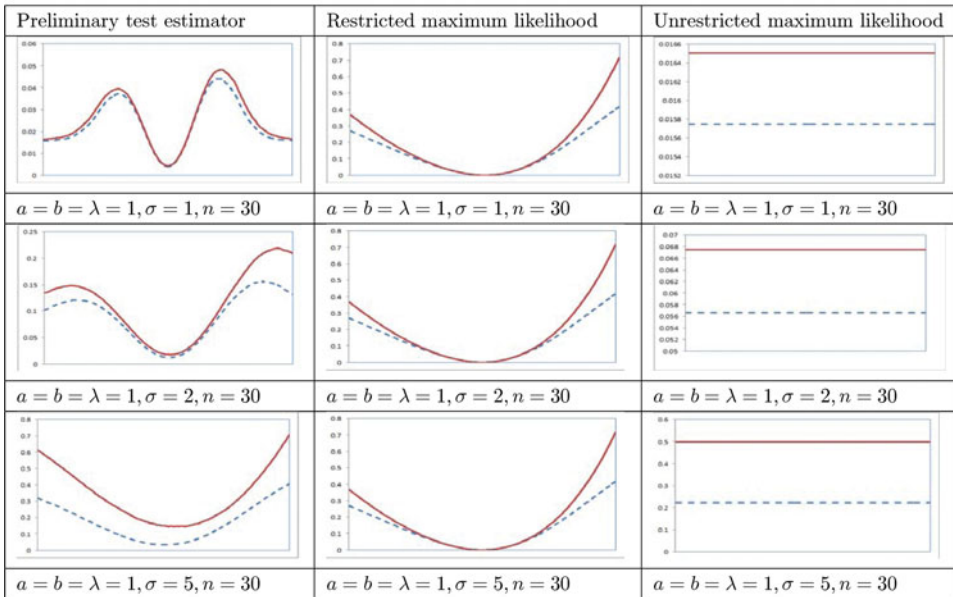
In this article, the PTE and its exact risk function is derived under BLINEX loss, It is then compared in terms of it's risk function, to the risk of the RMLE and UMLE. For specific choices of parameters and sample sizes it can be seen that the PTE is just as efficient as UMLE, these results were obtained when $\delta \rightarrow \pm\infty$ as the sample size increases and by keeping σ constant. When the sample size remains constant and σ increases, the interval where PTE dominates UMLE increases. In other words, when σ increases we rather prefer to use the PTE in our problems.

The FBPTE was also defined in Sec. 3.2 and the PTE was computationally compared to the FBPTE. For specific choices of parameters and sample sizes it can be seen that both estimators perform well under specific conditions and for specific values of δ . Since the feasible Bayes estimator, given by (14), is nothing more than the Bayes estimator under squared error loss, we can not expect that the risk of FBPTE has a special form. It is still unclear how one can find a Bayes estimator under BLINEX loss with stable risk performance, which leaves this problem open for further research. By examining the results obtained, we found that the PTE performs better than FBPTE for some specific parameter choices and vice versa.

The BLINEX and LINEX risk functions are illustrated in the following sketches (see Figs. 12a–c) for specific values of the bounding parameter, λ , the asymmetry parameter b , and the shape parameter a which influences the direction of the penalisation of the delta. In addition to these parameters, different sample sizes were selected and the population variance was also adjusted to determine the effect on the risk functions. In each sketch the

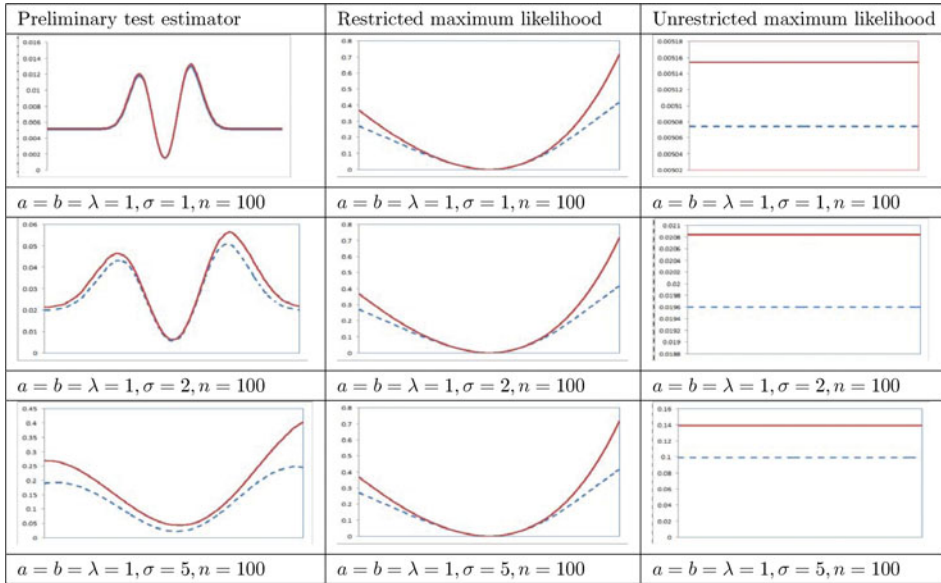


(a)



(b)

Figure 12. a. Risk functions under BLINEX and LINEX loss for different choices of the scale parameter and $n = 10$. b. Risk functions under BLINEX and LINEX loss for different choices of the scale parameter and $n = 30$. c. Risk functions under BLINEX and LINEX loss for different choices of the scale parameter and $n = 100$.



(c)

Figure 12. (continued)

corresponding risk functions for BLINEX and LINEX loss are compared to one another. The BLINEX risk function is represented by the dashed line curve and the LINEX risk function is represented by the solid line curve. All the risk functions are represented as a function of δ , where $-1 \leq \delta \leq 1$. In the first column of the panel the risk functions of the PTE under BLINEX and LINEX loss are given. In the second column the risk functions of the RMLE under BLINEX and LINEX loss are given and in the last column the risk functions of the UMLE under BLINEX and LINEX loss are given.

Even though only specific choices of the parameters a , b , λ , σ , and n were investigated, it can clearly be seen that the BLINEX risk functions represent lower risk than that of the corresponding LINEX risk function across all estimators. These results can be used to motivate why the BLINEX loss function should be used, rather than the LINEX loss function. The risk functions of RMLE remain the same for all choices of σ and n , since the risk function is not a function of these two parameters. The risk functions of UMLE are not a function of δ , therefore the risk is constant over the domain of δ . As shown in Wen and Levy (2001), the BLINEX loss function has the same flexibility as the LINEX, together with the added advantage of being a bounded function, now it is also shown that under BLINEX loss all the estimators performs better when compared to LINEX loss.

All the simulations in this article were done in SAS 9.2.

Acknowledgments

This authors are grateful for all the valuable comments and recommendations made by the anonymous reviewers.

Funding

This work is based upon research supported by the National Research Foundation, South Africa (Grant: Unlocking the future: FA2007043000003).

References

- Arashi, M. (2010). Idea of constructing an image of Bayes action. *Statist. Methodol.* 7:22–29.
- Arashi, M. (2009). Constrained Bayes estimator in elliptical models. *57th Proc Int. Statist. Instit. CPM*, Durban, South Africa.
- Arashi, M., Tabatabaey, S.M.M. (2010). Estimation of the location parameter under LINEX loss function: Multivariate case, *Metrika.* 72(1):51–57.
- Arashi, M., Tabatabaey, S.M. M., Khan, S. (2008). Estimation in multiple regression model with elliptically contoured errors under MLINEX loss. *J. Appl. Probab. Statist.* 3(1):23–35.
- Bancroft, T. A. (1964). Analysis and inference for incompletely specified models involving the use of preliminary test(s) of significance. *Biometrics* 20:427–442.
- Bancroft, T. A. (1944). On biases in estimation due to the use of preliminary tests of significance. *Ann. Mathemat. Statist.* 15(2): 190–204.
- Clarke, J. A. (1986). Some implications of estimating a regression scale parameter after a preliminary test of restrictions. M.Ec. Minor Thesis, Dept. of Econometrics and Operations Research, Monash University.
- Ghari, B., Arashi, M., Tabatabaey, S. M. M. (2009). Estimation of the vector of a multivariate normal model under reflected normal loss. *J. Statist. Res.* 43(1): 41–45.
- Giles, D. E. A. (2002). Preliminary test and bayes estimation of a location parameter under reflected normal loss. In: Wan, A., Chaturvedi, A., Eds., *Handbook of Applied Econometrics and Statistical Inference*. New York: Marcel Dekker, pp. 287–303.
- Giles, J. A., Giles, D. E. A. (1996). Risk of a homoscedasticity pre-test estimator of the regression scale under LINEX loss. *J. Statist. Plann. Infer.* 50: 21–35.
- Judge, G. G., Bock, M. E. (1978). *The Statistical Implications of Pre-test and Stein-rule Estimation in Econometrics*. Amsterdam: North-Holland Publishing Company.
- Kibria, B. M. G., Saleh, A. K. Md. Ehsanes. (2006). Optimum critical value for pre-test estimator. *Commun. Statist. Theor. Meth.* 35(2): 309–319.
- Ohtani, K., Giles, D. E. A., Giles, J. A. (1997). The exact risk performance of a pre-test estimator in a heteroskedastic linear regression model under a balanced loss function. *Econometric Reviews.* 16: 119–130.
- Parsian, A., Kirmani, S. N. U. A. (2002). Estimation under LINEX Loss Function. In: Ullah, A., Wan, A. T. K., Chaturvedi, A., Eds., *Handbook of Applied Econometrics and Statistical Inference*. New York: Marcel Dekker, pp. 53–76.
- Porosiński, Z., Kamińska, A. (2009). On robust Bayesian estimation under some asymmetric and bounded loss function. *Statistics* 43(3): 253–265.
- Saleh, A. K. Md. Ehsanes. (2006). Theory of preliminary test and Stein-type estimation with applications. Hoboken: NJ: Wiley.
- Saleh, A. K. Md. Ehsanes, Kibria, B. M. G. (1993). Performances of some new preliminary test ridge regression estimators and their properties. *Commun. Statist. Theor. Meth.* 22: 2747–2764.
- Varian, H. R. (1975). A Bayes approach to real estate assessment. In: Feinberg, S. E., Zellner, A., Eds., *Studies in Bayesian Econometrics and Statistics in honour of L.J. Savage*. Amsterdam: North Holland, pp. 195–208.
- Wen, D., Levy, S. (2001). BLINEX: A bounded asymmetric loss function with application to Bayesian estimation. *Commun. Statist. Theor. Meth.* 30(1): 147–153.
- Zellner, A. (1977). *An Introduction to Bayesian Inference in Econometrics*. New York: Wiley.