DISCUSSION OF

Joint Maximum Likelihood Estimators for Gutenberg-Richter Parameters λ_0 and β Using Subcatalogs

Manuscript Reference: M. Ordaz and S. Giraldo, *Earthquake Spectra*, vol. 34, no. 1 (February 2018): 301–312.

Petrus J. Vermeulen[a)] and Andrzej Kijko[a)]

INTRODUCTION

Ordaz and Giraldo (2018), further on referred to as O-G, attempt to improve upon the extended Aki-Utsu maximum likelihood estimator and that given by Kijko and Smit (2012) for β and λ. Given a seismic catalog that comprises more than one subcatalog with different levels of completeness (Figure 1), the extended Aki-Utsu estimators are given by the following:

$$\hat{\beta} = \left(\frac{r_1}{\hat{\beta}_1} + \frac{r_2}{\hat{\beta}_2} + \cdots + \frac{r_s}{\hat{\beta}_s}\right)^{-1}$$
$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{s} T_i \exp[-\beta(m_{min}^i - m_{min})]}, \tag{1}$$

where $\hat{\beta}_i$ denotes the classic Aki-Utsu estimator for each subcatalog, $1/\left(m_{min}^i - \langle m \rangle_i\right)$, with m_{min}^i the level of completeness of the i^{th} subcatalog, and m_{min} is the overall chosen minimum value of magnitude taken into consideration. The number r_i is defined as the ratio of the number n_i of events in the i^{th} subcatalog to the total number of events, n, in the entire catalog, that is $r_i = \frac{n_i}{n}$. The estimator $\hat{\lambda}$ denotes the rate of seismicity of the whole catalog, and T_i denotes the time span of the i^{th} subcatalog.

The improvement O-G propose is a joint maximum likelihood estimation of the pair (β, λ). This is indeed an improvement, as it uses not only the marginal likelihoods of β and λ but also their joint (simultaneous) likelihood. In addition, O-G show numerically that this estimation is superior to the separate, marginal maximum likelihood estimators as applied by Kijko and Smit (2012). However, it is interesting to note, as we have discovered, that the equations given by O-G turn out to simply be a special instance in the scheme developed by Kijko and Sellevoll (1989; further on referred to as K-S). If no extreme part of the catalog was used, and one supposed $m_{max} = \infty$, the O-G equations would look exactly the same as those from the scheme developed by K-S. The derivation of the likelihood functions differ slightly from K-S in that O-G use every single interval between consecutive earthquakes, whereas K-S use the total time span of each complete subcatalog. It turns out

[a)] Natural Hazard Center, University of Pretoria, Pretoria, South Africa; Email: u28038020@tuks.co.za (P. J. V.)

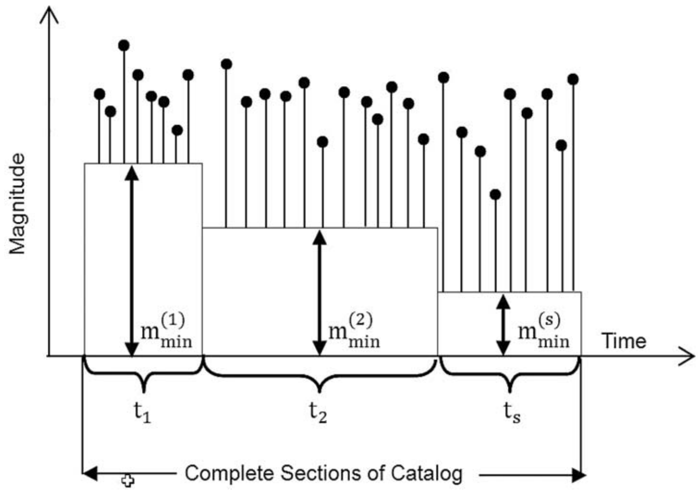

Figure 1. A schematic illustration of a seismic event catalog that can be used in the estimation of the Gutenberg-Richter β-value (after Kijko and Smit 2012).

that these differences in the derivation lead to equivalent likelihood functions. Specifically, this is because the construction of the likelihood function involves the multiplication of exponential terms, having time as a factor in the exponent. In other words, it is of the following form: $\prod_i (\lambda t_i) \exp(-\lambda t_i) = (\lambda t_i)^n \exp(\lambda \sum_i t_i) = (\lambda t_i)^n \exp(\lambda T)$, where T is the time span of a catalog or subcatalog, λ is the Poissonian rate, and n is the number of events. This is to be expected, as the likelihood functions are derived from the same initial distributions of magnitude and interevent time distribution. The attentive reader might note that K-S make use of the likelihood of observing n_i earthquakes in a time T_i [Kijko and Sellevoll (1989); Equation 9], and O-G use the the interevent time distribution. However, from a formal point of view, the equivalence of the descriptions of the Poisson process (as a counting process or a distribution of interevent times) tells us that the outcomes should be equivalent. Recall that the Poisson process as a counting process is characterized by the distribution as follows:

$$P[N(t) = n] = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \qquad (2)$$

On the other hand, the distribution of interevent times is given by the following equation:

$$P[\text{interarrival time} > t] = e^{-\lambda t} \qquad (3)$$

THE LIKELIHOOD EQUATIONS

In this section, the derivation of the likelihood equations of O-D and K-S will be reviewed and discussed in more detail. K-S derive the likelihood equation quite briefly and compactly and do not state the end result, as the reader is expected to be well acquainted with the procedures the authors are following. These equations will be derived in added detail here, and the result will be given explicitly. O-G give the full derivation up to the likelihood equation, but we will restate it here for the sake of completeness and comparison. To simplify and facilitate easy comprehension of the derivations, some notation borrowed from O-G is introduced here as follows:

$$T_j^* = T_j e^{-\hat{\beta}(m_{0j}-m_0)} \tag{4}$$

$$T^* = \sum_{j=1}^{L} T_j^* \tag{5}$$

$$Q = \sum_{j=1}^{s} n_j(m_{0j} - m_0) \tag{6}$$

$$S = \sum_{j=1}^{s} S_j \tag{7}$$

$$S_j = \sum_{i=1}^{n_j} (m_{i,j} - m_0) \tag{8}$$

APPROACH FOLLOWED BY KIJKO AND SELLEVOLL (1989)

Following K-S, let us assume that earthquake magnitudes follow the doubly truncated Gutenberg-Richter distribution. The likelihood function obtained from this distribution gives a likelihood in terms of β, and for a specific subcatalog, this is as follows:

$$\mathcal{L}_{i\beta}^{K-S} = \beta^{n_i} \frac{e^{-\beta \sum_j^{n_i} m_{ij}}}{e^{-\beta m_i}} = \beta^{n_i} e^{-\beta \sum_j^{n_i} (m_{ij} - m_i)} = \beta^{n_i} e^{-\beta S_i} \tag{9}$$

Not providing much motivation, as it is assumed that the reader is familiar with the Poisson process, the likelihood function involving λ and β, which is, in fact, merely the probability mass function for a time period T_i, takes the following form:

$$\mathcal{L}_{i\lambda}^{K-S} = \frac{T_i^{n_i}}{n!} \cdot e^{(-\lambda e^{-\beta \Delta_i} T_i)} (\lambda e^{-\beta \Delta_i} T_i)^{n_i} = \frac{T_i^{n_i}}{n!} \cdot e^{(-\lambda e^{-\beta \Delta_i} T_i)} \lambda^{n_i} e^{-\beta n_i \Delta_i} \tag{10}$$

where

$$\lambda_i = \lambda(1 - F(m_i)) = \frac{\lambda e^{-\beta m_i}}{e^{-\beta m_{min}}} = \lambda e^{-\beta \Delta_i}, \tag{11}$$

and

$$F(x) = \frac{e^{-\beta m_{min}} - e^{-\beta x}}{e^{-\beta m_{min}}} \text{ and } \Delta_i = m_i - m_{min}. \tag{12}$$

The joint likelihood function is obtained by combining functions of Equations 9 and 10 over the s complete subcatalogs:

$$\mathcal{L}_{i\beta}^{K-S}\mathcal{L}_{i\lambda}^{K-S} = \prod_{i}^{s} \beta^{n_i} e^{-\beta S_i} \frac{T_i}{n!} \cdot e^{\left(-\lambda e^{-\beta \Delta_i} T_i\right)} \lambda^{n_i} e^{-\beta n_i \Delta_i} = \frac{T}{n!} \cdot \beta^n \lambda^n e^{-\beta S} e^{-\beta Q} e^{-\lambda T^*}, \tag{13}$$

Note that coefficient $\frac{T}{n!}$ does not have any effect on the values of β and λ; therefore, the likelihood function might well be written as:

$$\mathcal{L}_{\beta}\mathcal{L}_{\lambda} = \beta^n \lambda^n e^{-\beta S} e^{-\beta Q} e^{-\lambda T^*} \tag{14}$$

where T^* is as defined in the previous section.

APPROACH FOLLOWED BY ORDAZ AND GIRALDO (2018)

Consider the probability density function of interarrival times as follows:

$$f_t = \lambda e^{-\lambda t} \tag{15}$$

Therefore, the likelihood of having n_i intervals between the events in a subcatalog of duration T_i is given by the following:

$$\mathcal{L}_{i\lambda}^{O-G} = \prod_{j=0}^{n_i} \lambda_i e^{-\lambda_i t_{j,i}} = (\lambda_i)^{n_i} e^{-\lambda_i \sum_{j=0}^{n_i} t_{j,i}} = (\lambda_i)^{n_i} e^{-\lambda_i T_i} \tag{16}$$

where $t_{j,i}$ denotes the j^{th} interevent time interval in the i^{th} subcatalog. Note that we are looking at the likelihood of observing n time intervals. This is, in fact, a counting process and turns out to be the counting process characterizing the Poisson process. This will reveal similarity (or stronger even, equivalence) of the likelihood function derived by Ordaz and Giraldo (2018) and Kijko and Sellevoll (1989). Note that:

$$\lambda_i = \lambda[1 - F_M(m_{min_i})] = \lambda e^{-\beta\left(m_{min}^{(i)} - m_0\right)} \tag{17}$$

Then, substituting Equation 17 in Equation 16 gives:

$$\begin{aligned}\mathcal{L}_{i\lambda}^{O-G} &= \prod_{j=0}^{n_i} \lambda_i e^{-\lambda_i t_{j,i}} = (\lambda_0)^{n_i} e^{-\beta\left(m_{min}^{(i)} - m_0\right)} e^{-\lambda_0 e^{-\beta\left(m_{min}^{(i)} - m_0\right)} T_i} \\ &= (\lambda_0)^{n_i} e^{-\beta \Delta_i} e^{-\lambda_0 e^{-\beta \Delta_i} T_i} \\ &= (\lambda_0)^{n_i} e^{-\beta \Delta_i} e^{-\lambda_0 T_i^*}\end{aligned} \tag{18}$$

For a given β value, the likelihood of observing n_i events in a subcatalog is given by the following:

$$\begin{aligned}\mathcal{L}_{i\beta}^{O-G} &= \prod_{j=1}^{n_i} \beta e^{-\beta\left(m_{j,i} - m_{min}^{(i)}\right)} = \beta^{n_i} e^{-\beta \sum_{j=1}^{n_i}\left(m_{j,i} - m_{min}^{(i)}\right)} \\ &= \beta^{n_i} e^{-\beta S_i}\end{aligned} \tag{19}$$

Combining the likelihood functions and considering the entire catalog gives the following:

$$\mathcal{L}_{i\beta}^{O-G}\mathcal{L}_{i\lambda}^{O-G} = \prod_{i}^{s} \beta^{n_i} e^{-\beta S_i} (\lambda_0)^{n_i} e^{-\beta \Delta_\mathrm{i}} e^{-\lambda_0 T_i^*}$$
$$= \beta^n (\lambda_0)^n e^{-\beta S} e^{-\beta Q} e^{-\lambda T^*} \tag{20}$$

MAXIMUM LIKELIHOOD EQUATIONS

Thus we see that the two likelihood functions, Equations 14 and 20, are the same. Kijko and Sellevoll (1989) give a general solution, which is rather cumbersome to work with, but they note that the equations are derived from maximizing the log-likelihood functions by setting the derivatives *of the* derivatives of the log-likelihood function with respect to λ and β equal to zero. Neither K-S nor O-G show the calculations explicitly. The derivation is not extensive and is given here. First, the log-likelihood function is:

$$\log(\mathcal{L}_\beta \mathcal{L}_\lambda) = \log(\beta^n \lambda^n e^{-\beta S} e^{-\beta Q} e^{-\lambda T^*})$$
$$= n \log(\beta) + n \log(\lambda) - \beta(Q+S) - \lambda T^* \tag{21}$$

where $\log(\cdot)$ denotes a natural logarithm. The derivative of the log-likelihood function with respect to λ is:

$$\frac{\partial}{\partial \lambda} \log(\mathcal{L}_\beta \mathcal{L}_\lambda) = \frac{\partial}{\partial \lambda}(n \log(\beta) + n \log(\lambda) - \beta(Q+S) - \lambda T^*)$$
$$= \frac{n}{\lambda} + T^*, \tag{22}$$

and the derivative of the log-likelihood function with respect to β is:

$$\frac{\partial}{\partial \beta} \log(\mathcal{L}_\beta \mathcal{L}_\lambda) = \frac{\partial}{\partial \beta}\left(n \log(\beta) + n \log(\lambda) - \beta(Q+S) - \lambda T^*\right)$$
$$= \frac{n}{\beta} - (Q+S) - \frac{\partial}{\partial \beta}\left(\lambda T^*\right)$$
$$= \frac{n}{\beta} - (Q+S) - \lambda \sum_{i}^{s} (T_i e^{-\beta \Delta_\mathrm{i}})$$
$$= \frac{n}{\beta} - (Q+S) - \frac{\partial}{\partial \beta} \lambda \sum_{i}^{s} \left(T_i e^{-\beta \Delta_i} (-\Delta_\mathrm{i})\right)$$
$$= \frac{n}{\beta} - (Q+S) + \lambda \sum_{i}^{s} \left(T_i e^{-\beta \Delta_i} (\Delta_\mathrm{i})\right)$$
$$= \frac{n}{\beta} - (Q+S) + \lambda \sum_{i}^{s} (T_i^* \Delta_\mathrm{i}) \tag{23}$$

Then, equating these derivatives to zero, the following is obtained:

$$\frac{\partial}{\partial\lambda}\log\left(\mathcal{L}_\beta\mathcal{L}_\lambda\right) = 0$$

$$\frac{n}{\lambda} + T^* = 0 \tag{24}$$

and

$$\frac{\partial}{\partial\beta}\log\left(\mathcal{L}_\beta\mathcal{L}_\lambda\right) = 0$$

$$\frac{n}{\beta} - (Q + S) + \lambda\sum_{i}^{s}\left(T_i^*\Delta_\mathrm{i}\right) = 0 \tag{25}$$

Substituting the estimate of λ obtained from Equation 24 in Equation 25, the nonlinear system of equations given by O-G is obtained:

$$\begin{cases} \frac{n}{\lambda} + T^* = 0 \\ \frac{n}{\beta} - (Q + S) + \frac{n}{T^*}\sum_{i}^{s}\left(T_i^*\Delta_\mathrm{i}\right) = 0 \end{cases} \tag{26}$$

Considering that K-S also start from the same likelihood equation and use the same method of maximizing the likelihood, the two solutions must be the same. This shows that the work of O-G is a special case of the work of K-S.

CONCLUSIONS

Ordaz and Giraldo (2018) give an improved (joint) maximum likelihood estimator for the parameters β and λ, compared with the extended Aki-Utsu estimator developed by Kijko and Smit (2012), in Equation 1. It was shown here that the joint maximum likelihood estimator of O-G is simply a special case of the joint maximum likelihood equation given by Kijko and Sellevoll (1989).

ACKNOWLEDGMENTS

This work was supported by NRF Incentive Funding (103724:2017), NRF THRIP MMI Holdings (96412:2018), and NRF Grant Number SFH13081828987.

REFERENCES

Kijko, A., and Sellevoll, M. A., 1989. Estimation of earthquake hazard parameters from incomplete data files. Part I. Utilization of extreme and complete catalogs with different threshold magnitudes, *Bulletin of the Seismological Society of America* **79**, 645–654.

Kijko, A., and Smit, A., 2012. Extension of the Aki-Utsu b-value estimator for incomplete catalogs, *Bulletin of the Seismological Society of America* **102**, 1283–1287.

Ordaz, M., and Giraldo, S., 2018. Joint maximum likelihood estimators for Gutenberg-Richter parameters λ_0 and β using subcatalogs, *Earthquake Spectra* **34**, 301–312.

(Received 17 July 2018; Accepted 19 October 2018)