*Cuadernos de Ingeniería. Nueva Serie*. Revista de la Facultad de Ingeniería
de la Universidad Católica de Salta (Argentina), Volumen 15, 2024

# Analysis of Gutenberg-Richter $b$-value and $m_{max}$ Part IV: New technique to estimate the parameters $m_{max}$, $m_{min}$ and $b$-value

## Análisis del parámetro $b$ y $m_{max}$ del Modelo de Gutenberg-Richter Parte IV: Nueva técnica para estimar los parámetros $m_{max}$, $m_{min}$ y $b$

**Mika Haarala**[1]

*Ingeniería Sísmica / artículo científico*

Citar: Haarala, M. (2024). Analysis of Gutenberg-Richter $b$-value and $m_{max}$. Part IV: New technique to estimate the parameters $m_{max}$, $m_{min}$ and $b$-value. *Cuadernos de Ingeniería* (15). http://revistas.ucasal.edu.ar

*Recibido: setiembre/2024*
*Aceptado: diciembre/2024*

### Abstract

In this paper, we carried out an analysis of the Gutenberg-Richer distribution function, which provides a new technique to estimate the expected values. This estimator, which we referred to as the expected value curve estimator, offers a relationship between the Order Statistic and the Statistic of the Maximums. Also, it was observed that the sample size is an unknown parameter in the real seismic catalogues. In addition, we demonstrate an algebraic solution for the Gutenberg-Richter distribution problem, where all the unknown parameters $\beta$, $m_{max}$ and $m_{min}$ are estimated using four estimators of the expected values. We also discuss the role of the negative $b$-value.

**Keywords**: General Gutenberg-Richter distribution function, Kijko-Sellevoll functions, estimators for the $m_{max}$, $m_{min}$ and $b$-value.

### Resumen

En este artículo, llevamos a cabo un análisis de la función de distribución de Gutenberg-Richter, lo que resultó en una nueva técnica para estimar el valor esperado. Este estimador, al que denominamos curva esperada, da una relación entre el orden estadístico y el máximo estadístico. Además, se observa que la medida de la muestra es un parámetro desconocido en los catálogos sísmicos reales. Por otra parte, demostramos una solución algebraica para el problema de la función de distribución de Gutenberg-Richter, donde todos los parámetros $\beta$, $m_{max}$ y $m_{min}$ son estimados haciendo uso de cuatro estimadores para los valores esperados. Por último, analizamos el papel del valor negativo de $b$.

**Palabras clave**: función general de distribución de Gutenberg-Richter, funciones Kijko-Sellevoll, estimadores de $m_{max}$, $m_{min}$ y $b$.

[1] Investigador Independiente

## 1. Introduction

The article series "Analysis of Gutenberg-Richter $b$-value and $m_{\max}$" has reached to a fourth part. This part IV (originally intended to be part III of the article series) originated precisely because in the development of the theory the need arose to consider a negative value for beta, which was analyzed in part III. Negative $b$ sounds illogical in the field of seismology, but the theory which contradicts what was established, surprisingly delivers results that are in good agreement with the real physical world. The first paper further developed the theory in order to include subcatalogues of different sizes and their expected values (Haarala and Orosco, 2016a). The classical estimators are mostly from the subcatalogue size 1, but what do we really know about the generalized estimators for subcatalogue sizes above 1 and their applications?

From the beginning of the series of these papers, the double truncated Gutenberg-Richter (GR) distribution function was analyzed (Haarala and Orosco, 2016a, 2016b, 2019). The definition of the GR model was incomplete, so it was necessary to extend to the General Gutenberg-Richter (GGR) probability density function (PDF) (Haarala, 2021):

$$f(m)=\begin{cases}\dfrac{\beta\exp\left[-\beta\left(m-m_{\min}\right)\right]}{1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]}, & \text{for } m_{\min}\leq m\leq m_{\max}\wedge\beta\neq 0,\\ \dfrac{1}{m_{\max}-m_{\min}}, & \text{for } m_{\min}\leq m\leq m_{\max}\wedge\beta=0,\\ 0, & \text{for } m\notin\left[m_{\min},m_{\max}\right],\end{cases}\tag{1}$$

where $\beta=b\log(10)$. The reason for this definition is the discontinuity of the term

$$\frac{\beta\exp\left[-\beta\left(m-m_{\min}\right)\right]}{1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]}$$

at $\beta=0$. Without this extended definition we must assume $\beta>0$ (or more generally $\beta\neq 0$). According to Part III (Haarala, 2021), the definition of GGR holds for $-\infty\leq\beta\leq\infty$ and $-\infty\leq m_{\min}<m_{\max}\leq\infty$. A cumulative distribution function (CDF) of GGR can be written as

$$F_M(m)=\begin{cases}0, & \text{for } m<m_{\min},\\ \dfrac{1-\exp\left[-\beta\left(m-m_{\min}\right)\right]}{1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]}, & \text{for } m_{\min}\leq m<m_{\max}\wedge\beta\neq 0,\\ \dfrac{m-m_{\min}}{m_{\max}-m_{\min}} & \text{for } m_{\min}\leq m<m_{\max}\wedge\beta=0,\\ 1, & \text{for } m\geq m_{\max}.\end{cases}\tag{2}$$

These definitions allow handling also negative $\beta$-values.

Let $M_1, M_2, \ldots, M_N \in [m_{\min}, m_{\max}]$ be a set of random variables from the catalogue $C_N$ of size $N$. Let $C_n$ be a subcatalogue of $C_N$ such that $C_n \subseteq C_N$ for $1 \le n \le N$. Now the maximum function $\max\{M_1, M_2, \ldots, M_n\}$ has a CDF (Haarala, 2021)

$$F_{M_n}(m) = \begin{cases} 0, & \text{for } m < m_{\min}, \\ \left[F_M(m)\right]^n & \text{for } m_{\min} \le m < m_{\max}, \\ 1, & \text{for } m_{\max} \le m. \end{cases} \tag{3}$$

A Kijko-Sellevoll function 1 (KS-1) is defined as (Haarala and Orosco, 2016a)

$$f_n^{KS-1}(x) = \sum_{k=1}^{\infty} \frac{\left(1 - \exp[-x]\right)^k}{k+n} \tag{4}$$

and a Kijko-Sellevoll function 2 (KS-2) as (Haarala and Orosco, 2016a)

$$f_n^{KS-2}(x) = \sum_{k=1}^{\infty} \frac{n}{k} \frac{\left(1 - \exp[-x]\right)^k}{k+n} \tag{5}$$

for $x > -\log(2)$. We can write the expected value of the maximum $M_n$ (at a subcatalogue size $n$) of the catalogue as

$$\begin{aligned} E(M_n) &= m_{\max} - \frac{1}{\beta} f_n^{KS-1}\left(\beta\left(m_{\max} - m_{\min}\right)\right) \\ &= m_{\min} + \frac{1}{\beta} f_n^{KS-2}\left(\beta\left(m_{\max} - m_{\min}\right)\right). \end{aligned} \tag{6}$$

We call this the Expected Value Curve (EVC). The relationship between $f_n^{KS-1}(x)$ and $f_n^{KS-2}(x)$ is

$$x = f_n^{KS-1}(x) + f_n^{KS-2}(x). \tag{7}$$

In our work, the number $n \in \mathbb{N}$ is an index of the event. In any case, the formulae (4)-(7) are true when $n \in \mathbb{R}_+$ (Haarala and Orosco, 2019). The EVC is a continuously increasing function. In this case, when the KS function is continuous it is used as a variable $\eta$ instead of $n$.

## 2. Expected Value Curve Estimator

In our previous work (Haarala and Orosco, 2016b), we divided the catalogue $C_N$ into the subcatalogues $C_{n:k}$, $k = 1, 2, \ldots, \lfloor N/n \rfloor$, $n = 1, 2, \ldots, N$, where $\lfloor \cdot \rfloor$ is a floor function i.e., a maximum integer value $m$ such that $m \leq N/n$. The mean value estimator for the expected value can be written as

$$\hat{E}^*\left(M_n\right) = \frac{\sum_{k=1}^{\lfloor N/n \rfloor} m_{(n):k}}{\lfloor N/n \rfloor} = \bar{m}_{(n} \tag{8}$$

where $m_{(n):k} \in C_{n:k}$ is a maximum observed value in the subcatalogue $n:k$ ($n$ means how many elements belong to the subcatalogue and $k$ is an index of the subcatalogue, meanwhile the magnitude $m_i \in C_N$ can belong to at most one subcatalogue at the same time and there could be events that do not belong to any subcatalogue). For example, in the case $n = 1$ it is $\hat{E}^*\left(M_{(1)}\right) = \bar{m}_{(1)} = \bar{m}$, which is a mean value of all the magnitudes in catalogue $C_N$, or in the case $n = N$, we have $\hat{E}^*\left(M_N\right) = \bar{m}_{(N)} = m_{(N)}$, which is a maximum observed value in catalogue $C_N$. When $1 < n < N$ we can choose the sub-catalogues randomly, but in such a way that each event is used at most once. A problem with this estimator is that there is no guarantee for the order of the estimators: $\bar{m}_{(1)} \leq \bar{m}_{(2)} \leq \ldots \leq \bar{m}_{(N)}$.

To be clear, in the Order Statistic a set of values $m_1, m_2, \ldots, m_N$ is ordered as $m_{(1)} \leq m_{(2)} \leq \ldots \leq m_{(N)}$. So, the maximum is $m_{(N)}$. In our case, all subcatalogues are the same size $n$, so symbolically their maximum is $m_{(n)}$, which depends on the subcatalogue. The mean is calculated from those maximums. The symbol $\bar{m}_{(n)}$means that it is necessary to take first the order (here it is the maximum from the subcatalogue) and then the mean, but it doesn't suggest that first the mean is taken, and then it is essential to order the values. The estimator (8) doesn't order the estimates.

Let's see an example. Suppose that we have a catalogue $C_4 = \{m_1, m_2, m_3, m_4\}$. Taking randomly the magnitudes, we could get the subcatalogues

$$\begin{aligned}
&C_{1:1} = \{m_3\}, \quad C_{1:2} = \{m_1\}, \quad C_{1:3} = \{m_4\}, \quad C_{1:4} = \{m_2\}, \\
&C_{2:1} = \{m_4, m_1\}, \quad C_{2:2} = \{m_2, m_3\}, \\
&C_{3:1} = \{m_3, m_1, m_4\}, \\
&C_{4:1} = \{m_2, m_4, m_3, m_1\}.
\end{aligned}$$

It's worth mentioning that there is only one subcatalogue of size 3, so no event $m_2$ is selected (in this case). The maximum observed values of the subcatalogues are now

$$\begin{aligned}
&m_{(1):1} = \max\{m_3\}, \quad m_{(1):2} = \max\{m_1\}, \quad m_{(1):3} = \max\{m_4\}, \quad m_{(1):4} = \max\{m_2\},\\
&m_{(2):1} = \max\{m_4, m_1\}, \quad m_{(2):2} = \max\{m_2, m_3\},\\
&m_{(3):1} = \max\{m_3, m_1, m_4\},\\
&m_{(4):1} = \max\{m_2, m_4, m_3, m_1\}.
\end{aligned}$$

providing estimators as

$$\bar{m}_{(1)} = \frac{1}{4}\sum_{k=1}^{4} m_{(1):k} = \bar{m}, \quad \bar{m}_{(2)} = \frac{1}{2}\sum_{k=1}^{2} m_{(2):k}, \quad \bar{m}_{(3)} = m_{(3):1}, \quad \bar{m}_{(4)} = m_{(4):1} = m_{(4)}.$$

Any set of randomly selected subcatalogues can provide a different result except at size 1 and $N$. If we repeat this process 1000 or 100000 times, we could see that it converges to some mean, an expected value.

Let's consider another example. Suppose that $m_{(k)} = k$, $k = 1,2,3,4$. The sets of size $n = 2$ are $\{1,2\}$, $\{1,3\}$, $\{2,3\}$, $\{1,4\}$, $\{2,4\}$, $\{3,4\}$. In the case of $n = 3$, the set are $\{1,2,3\}$, $\{1,2,4\}$, $\{1,3,4\}$, $\{2,3,4\}$ and $n = 4$ there is only $\{1,2,3,4\}$. There is no set $\{2,2\}$, because our example has only an event with value 2. Now the weighted means are

$$\begin{aligned}
\bar{m}_{(1)} &= \frac{1}{4}\cdot 1 + \frac{1}{4}\cdot 2 + \frac{1}{4}\cdot 3 + \frac{1}{4}\cdot 4 = 2\frac{1}{2},\\
\bar{m}_{(2)} &= \frac{1}{6}\cdot 2 + \frac{2}{6}\cdot 3 + \frac{3}{6}\cdot 4 = 3\frac{1}{3},\\
\bar{m}_{(3)} &= \frac{1}{4}\cdot 3 + \frac{3}{4}\cdot 4 = 3\frac{3}{4},\\
\bar{m}_{(4)} &= 4.
\end{aligned}$$

In general terms, suppose that we have ordered events $m_{(1)} \le m_{(2)} \le \cdots \le m_{(N)}$. Given that the order of events doesn't matter for the maximum function, we can choose

$$\binom{N}{n} = \frac{N!}{n!(N-n)!} = \binom{N}{N-n} \tag{9}$$

sets of size $n$. Let $p$ be an integer such that $n \le p \le N$. Now the observed maxima can have values $m_{(n)} \le \cdots \le m_{(p)} \le \cdots \le m_{(N)}$. If we now have the event $m_{(p)}$, that is the largest member of the set of size $p$, so there are $p-1$ elements, from which we can choose the rest of the $n-1$ members of the subset of size $n$ with maximum observed value $m_{(p)}$. Thus, we have

$$\binom{p-1}{n-1}$$

sets of size $n$, where the maximum observed value is $m_{(p)}$. The EVC estimator, as we call it, is now

$$\hat{E}(M_n) = \sum_{p=n}^{N} \frac{\binom{p-1}{n-1}}{\binom{N}{n}} m_{(p)} = \binom{N}{n}^{-1} \sum_{p=n}^{N} \binom{p-1}{n-1} m_{(p)} = \bar{m}_{(n)}. \tag{10}$$

Appendix A provides a method to compute (10) with MATLAB in a quite powerful way. For example, it is easy to find the estimators in the cases $n = 1$ and $n = N$:

$$\hat{E}(M_1) = \bar{m}_{(1)} = \binom{N}{1}^{-1} \sum_{p=1}^{N} \binom{p-1}{0} m_{(p)} = \frac{1}{N} \sum_{p=1}^{N} m_{(p)} = \bar{m}\,,$$

$$\hat{E}(M_N) = \bar{m}_{(N)} = \binom{N}{N}^{-1} \sum_{p=N}^{N} \binom{p-1}{N-1} m_{(p)} = m_{(N)}\,.$$

The first of the above expressions is a classical mean estimator, and the second one is a maximum, also known as the Pisarenko estimator (Pisarenko et al., 1996; Haarala and Orosco, 2016a).

Table 1 shows the example of the factors in cases $N = 3$ and $N = 4$.

**Table 1.** Example of the EVC estimator

| $n$ | $N = 3$ | $N = 4$ |
|---|---|---|
| 1 | $\bar{m}_{(1)} = \frac{1}{3} m_{(1)} + \frac{1}{3} m_{(2)} + \frac{1}{3} m_{(3)}$ | $\bar{m}_{(1)} = \frac{1}{4} m_{(1)} + \frac{1}{4} m_{(2)} + \frac{1}{4} m_{(3)} + \frac{1}{4} m_{(4)}$ |
| 2 | $\bar{m}_{(2)} = \frac{1}{3} m_{(2)} + \frac{2}{3} m_{(3)}$ | $\bar{m}_{(2)} = \frac{1}{6} m_{(2)} + \frac{2}{6} m_{(3)} + \frac{3}{6} m_{(4)}$ |
| 3 | $\bar{m}_{(3)} = m_{(3)}$ | $\bar{m}_{(4)} = \frac{1}{4} m_{(3)} + \frac{3}{4} m_{(4)}$ |
| 4 | | $\bar{m}_{(4)} = m_{(4)}$ |

## 3. Some Properties of the EVC Estimator

We can see from the recurrence formula (Abramowitz et al., 1972) that

$$\begin{aligned}\sum_{p=n}^{N}\binom{p-1}{n-1} &= \binom{n-1}{n-1}+\binom{n}{n-1}+\cdots+\binom{N-1}{n-1}\\ &= \binom{n-1}{0}+\binom{n}{1}+\cdots+\binom{N-1}{N-n} = \binom{N}{N-n} = \binom{N}{n}.\end{aligned}$$

Now, it is possible to find the next properties for the EVC estimator (10):

i) If $m_{(n)}=\cdots=m_{(N)}$, then

$$\bar{m}_{(n)} = m_{(N)}\binom{N}{n}^{-1}\sum_{p=n}^{N}\binom{p-1}{n-1} = m_{(N)} = \bar{m}_{(N)}\ . \tag{11}$$

Hence, $\bar{m}_{(n)}=\cdots=\bar{m}_{(N)}$.

ii) If $m_{(n)}=\cdots=m_{(k-1)}<m_{(k)}=\cdots=m_{(N)}$, then

$$\begin{aligned}\bar{m}_{(n)} &= \binom{N}{n}^{-1}\left\{\sum_{p=n}^{k-1}\binom{p-1}{n-1}m_{(k-1)}+\sum_{p=k}^{N}\binom{p-1}{n-1}m_{(k)}\right\}\\ &= \binom{N}{n}^{-1}\left\{\binom{k-1}{n}m_{(k-1)}+\left(\binom{N}{n}-\binom{k-1}{n}\right)m_{(k)}\right\}\\ &= m_{(k)}-\binom{N}{n}^{-1}\binom{k-1}{n}\underbrace{\left(m_{(k)}-m_{(k-1)}\right)}_{>0}.\end{aligned}$$

Because

$$\begin{aligned}-\binom{N}{n}^{-1}\binom{k-1}{n} &= -\frac{k-1}{N}\frac{k-2}{N-1}\cdots\frac{k-n}{N+1-n}\\ &= -\frac{k-1}{N}\frac{k-2}{N-1}\cdots\frac{k-n}{N+1-n}\frac{k-(n+1)}{N+1-(n+1)}\frac{N+1-(n+1)}{k-(n+1)}\\ &= -\binom{N}{n+1}^{-1}\binom{k-1}{n+1}\frac{N+1-(n+1)}{k-(n+1)}\\ &< -\binom{N}{n+1}^{-1}\binom{k-1}{n+1}\end{aligned}$$

where $n<k\leq N$. Therefore $\bar{m}_{(n)}<\cdots<\bar{m}_{(k-1)}<\bar{m}_{(k)}=\cdots=\bar{m}_{(N)}$.

iii) Suppose, that $m_{(n)} \le \cdots \le m_{(k-1)} < m_{(k)} = \cdots = m_{(N)}$, $n < k \le N$. The binomial factors can be written as

$$\binom{N}{n} = \frac{N!}{n!(N-n)!} = \frac{n+1}{N-n}\frac{N!}{(n+1)!(N-n-1)!} = \frac{n+1}{N-n}\binom{N}{n+1}$$

and

$$\begin{aligned}\binom{p-1}{(n+1)-2} &= \frac{(p-1)!}{((n+1)-2)!((p-1)-((n+1)-2))!}\\ &= \frac{n}{p-n}\frac{(p-1)!}{((n+1)-1)!((p-1)-((n+1)-1))!}\\ &= \frac{n}{p-n}\binom{p-1}{(n+1)-1},\end{aligned}$$

where $n < p \le N$. Now, we can write

$$\begin{aligned}\overline{m}_{(n)} &= \binom{N}{n}^{-1}\left\{m_{(n)} + \sum_{p=n+1}^{N}\binom{p-1}{n-1}m_{(p)}\right\}\\ &= \frac{N-n}{n+1}\binom{N}{n+1}^{-1}\left\{m_{(n)} + \sum_{p=n+1}^{N}\frac{n}{p-n}\binom{p-1}{(n+1)-1}m_{(p)}\right\}\\ &= \binom{N}{n+1}^{-1}\left\{\frac{N-n}{n+1}m_{(n)} + \sum_{p=n+1}^{N}\frac{N-n}{n+1}\frac{n}{p-n}\binom{p-1}{(n+1)-1}m_{(p)}\right\}\\ &= \binom{N}{n+1}^{-1}\left\{\frac{N-n}{n+1}m_{(n)} + \sum_{p=n+1}^{N}\left(\frac{N-n}{n+1}\frac{n}{p-n}-1\right)\binom{p-1}{(n+1)-1}m_{(p)} + \sum_{p=n+1}^{N}\binom{p-1}{(n+1)-1}m_{(p)}\right\}\\ &= \binom{N}{n+1}^{-1}\left\{\frac{N-n}{n+1}m_{(n)} + \sum_{p=n+1}^{N}\left(\frac{N-n}{n+1}\frac{n}{p-n}-1\right)\binom{p-1}{(n+1)-1}m_{(p)}\right\} + \overline{m}_{(n+1)}\end{aligned} \tag{12}$$

If it is set $m_{(n)} = m_{(n+1)} = \cdots = m_{(N)}$, then using the property (i) we get

$$m_{(N)} = \binom{N}{n+1}^{-1}\left\{\frac{N-n}{n+1} + \sum_{p=n+1}^{N}\left(\frac{N-n}{n+1}\frac{n}{p-n}-1\right)\binom{p-1}{(n+1)-1}\right\}m_{(N)} + m_{(N)}$$

which shows that

$$0 = \frac{N-n}{n+1} + \sum_{p=n+1}^{N}\left(\frac{N-n}{n+1}\frac{n}{p-n}-1\right)\binom{p-1}{(n+1)-1}. \tag{13}$$

Since

$$0 \le \frac{N-n}{n+1}\frac{n}{p-n} - 1 \quad \Leftrightarrow \quad p \le \frac{(N+1)n}{n+1},$$

$$\frac{N-n}{n+1} \ge 0, \quad \binom{p-1}{(n+1)-1} > 0, \quad n+1 \le p \le N,$$

and by setting

$$q = \left\lfloor \frac{(N+1)n}{n+1} \right\rfloor \in \mathbb{N}, \quad 1 \le q \le N,$$

we can separate the positive and negative factors of (13) as

$$0 \le \frac{N-n}{n+1} + \sum_{p=n+1}^{q} \left( \frac{N-n}{n+1}\frac{n}{p-n} - 1 \right) \binom{p-1}{(n+1)-1},$$
$$0 \le \sum_{p=q+1}^{N} \left( 1 - \frac{N-n}{n+1}\frac{n}{p-n} \right) \binom{p-1}{(n+1)-1}.$$

Hence, the equation (13) can write as

$$\frac{N-n}{n+1} + \sum_{p=n+1}^{q} \left( \frac{N-n}{n+1}\frac{n}{p-n} - 1 \right) \binom{p-1}{(n+1)-1} = \sum_{p=q+1}^{N} \left( 1 - \frac{N-n}{n+1}\frac{n}{p-n} \right) \binom{p-1}{(n+1)-1}.$$

To complete the proof, it holds for the term in (12)

$$\begin{aligned}
&\frac{N-n}{n+1} m_{(n)} + \sum_{p=n+1}^{N} \left( \frac{N-n}{n+1}\frac{n}{p-n} - 1 \right) \binom{p-1}{(n+1)-1} m_{(p)} \\
&= \frac{N-n}{n+1} m_{(n)} + \sum_{p=n+1}^{q} \left( \frac{N-n}{n+1}\frac{n}{p-n} - 1 \right) \binom{p-1}{(n+1)-1} m_{(p)} - \sum_{p=q+1}^{N} \left( 1 - \frac{N-n}{n+1}\frac{n}{p-n} \right) \binom{p-1}{(n+1)-1} m_{(p)} \\
&\overset{(a)}{\le} \left[ \frac{N-n}{n+1} + \sum_{p=n+1}^{q} \left( \frac{N-n}{n+1}\frac{n}{p-n} - 1 \right) \binom{p-1}{(n+1)-1} \right] m_{(q)} - \left[ \sum_{p=q+1}^{N} \left( 1 - \frac{N-n}{n+1}\frac{n}{p-n} \right) \binom{p-1}{(n+1)-1} \right] m_{(q+1)} \\
&\overset{(b)}{\le} 0.
\end{aligned}$$

The strict inequality

$$\frac{N-n}{n+1} m_{(n)} + \sum_{p=n+1}^{q} \left( \frac{N-n}{n+1}\frac{n}{p-n} - 1 \right) \binom{p-1}{(n+1)-1} m_{(p)} < \left[ \frac{N-n}{n+1} + \sum_{p=n+1}^{q} \left( \frac{N-n}{n+1}\frac{n}{p-n} - 1 \right) \binom{p-1}{(n+1)-1} \right] m_{(q)},$$

in the case of $m_{(n)} \le \cdots \le m_{(k-1)} < m_{(k)} = \cdots = m_{(q)} = \cdots = m_{(N)}$, implies the strict inequality at (a). The same way, the strict inequality

$$\left[\sum_{p=q+1}^{N}\left(1-\frac{N-n}{n+1}\frac{n}{p-n}\right)\binom{p-1}{(n+1)-1}\right]m_{(q+1)} < \sum_{p=q+1}^{N}\left(1-\frac{N-n}{n+1}\frac{n}{p-n}\right)\binom{p-1}{(n+1)-1}m_{(p)} ,$$

when $n \le \cdots \le m_{(q+1)} \le \cdots \le m_{(k-1)} < m_{(k)} = \cdots = m_{(N)}$, implies the strict inequality at (a). In this case we have

$$\frac{N-n}{n+1}m_{(n)} + \sum_{p=n+1}^{q}\left(\frac{N-n}{n+1}\frac{n}{p-n}-1\right)\binom{p-1}{(n+1)-1}m_{(p)} \le \left[\frac{N-n}{n+1} + \sum_{p=n+1}^{q}\left(\frac{N-n}{n+1}\frac{n}{p-n}-1\right)\binom{p-1}{(n+1)-1}\right]m_{(q)}.$$

And if $q = k-1$, then (b) has strict inequality. Thus, the equation (12) shows that $\bar{m}_{(n)} < \bar{m}_{(n+1)}$, when $m_{(k-1)} < m_{(k)}$, and it holds for all $\bar{m}_{(n)} < \cdots < \bar{m}_{(k-1)} < \bar{m}_{(k)} = \cdots = \bar{m}_{(N)}$ since $n$ is arbitrary index, $n < k$.

These three properties demonstrate that the EVC estimator is ordered. Moreover, if we have only two events such as $m_{(k-1)} < m_{(k)}$, then all estimates are strictly increasing with indices $n, n+1, \ldots, k$, and it is $m_{\min} < m_{\max}$ (Haarala, 2021). If we have three maximum events like $m_{(N-2)} = m_{(N-1)} = m_{(N)}$ with their estimates $\bar{m}_{(N-2)} = \bar{m}_{(N-1)} = \bar{m}_{(N)}$, the curve that passes through these three points is a straight line that is parallel to the $x$ axes. This means that the estimates are then $\hat{m}_{\min} = \hat{m}_{\max} = m_{(N)}$ and $\beta = -\infty$ (Haarala, 2021; Haarala and Vermeulen, 2022). The theory shows that we should not get $\beta = \pm\infty$ (Haarala, 2021) and the EVC estimator cannot get the estimates that result in $\beta = \infty$, but unfortunately, $\beta = -\infty$ is possible. This can happen in the real data set because of the binning of the magnitudes (the magnitudes are rounded to the first decimal).

## 4. Connection to the Order Statistics

The maximum function creates the order. Let $1 \le i \le N,\ i \in \mathbb{N}$, be the index of the random variable $M_i$, and let $I_n$ be a set of a size $n$ of indices $I_n = \{i_1, i_2, \ldots, i_n\}$ without regarding the order, where $1 \le i_p \le N$ and $i_p \ne i_q$ for all $p \ne q$. Let $\mathfrak{I}_n$ be a set containing all combination sets of $n$ indices from the $N$ distinct indices. Note: The index $i_p$ can belong to set $I_n$ only once, but it can belong to different sets of $I_n$ with different $n$ values.

We can now set

$$M_{(n)} = \min_{I_n \in \mathfrak{I}_n}\left(\max_{i \in I_n}\{M_i\}\right).$$

For example, the set $I_1$ has only one index then

$$M_{(1)} = \min_{I_1 \in \mathfrak{I}_1}\left(\max_{i \in I_1}\{M_i\}\right) = \min(M_1, M_2, \ldots, M_N),$$

because it is $\mathfrak{I}_1 = \{\{1\},\{2\},\ldots,\{N\}\}$ and $\max(M_i) = M_i$.

We can write $I_n^* \in \mathfrak{I}_n$ as an index set, which gives the minimum

$$M_{(n)} = \min_{I_n \in \mathfrak{I}_n} \left( \max_{i \in I_n} \{M_i\} \right) = \max_{i \in I_n^*} \{M_i\}.$$

Now it follows

$$M_{(n)} = \max_{i \in I_n^*} \{M_i\} \le \max_{i \in I_{n+1}^*} \{M_i\} = M_{(n+1)}.$$

This holds because if $M_{(n+1)} < M_{(n)}$ then

$$\max_{i \in I_n} \{M_i\} \le \max_{i \in I_{n+1}^*} \{M_i\} = M_{(n+1)} < M_{(n)} = \max_{i \in I_n^*} \{M_i\},$$

where $I_n \subset I_{n+1}^*$. Now it can be found out that an index set $I_n$ gives smaller maximum than $I_n^*$. It is a contradiction for the definition of $I_n^*$. Because $M_{(1)}$ is the smallest value of all, we find the order $M_{(1)} \le M_{(2)} \le \cdots \le M_{(N)}$.

The distribution function of Order Statistics can be shown as follows:

$$\begin{aligned}
P\left(M_{(n)} \le m\right) &= P\left(M_{(n)} \le m \vee M_{(n+1)} \le m \vee \ldots \vee M_{(N)} \le m\right) \\
&= \sum_{k=n}^{N} P\left(\max_{i \in I_k^*} \{M_i\} \le m\right) \\
&= \sum_{k=n}^{N} \binom{N}{k} P\left(M_1 \le m \wedge \ldots \wedge M_k \le m \wedge m < M_{k+1} \wedge \ldots \wedge m < M_N\right) \\
&= \sum_{k=n}^{N} \binom{N}{k} \prod_{i=1}^{k} P\left(M_i \le m\right) \prod_{i=k+1}^{N} \left[1 - P\left(M_i \le m\right)\right] \\
&= \sum_{k=n}^{N} \binom{N}{k} \left[P\left(M_i \le m\right)\right]^k \left[1 - P\left(m \le M_i\right)\right]^{N-k} \\
&= \sum_{k=n}^{N} \binom{N}{k} \left[F(m)\right]^k \left[1 - F(m)\right]^{N-k}.
\end{aligned} \tag{14}$$

The first equality follows because $M_{(n)} \le m$, thus, $M_{(n)} \le m \ \left(\le M_{(n+1)}\right), M_{(n)} \le M_{(n+1)} \le m \ \left(\le M_{(n+2)}\right)$ and so on until $M_{(n)} \le \cdots \le M_{(N)} \le m$. The second equation follows because $\mathfrak{I}_n \cap \mathfrak{I}_m = \varnothing, \ n \ne m$. This means that the ordered random variables are independent. The third equality comes from the binomial probability because the maximum doesn't depend on the order, so all the combinations that give the maximum have to be considered.

The distribution (14) is an expression that states how the CDF is presented for the Order Statistics. The result (14) can be rewritten as follows:

$$\begin{aligned}\sum_{k=n}^{N}\binom{N}{k}[F(m)]^{k}[1-F(m)]^{N-k} &= \sum_{k=n}^{N}\binom{N}{k}[F(m)]^{k}\sum_{p=0}^{N-k}(-1)^{N-k-p}\binom{N-k}{p}[F(m)]^{N-k-p} \\ &= \sum_{k=n}^{N}\sum_{p=0}^{N-k}(-1)^{N-k-p}\binom{N}{k}\binom{N-k}{p}[F(m)]^{N-p} \\ &= \cdots\end{aligned}$$

It is necessary to take a common factor $[F(x)]^n$ in each case, when the terms of the double sum can be found

$$\begin{aligned} p=0: &\quad [F(x)]^{N}\sum_{k=n}^{N}(-1)^{N-k}\binom{N}{k}\binom{N-k}{0} \\ p=1: &\quad [F(x)]^{N-1}\sum_{k=n}^{N-1}(-1)^{N-k-1}\binom{N}{k}\binom{N-k}{1} \\ &\quad \vdots \\ p=N-n: &\quad [F(x)]^{N-(N-n)}\sum_{k=n}^{N-(N-n)}(-1)^{(N-k)-(N-n)}\binom{N}{k}\binom{N-k}{N-n}. \end{aligned}$$

Thus, by changing the order of summations it results in

$$\begin{aligned}\cdots &= \sum_{p=0}^{N-n}(-1)^{N-p}\left[\sum_{k=n}^{N-p}(-1)^{k}\binom{N}{k}\binom{N-k}{p}\right][F(m)]^{N-p} \\ &= \sum_{q=n}^{N}(-1)^{q}\left[\sum_{k=n}^{q}(-1)^{k}\binom{N}{k}\binom{N-k}{N-q}\right][F(m)]^{q} \\ &= \cdots.\end{aligned} \tag{15}$$

Now it holds

$$\binom{N}{k}\binom{N-k}{N-q}=\frac{N!}{k!(N-k)!}\frac{(N-k)!}{(N-q)!(q-k)!}=\frac{N!}{q!(N-q)!}\frac{q!}{k!(q-k)!}=\binom{N}{q}\binom{q}{k}.$$

Hence,

$$\sum_{k=n}^{q}(-1)^{k}\binom{N}{k}\binom{N-k}{q-k}=\binom{N}{q}\sum_{k=n}^{q}(-1)^{k}\binom{q}{k}.$$

Using a check for binomial coefficients (Abramowitz et al., 1972), the sum can result in

$$\sum_{k=0}^{q}(-1)^k\binom{q}{k}=(-1)^{q-1}\sum_{k=0}^{q-1}(-1)^{q-1-k}\binom{q}{k}+(-1)^q\binom{q}{q}=(-1)^{q-1}+(-1)^q=0$$
$$\sum_{k=0}^{n-1}(-1)^k\binom{q}{k}=(-1)^{n-1}\sum_{k=0}^{n-1}(-1)^{n-1-k}\binom{q}{k}=(-1)^{n-1}\binom{q-1}{n-1}$$
$$\sum_{k=n}^{q}(-1)^k\binom{q}{k}=\sum_{k=0}^{q}(-1)^k\binom{q}{k}-\sum_{k=0}^{n-1}(-1)^k\binom{q}{k}=(-1)^n\binom{q-1}{n-1}.$$

Thus, the equation (15) can be completed by writing

$$\cdots=\sum_{q=n}^{N}(-1)^{n+q}\binom{N}{q}\binom{q-1}{n-1}[F(m)]^q.$$

The CDF of the $n^{\text{th}}$ Order Statistic could be given by

$$F_{M_{(n)}}(m)=\sum_{p=n}^{N}(-1)^{n+p}\binom{N}{p}\binom{p-1}{n-1}F_{M_p}(m). \tag{16}$$

It provides an easy way to calculate the CDF of the $n^{\text{th}}$ Order Statistic for the GGR distribution. It is essential to replace the $F_{M_p}(m)$ by (3).

The estimator for the EVC (10) gives an idea of what to set

$$F_{M_n}(m)=\sum_{p=n}^{N}\binom{N}{n}^{-1}\binom{p-1}{n-1}F_{M_{(p)}}(m). \tag{17}$$

Applying (16) gives

$$\begin{aligned}F_{M_n}(m)&=\binom{N}{n}^{-1}\sum_{p=n}^{N}\binom{p-1}{n-1}F_{M_{(p)}}(m)\\&=\binom{N}{n}^{-1}\sum_{p=n}^{N}\binom{p-1}{n-1}\sum_{q=p}^{N}(-1)^{p+q}\binom{N}{q}\binom{q-1}{p-1}F_{M_q}(m)\\&=\binom{N}{n}^{-1}\sum_{p=n}^{N}\sum_{q=p}^{N}(-1)^{p+q}\binom{p-1}{n-1}\binom{N}{q}\binom{q-1}{p-1}F_{M_q}(m)\\&=\cdots.\end{aligned}$$

Similar to (15), the double summation can change the order. Thus,

$$\begin{aligned}\cdots&=\binom{N}{n}^{-1}\sum_{q=n}^{N}\sum_{p=n}^{q}(-1)^{p+q}\binom{p-1}{n-1}\binom{N}{q}\binom{q-1}{p-1}F_{M_q}(m)\\&=\binom{N}{n}^{-1}\sum_{q=n}^{N}(-1)^q\left[\sum_{p=n}^{q}(-1)^p\binom{p-1}{n-1}\binom{q-1}{p-1}\right]\binom{N}{q}F_{M_q}(m).\end{aligned} \tag{18}$$

The next step is to find a recurrence formula:

$$\begin{aligned}\binom{p-1}{n-1}\binom{q-1}{p-1} &= \frac{(p-1)!}{(n-1)!(p-n)!}\frac{(q-1)!}{(p-1)!(q-p)!}\\ &= \frac{n}{q}\frac{p!}{n!(p-n)!}\frac{q!}{p!(q-p)!}\\ &= \frac{n}{q}\binom{p}{n}\binom{q}{p}\end{aligned}$$

or to be rewritten as

$$\binom{p}{n}\binom{q}{p} = \frac{q}{n}\binom{p-1}{n-1}\binom{q-1}{p-1}. \tag{19}$$

Knowing that $q \geq p \geq n$, it can apply the recurrence formula (19) $n-1$ times which gives

$$\binom{p-1}{n-1}\binom{q-1}{p-1} = \frac{q-1}{n-1}\cdot\frac{q-2}{n-2}\cdots\frac{q-(n-1)}{1}\cdot\frac{(q-n)!}{(q-n)!}\cdot\binom{p-n}{0}\binom{q-n}{p-n} = \binom{q-1}{n-1}\binom{q-n}{p-n}.$$

To solve the summation on the (18) it can be written as

$$\begin{aligned}\sum_{p=n}^{q}(-1)^p\binom{p-1}{n-1}\binom{q-1}{p-1} &= \binom{q-1}{n-1}\sum_{p=n}^{q}(-1)^p\binom{q-n}{p-n}\\ &= (-1)^n\binom{q-1}{n-1}\sum_{k=0}^{q-n}(-1)^k\binom{q-n}{k}.\end{aligned}$$

If $q > n$, then by using the Binomial Coefficient Check (Abramowitz et al., 1972) it is

$$\begin{aligned}\sum_{k=0}^{q-n}(-1)^k\binom{q-n}{k} &= (-1)^{q-n}\binom{q-n}{q-n} + (-1)^{q-n-1}\sum_{k=0}^{q-n-1}(-1)^{(q-n-1)-k}\binom{q-n}{k}\\ &= (-1)^{q-n}\binom{q-n}{q-n} + (-1)^{q-n-1}\binom{q-n-1}{q-n-1} = 0,\end{aligned}$$

and if $q = n$, then

$$(-1)^n\binom{q-1}{n-1}\sum_{k=0}^{q-n}(-1)^k\binom{q-n}{k} = (-1)^n\binom{n-1}{n-1}(-1)^0\binom{0}{0} = (-1)^n.$$

This shows how all other terms vanish except the $n^{\text{th}}$ term. So,

$$\begin{aligned}
F_{M_n}(m) &= \binom{N}{n}^{-1} \sum_{q=n}^{N} (-1)^q \left[ \sum_{p=n}^{q} (-1)^p \binom{p-1}{n-1} \binom{q-1}{p-1} \right] \binom{N}{q} F_{M_q}(m) \\
&= \binom{N}{n}^{-1} (-1)^n (-1)^n \binom{N}{n} F_{M_n}(m) \\
&= F_{M_n}(m).
\end{aligned}$$

In order to complete the analysis, let's consider the distribution of the $n^{\text{th}}$ Order Static:

$$\begin{aligned}
F_{M_{(n)}}(m) &= \sum_{q=n}^{N} (-1)^{n+q} \binom{N}{q} \binom{q-1}{n-1} F_{M_q}(m) \\
&= \sum_{q=n}^{N} (-1)^{n+q} \binom{N}{q} \binom{q-1}{n-1} \left[ \binom{N}{q}^{-1} \sum_{p=q}^{N} \binom{p-1}{q-1} F_{M_{(p)}}(m) \right] \\
&= (-1)^n \sum_{q=n}^{N} \sum_{p=q}^{N} (-1)^q \binom{q-1}{n-1} \binom{p-1}{q-1} F_{M_{(p)}}(m) \\
&= (-1)^n \sum_{p=n}^{N} \left[ \sum_{q=n}^{p} (-1)^q \binom{q-1}{n-1} \binom{p-1}{q-1} \right] F_{M_{(p)}}(m) \\
&= F_{M_{(n)}}(m).
\end{aligned}$$

It has been shown that (16) is an inverse of (17) and vice versa. Moreover,

$$\begin{aligned}
E\left(X\left(M_{(n)}\right)\right) &= \int x(m) \frac{\partial}{\partial m} \left[ \sum_{p=n}^{N} (-1)^{n+p} \binom{N}{p} \binom{p-1}{n-1} F_{M_p}(m) \right] dm \\
&= \sum_{p=n}^{N} (-1)^{n+p} \binom{N}{p} \binom{p-1}{n-1} \int x(m) F'_{M_p}(m)\, dm \\
&= \sum_{p=n}^{N} (-1)^{n+p} \binom{N}{p} \binom{p-1}{n-1} \int x(m) \frac{\partial}{\partial m} \left[ \sum_{q=p}^{N} \binom{N}{p}^{-1} \binom{q-1}{p-1} E\left(M_{(q)}\right) \right] dm \\
&= (-1)^n \sum_{p=n}^{N} \sum_{q=p}^{N} (-1)^p \binom{p-1}{n-1} \binom{q-1}{p-1} \int x(m) F'_{M_{(q)}}(m)\, dm \\
&= \int x(m)\; dF_{M_{(n)}}(m),
\end{aligned} \tag{20}$$

where $X\left(M_{(n)}\right)$ can be, for example, a $p^{\text{th}}$ central moment $X\left(M_{(n)}\right) = \left(M_{(n)} - E\left(M_{(n)}\right)\right)^p$ or a $p^{\text{th}}$ moment $X\left(M_{(n)}\right) = M_{(n)}^p$. Using expected value (6) and variance (Haarala and Orosco, 2016b)

$$Var\left(M_n\right)=\frac{1}{\beta^2}\sum_{k=2}^{\infty}\frac{2n}{2n+k}\left\{\sum_{j=1}^{k-1}\frac{1}{n+j}\right\}\frac{\left(1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]\right)^k}{n+k}, \tag{21}$$

it is possible to calculate the expected value for $n^{\text{th}}$ Order Static by

$$E\left(M_{(n)}\right)=\sum_{p=n}^{N}(-1)^{n+p}\binom{N}{p}\binom{p-1}{n-1}E\left(M_p\right) \tag{22}$$

and the variance by using the formula

$$\begin{aligned}
Var\left(M_{(n)}\right)&=E\left[\left(M_{(n)}-E\left(M_{(n)}\right)\right)^2\right]\\
&=\sum_{p=n}^{N}(-1)^{n+p}\binom{N}{p}\binom{p-1}{n-1}\int\left(m-E\left(M_{(n)}\right)\right)^2 dF_{M_p}(m)\\
&=\sum_{p=n}^{N}(-1)^{n+p}\binom{N}{p}\binom{p-1}{n-1}\int m^2-2m\,E\left(M_{(n)}\right)+\left[E\left(M_{(n)}\right)\right]^2 dF_{M_p}(m)\\
&=\sum_{p=n}^{N}(-1)^{n+p}\binom{N}{p}\binom{p-1}{n-1}\Big[Var\left(M_p\right)+2\left[E\left(M_p\right)-E\left(M_{(n)}\right)\right]\int m\,dF_{M_p}(m)\\
&\qquad\qquad -\left(\left[E\left(M_p\right)\right]^2-\left[E\left(M_{(n)}\right)\right]^2\right)\int dF_{M_p}(m)\Big]\\
&=\sum_{p=n}^{N}(-1)^{n+p}\binom{N}{p}\binom{p-1}{n-1}\left[Var\left(M_p\right)+\left[E\left(M_p\right)-E\left(M_{(n)}\right)\right]^2\right].
\end{aligned} \tag{23}$$

Table 2 offers a practical example about the Order Statistic. In this case, it is set the $\beta=\log(10)$, $m_{\max}=8$, $m_{\min}=5$, $N=5$ and the sample size is 100000. The expected value and the variance for the Statistic of Maximums are calculated by (6) and (21), respectively. Similarly, the expected value and the variance for the Order Statistic are calculated by (22) and (23), respectively. Using the random generator for the GGR distributed random numbers (Haarala and Vermeulen, 2022) we can estimate the expected values for the Order Statistic as follows:

```
C = sort( GRdistribution(log(10),8,5,5,100000) );
E = mean(C,2)
V = var(C,[],2)
```

**Table 2.** Example of the Order Statistic

| $n$ | $E(M_n)$ | $E(M_{(n)})$ | $\hat{E}(M_{(n)})$ | $Var(M_n)$ | $Var(M_{(n)})$ | $Var(\hat{M}_{(n)})$ |
|---|---|---|---|---|---|---|
| 1 | 5,4313 | 5,0868 | 5,0867 | 0,1796 | 0,0075 | 0,0075 |
| 2 | 5,6459 | 5,1951 | 5,1950 | 0,2203 | 0,0192 | 0,0192 |
| 3 | 5,7883 | 5,3395 | 5,3395 | 0,2359 | 0,0400 | 0,0402 |
| 4 | 5,8946 | 5,5556 | 5,5561 | 0,2431 | 0,0860 | 0,0864 |
| 5 | 5,9794 | 5,9794 | 5,9798 | 0,2464 | 0,2464 | 0,2451 |

## 5. An Ideal Catalogue

We define an Ideal Catalogue to be a catalogue $\{m_1^*, m_2^*, \ldots, m_N^*\}$ of size $N$ such as the EVC estimator (10) with these events has a property $\bar{m}_{(n)} = E(M_{(n)})$ for all $n$. Using the factors of (10) we have

$$\begin{pmatrix} m_1^* \\ \vdots \\ m_{N-1}^* \\ m_N^* \end{pmatrix} = \begin{pmatrix} \frac{1}{N} & \cdots & \frac{1}{N} & \frac{1}{N} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \frac{1}{N} & \frac{N-1}{N} \\ 0 & \cdots & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} E(M_1) \\ \vdots \\ E(M_{N-1}) \\ E(M_N) \end{pmatrix}. \tag{24}$$

The inverse matrix becomes unstable when $N > 10$.

As it was shown in the section above, the Order Statistic is an inverse for the Statistic of Maximums. So, the ideal catalogue is same as the Order Statistic. We can write for the case of the Ideal Catalogue as

$$m_n^* = \sum_{p=n}^{N} (-1)^{n+p} \binom{N}{p} \binom{p-1}{n-1} E(m_n). \tag{25}$$

However, the factors grow rapidly. For example, the factor matrix in the case $N = 15$ looks like

$$\begin{pmatrix} m_1^* \\ m_2^* \\ \vdots \\ m_{14}^* \\ m_{15}^* \end{pmatrix} = \begin{pmatrix} 15 & -105 & 455 & -1365 & 3003 & -5005 & 6435 & -6435 & 5005 & -3003 & 1365 & -455 & 105 & -15 & 1 \\ 0 & 105 & -910 & 4095 & -12012 & 25025 & -38610 & 45045 & -40040 & 27027 & -13650 & 5005 & -1260 & 195 & -14 \\ 0 & 0 & 455 & -4095 & 18018 & -50050 & 96525 & -135135 & 140140 & -108108 & 61425 & -25025 & 6930 & -1170 & 91 \\ 0 & 0 & 0 & 1365 & -12012 & 50050 & -128700 & 225225 & -280280 & 252252 & -163800 & 75075 & -23100 & 4290 & -364 \\ 0 & 0 & 0 & 0 & 3003 & -25025 & 96525 & -225225 & 350350 & -378378 & 286650 & -150150 & 51975 & -10725 & 1001 \\ 0 & 0 & 0 & 0 & 0 & 5005 & -38610 & 135135 & -280280 & 378378 & -343980 & 210210 & -83160 & 19305 & -2002 \\ 0 & 0 & 0 & 0 & 0 & 0 & 6435 & -45045 & 140140 & -252252 & 286650 & -210210 & 97020 & -25740 & 3003 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6435 & -40040 & 108108 & -163800 & 150150 & -83160 & 25740 & -3432 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5005 & -27027 & 61425 & -75075 & 51975 & -19305 & 3003 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3003 & -13650 & 25025 & -23100 & 10725 & -2002 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1365 & -5005 & 6930 & -4290 & 1001 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 455 & -1260 & 1170 & -364 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 105 & -195 & 91 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 15 & 14 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} E(M_1) \\ E(M_2) \\ \vdots \\ E(M_{14}) \\ E(M_{15}) \end{pmatrix}.$$

Since the expected values are approximations, the double precision arithmetic becomes unstable even if we could compute the exact values for the factors. This could work until $N = 30$ even the sum of all factors in each row is 1. This is not hard to show by induction. If $n = N$, then

$$\sum_{p=N}^{N} (-1)^{N+p} \binom{N}{p} \binom{p-1}{N-1} = (-1)^{2N} \binom{N}{N} \binom{N-1}{N-1} = 1.$$

Suppose that

$$\sum_{p=n}^{N} (-1)^{n+p} \binom{N}{p} \binom{p-1}{n-1} = 1.$$

Then

$$\begin{aligned} \sum_{p=n}^{N+1} (-1)^{n+p} \binom{N+1}{p} \binom{p-1}{n-1} &= \sum_{p=n}^{N} (-1)^{n+p} \binom{N+1}{p} \binom{p-1}{n-1} + (-1)^{n+N+1} \binom{N+1}{N+1} \binom{N+1-1}{n-1} \\ &= \sum_{p=n}^{N} (-1)^{n+p} \frac{N+1}{N+1-p} \binom{N}{p} \binom{p-1}{n-1} - (-1)^{n+N} \binom{N}{n-1} \\ &= \sum_{p=n}^{N} (-1)^{n+p} \left(1 + \frac{p}{N+1-p}\right) \binom{N}{p} \binom{p-1}{n-1} - (-1)^{n+N} \binom{N}{n-1} \\ &= \sum_{p=n}^{N} (-1)^{n+p} \binom{N}{p} \binom{p-1}{n-1} + \sum_{p=n}^{N} (-1)^{n+p} \frac{p}{N-p+1} \binom{p-1}{n-1} \binom{N-1}{p-1} - (-1)^{n+N} \binom{N}{n-1} \\ &= 1 + (-1)^{n} \binom{N-1}{n-1} \frac{N}{N-n+1} (-1)^{N} \sum_{p=0}^{N-n} (-1)^{(N-n)-p} \binom{N-n+1}{p} - (-1)^{n+N} \binom{N}{n-1} \\ &= 1 + (-1)^{n+N} \binom{N}{n-1} \binom{N-n}{N-n} - (-1)^{n+N} \binom{N}{n-1} = 1. \end{aligned}$$

The third technique (which uses a random generator) can be applied to estimate the expected values for the Order Statistic

$$\begin{pmatrix} m_1^* \\ \vdots \\ m_{N-1}^* \\ m_N^* \end{pmatrix} = \texttt{mean(sort(} \ \texttt{GRdistribution}\left(\beta, m_{\max}, m_{\min}, N, 1000\right) \ \texttt{),2);} \tag{26}$$

Thus, the conclusion is

$$m_n^* = E\left(M_{(n)}\right).$$

The error

$$\hat{E}\left(M_n\right) - E\left(M_n\right) = \binom{N}{n}^{-1} \sum_{p=n}^{N} \binom{p-1}{n-1} m_n^* - E\left(M_n\right)$$

calculated by the different methods (24)-(26) for the ideal catalogue of size 10 ($\beta = \log\left(10\right)$, $m_{\max} = 8$ and $m_{\min} = 5$) is presented in Table 3.

**Table 3.** Error between the EVC estimators of the ideal catalogue and the theoretical expected value

| $n$ | **Inverse Matrix Method (24)** | **Order Statistic Method (25)** | **Random Generator Method (26)** |
|---|---|---|---|
| 1 | 0,1776E−14 | 0,2780E−12 | 0,0027 |
| 2 | 0 | 0,1474E−12 | 0,0019 |
| 3 | 0 | 0,0258E−12 | 0,0021 |
| 4 | 0 | 0,0142E−12 | 0,0031 |
| 5 | 0,0888E−14 | 0,0098E−12 | 0,0045 |
| 6 | 0 | 0,0044E−12 | 0,0064 |
| 7 | 0 | 0,0018E−12 | 0,0085 |
| 8 | 0,0888E−14 | 0,0009E−12 | 0,0109 |
| 9 | 0 | 0 | 0,0133 |
| 10 | 0 | 0 | 0,0159 |

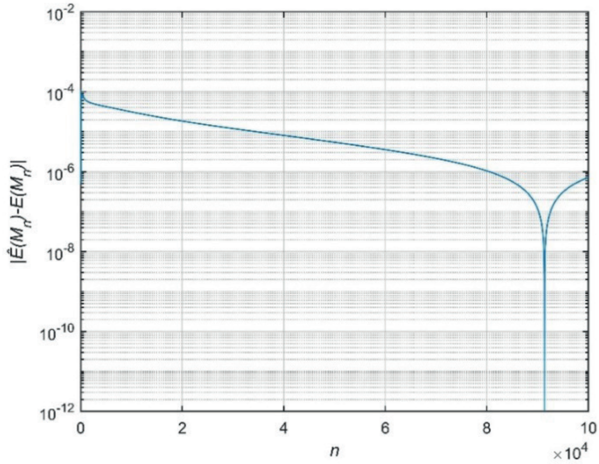


**Figure 1.** Error between the EVC estimator and the real expected value

The ideal catalogue is only a theoretical tool. For example, by using the Random Generator Method that can be seen in Figure 1 that in the case of $\beta = \log(10)$, $m_{\max} = 8$ and $m_{\min} = 5$ the EVC estimator (see Appendix A) is numerically stable and appears to be unbiased up to $N = 10000$.

Table 4 shows a numerical example of the bias introduced by the sample size. In the first column, we have the ideal catalogue of size 6 ($\beta = \log(10)$, $m_{\max} = 8$ and $m_{\min} = 5$). As it was pointed out above, the second column calculated by the EVC estimator (10) is equal to the theoretical expected values. It is important to note that it is not necessary to know all the events, only the $N-n+1$ largest ones. For example, we need the ideal events $m_2^*$, $m_3^*$, $m_4^*$, $m_5^*$ to calculate the $\overline{m}_{(2)}$, but not the event $m_1^*$.

Even though, the smallest value is not used, it is crucial to know the total number of events $(N = 6)$ to obtain unbiased estimates. Let's assume that there exists a threshold that prevents us from recording the events less than $5{,}10$. In this case, the sample size $N$ is $5$ instead of $6$. These estimates for the maximum are presented in the fourth column. Similarly, if the threshold is $5{,}20$, the last column shows the estimates calculated by $N = 4$. The table is also presented in Figure 2.

**Table 4.** Bias of the expected-value-curve estimator

| $n$ | Ideal catalogue | All events | 5 biggest events | 4 biggest events |
|---|---|---|---|---|
| 1 | 5,0723 | 5.4313 | | |
| 2 | 5.1590 | 5.6459 | 5.5031 | |
| 3 | 5.2674 | 5.7883 | 5.7173 | 5.5891 |
| 4 | 5.4117 | 5.8946 | 5.8593 | 5.8027 |
| 5 | 5.6276 | 5.9794 | 5.9653 | 5.9442 |
| 6 | 6.0497 | 6.0497 | 6.0497 | 6.0497 |

In earlier work (Haarala 2021) we presented the estimators for the $m_{\max}$ by using the recurrence formula

$$f_n^{KS-1}\left(\beta\left(m_{\max}-m_{\min}\right)\right)=\frac{f_{n-1}^{KS-1}\left(\beta\left(m_{\max}-m_{\min}\right)\right)}{1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]}-\frac{1}{n}.$$

In Appendix B it was proved that

$$\hat{\beta}=-\frac{(n-2)\hat{E}\left(M_{n-3}\right)-2(n-1)\hat{E}\left(M_{n-2}\right)+n\hat{E}\left(M_{n-1}\right)}{n(n-1)(n-2)\left(\hat{E}^2\left(M_{n-2}\right)+\hat{E}^2\left(M_{n-1}\right)+\hat{E}\left(M_{n-3}\right)\left(\hat{E}\left(M_n\right)-\hat{E}\left(M_{n-1}\right)\right)-\hat{E}\left(M_{n-2}\right)\left(\hat{E}\left(M_{n-1}\right)+\hat{E}\left(M_n\right)\right)\right)},$$

$$\hat{m}_{\max}=\frac{(n-1)\hat{E}\left(M_{n-2}\right)\left[\hat{\beta}n\hat{E}\left(M_n\right)-1\right]-n\hat{E}\left(M_{n-1}\right)\left[\hat{\beta}(n-1)\hat{E}\left(M_{n-1}\right)-1\right]}{\hat{\beta}n(n-1)\left[\hat{E}\left(M_{n-2}\right)-2\hat{E}\left(M_{n-1}\right)+\hat{E}\left(M_n\right)\right]+1},$$

$$\hat{m}_{\min}=\hat{m}_{\max}+\frac{1}{\hat{\beta}}\log\left(1-\frac{\hat{\beta}\left(\hat{m}_{\max}-\hat{E}\left(M_{n-1}\right)\right)}{\hat{\beta}\left(\hat{m}_{\max}-\hat{E}\left(M_n\right)\right)+\frac{1}{n}}\right).$$

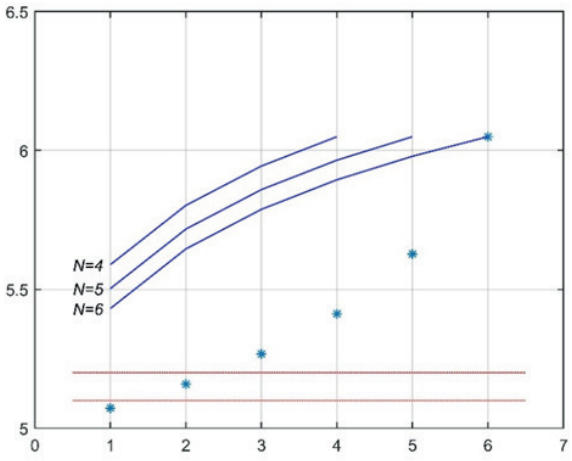


**Figure 2.** Estimated EVC values with different sample size

This method needs 4 points, $\hat{E}(M_{n-3}), \hat{E}(M_{n-2}), \hat{E}(M_{n-1})$ and $\hat{E}(M_n)$ to solve the estimates for the $\hat{\beta}$, $\hat{m}_{\max}$ and $\hat{m}_{\min}$, which means that it is overdetermined. Thus, this method cannot provide solutions in cases where it is still possible to solve the parameters with fewer points. Finding out the method to solve the parameters at the general points $E(M_{n_1}), E(M_{n_2}), E(M_{n_3})$is beyond the scope of this paper.

Table 5 shows the estimates for the $\hat{\beta}$, $\hat{m}_{\max}$ and $\hat{m}_{\min}$. It can be seen that $\hat{\beta}$ has a slight bias with $N = 5$ and $N = 4$, but it gives the right answer for $N = 6$, where the error is in the $10^{\text{th}}$ decimal ($\log(10) = 2{,}30258509299404$).

**Table 5.** Estimates for the $\hat{\beta}$ , $\hat{m}_{\max}$ and $\hat{m}_{\min}$

| **N** | $n$ | $\hat{\beta}$ | $\hat{m}_{\max_2}$ | $\hat{m}_{\min_2}$ |
|---|---|---|---|---|
| 6 | 4 | 2,302585093001670 | 8,0000 | 5,0000 |
| | 5 | 2,302585093001670 | 8,0000 | 5,0000 |
| | 6 | 2,302585093001670 | 8,0000 | 5,0000 |
| 5 | 5 | 2,302590307555657 | 7,9880 | 5,0666 |
| | 6 | 2,302590307555657 | 7,9884 | 5,0669 |
| 4 | 6 | 2,302604754044397 | 7,9710 | 5,1454 |

This example shows how the unbiased estimators have biased results because of the bias of the sample size. Moreover, this shows that the EVC estimator is really related to the expected value curve (6) because the estimates of the EVC parameters with the correct sample size are exactly the same with the parameters of the theoretical model. This is not a surprise given the way in which the EVC estimator was derived above. It also explains why the expected value doesn't change even if the catalogue size $N$ increases, while the magnitudes of the Ideal Catalogue (or in other words the expected values of the $n^{th}$ order statistic), decrease at the same time (Table 6).

**Table 6.** Bias of the expected-value-curve estimator

| $n$ | **Ideal catalogue** | $\hat{E}(M_n)$ | **Ideal catalogue** | $\hat{E}(M_n)$ | **Ideal catalogue** | $\hat{E}(M_n)$ |
|---|---|---|---|---|---|---|
| 1 | 5,1445 | 5,4313 | 5,1084 | 5,4313 | 5,0868 | 5,4313 |
| 2 | 5,3611 | 5,6459 | 5,2529 | 5,6459 | 5,1951 | 5,6459 |
| 3 | 5,7883 | 5,7883 | 5,4692 | 5,7883 | 5,3395 | 5,7883 |
| 4 | | | 5,8946 | 5,8946 | 5,5556 | 5,8946 |
| 5 | | | | | 5,9794 | 5,9794 |

The analysis above leads to three main observations. The distribution model (1) is, mathematically speaking, a double truncated exponential distribution. Given that it is not possible to measure all events, there is a threshold called corner magnitude $m^c \in \left]m_{\min}, m_{\max}\right[$. Corner magnitude is a magnitude limit from which the catalogue is complete upwards. It means that the catalogue has missing events less than the corner magnitude $m^c$. The definition of the EVC estimator (10) and the examples above show that the estimates for the EVC can be computed for any threshold $m^{tr} \in \left[m^c, m_{\max}\right[$. For the four-point method above, it could be $m^c \le m^{tr} \le m_{(N-3)}$. The only requirement is that we must know the real catalogue size, with missing or unknown events. Furthermore, it was shown above that it is possible to find unbiased estimates for all parameters of the double truncated exponential distribution if we know the sample size.

The Gutenberg-Richter law (Ishimoto, Iida, 1939; Gutenberg, Richter, 1944) was presented as

$$\log_{10} N(m) = a - b\,m\,, \tag{27}$$

where $N(m)$ is a number of events with magnitudes greater than or equal to $m$. The magnitude threshold $m_{\min}$ is set so that all events are greater than or equal to $m_{\min}$ and there are no events with magnitudes less than this threshold, rewriting it (27) as

$$\log_{10} N(m) = a - b(m - m_{\min}),$$

we can see

$$\log_{10} N(m_{\min}) = a .$$

So, the unknown sample size must be an estimate of the number of events in the catalogue:

$$\hat{N} = \hat{N}(m_{\min}) = 10^{\hat{a}}.$$

This shows that the $a$ plays a role of the unknown sample size.

The second observation tells us that (1) is not just the double truncated exponential distribution when we talk about the Gutenberg-Richter law (27). Instead of the three parameters, we should estimate four parameters: $\beta$, $m_{\max}$, $m_{\min}$ and $N \geq N(m^c)$ to get unbiased estimators. Thus, when naming (1) as the GGR distribution, it must be considered more than only a double truncated exponential distribution where both parameters, $a$ and $b$ of the Gutenberg-Richter law play a role in the distribution. It is to see the GGR distribution as a subfamily of the double truncated exponential distribution.

The third observation is that we should never use the Aki-Utsu estimator $\hat{\beta}_{AU}\left[=\hat{\beta}_{GAU}^{(1)}\right]$ and the Page's estimator $\hat{\beta}_{P}\left[=\hat{\beta}_{GP}^{(1)}\right]$ because the unknown sample size is bigger than the observed sample size $N > N(m^c)$, and these estimators are defined at $n = 1$ which has missing events below the threshold $m^c$ (see Figure 2). (Note: Definitions of these estimators can be found in Haarala and Orosco (2016b)).

## 6. The Expected Value for the EVC Estimator

As noted above, the expected value for the Statistic of Maximums can be written using the $n^{\text{th}}$ Order Statistic such as

$$\begin{aligned}
E(M_n) &= \int m \, dF_{M_n}(m) \\
&= \int m \frac{\partial}{\partial m}\left[\sum_{p=n}^{N}\binom{N}{n}^{-1}\binom{p-1}{n-1}F_{M_{(p)}}(m)\right] dm \\
&= \sum_{p=n}^{N}\binom{N}{n}^{-1}\binom{p-1}{n-1}\int m \, dF_{M_{(p)}}(m) \\
&= \sum_{p=n}^{N}\binom{N}{n}^{-1}\binom{p-1}{n-1}E(M_{(p)}) .
\end{aligned}$$

Because of that we can present the EVC estimator (10) as

$$\bar{M}_n = \sum_{p=n}^{N} \binom{N}{n}^{-1} \binom{p-1}{n-1} M_{(p)},$$

it follows

$$E\left(\bar{M}_n\right) = \sum_{p=n}^{N} \binom{N}{n}^{-1} \binom{p-1}{n-1} E\left(M_{(p)}\right) = E\left(M_n\right).$$

Therefore, the EVC estimator is an unbiased estimator as it was shown above in an experimental way.

## 7. More Theoretical Observations

In Part I (Haarala and Orosco, 2016a), it was mentioned that $\beta\left(m_{\max} - m_{\min}\right)$ can be associated with a pseudo distribution function like a pseudo maximum $\chi_{\max}$. This is also known as normalized maximum. We can see the expected value (6) as

$$E\left(\beta\left(M_{(\eta)} - m_{\min}\right)\right) = f_{\eta}^{KS-2}\left(\beta\left(m_{\max} - m_{\min}\right)\right) = f_{\eta}^{KS-2}\left(\chi_{\max}\right). \tag{28}$$

or, if we define $\mathrm{X}_{(\eta)} = \beta\left(M_{(\eta)} - m_{\min}\right)$,

$$E\left(\mathrm{X}_{(\eta)}\right) = f_{\eta}^{KS-2}\left(\chi_{\max}\right).$$

This normalized maximum $\chi_{\max}$ defines a class of distribution functions. If $M_{(\eta),1}$ is from the distribution function $\beta_1\left(m_{\max,1} - m_{\min,1}\right)$ and if $M_{(\eta),2}$ is from the distribution function $\beta_2\left(m_{\max,2} - m_{\min,2}\right)$, they have the same expected values

$$E\left(\beta_1\left(M_{(\eta),1} - m_{\min,1}\right)\right) = E\left(\beta_2\left(M_{(\eta),2} - m_{\min,2}\right)\right)$$

for all $\eta$ if and only if $\beta_1\left(m_{\max,1} - m_{\min,1}\right) = \beta_2\left(m_{\max,2} - m_{\min,2}\right)$. This means that these distribution functions belong to the same class. Figure 3 illustrates this situation.

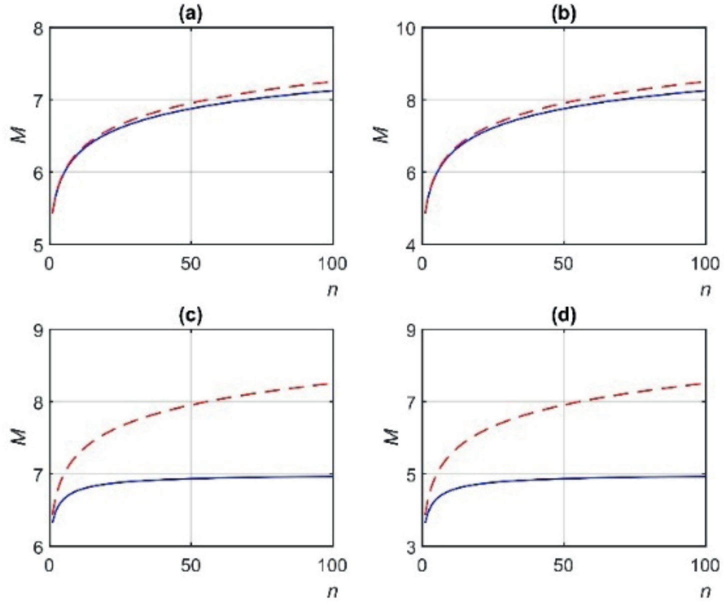


**Figure 2.** Comparison between expected values

Subplot (a) is drawn with $b=1$, $m_{\min}=5$, $m_{\max}=8$ when $\chi_{\max}=1\log(10)(8-5)=3\log(10)$ and subplot (b) is drawn with $b=0.5$, $m_{\min}=4$, $m_{\max}=10$ when $\chi_{\max}=0.5\log(10)(10-4)=3\log(10)$. We can see that the curves are similar and the only difference between them is the scale of the y-axis. Subplot (c) is drawn with $b=1$, $m_{\min}=7$, $m_{\max}=8$ when $\chi_{\max}=1\log(10)(8-7)=\log(10)$ meanwhile subplot (d) with $b=0.5$, $m_{\min}=6$, $m_{\max}=8$ when $\chi_{\max}=0.5\log(10)(8-6)=\log(10)$. We can see the same equality between subplot (c) and (d) as we saw between subplots (a) and (b).

The red curve marks the unbounded expected values. That is to say that in all cases, those unbounded curves have the pseudo maximum $\chi_{\max}=\beta(\infty-m_{\min})=\infty$.

These examples show how the shape of the curve depends only on the normalized maximum $\chi_{\max}$ and the parameters $\beta$, $m_{\max}$ and $m_{\min}$ just scale that curves. The same can be seen from (28) where the expected curve is defined in the normalized form, and it only depends on the normalized maximum $\chi_{\max}$ giving the linear transformation between the normalized distribution function and the final distribution function.

We call $\mathfrak{s}$ as a shape factor defined as

$$\mathfrak{s}=\chi_{\max}/\log(10)=b(m_{\max}-m_{\min}).$$

## 8. Convergence of the KS-2 Function

The expected value of (6) using the series (5) can be written as

$$E\left(M_{(n)}\right)=m_{\min}+\frac{1}{\beta}\sum_{k=1}^{\infty}\frac{\left(1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]\right)^{k}}{k\left(\frac{k}{n}+1\right)}$$

$$\underset{n\to\infty}{\to} m_{\min}+\frac{1}{\beta}\sum_{k=1}^{\infty}\frac{\left(1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]\right)^{k}}{k}.$$

This series is the same as (writing $z=1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]$ )

$$\sum_{k=1}^{\infty}\frac{z^{k}}{k}=-\log\left(1-z\right)=\beta\left(m_{\max}-m_{\min}\right).$$

Thus, the expected value is asymptotically

$$E\left(M_{(\infty)}\right)=m_{\max}. \tag{29}$$

This is natural since the expected maximum magnitude of all possible magnitudes is a maximum possible magnitude. It means that the estimator $\hat{m}_{\max}\to m_{\max}$ when $n\to\infty$. Using the Pisarenko estimator $E\left(M_{N}\right)=m_{N}$ to estimate $\hat{m}_{\max}$, it makes the result robust. The Pisarenko estimator is an unbiased estimator even with the bias of the catalogue size $N$. Moreover, the convergence fixes the bias on the estimate $\hat{m}_{\max}$ produced by the $\beta$ and $m_{\min}$ values.

Let's assume that we know the EVC $E\left(M_{n}\right)$ for the shape factor $\varsigma=1\left(8-5\right)$. It is assumed also that we know the minimum: $m_{\min}=5$. The real $b=1$ is perturbed as it is shown in Figure 4. The estimate for the maximum $\hat{m}_{\max}$ could be solved by using an iterative solution as it was presented in Part I in the Analysis of Gutenberg-Richter b-value and $m_{\max}$ (Haarala and Orosco, 2016a). Figure 4 shows how the estimate of the maximum converges towards the maximum with all $b$-values.

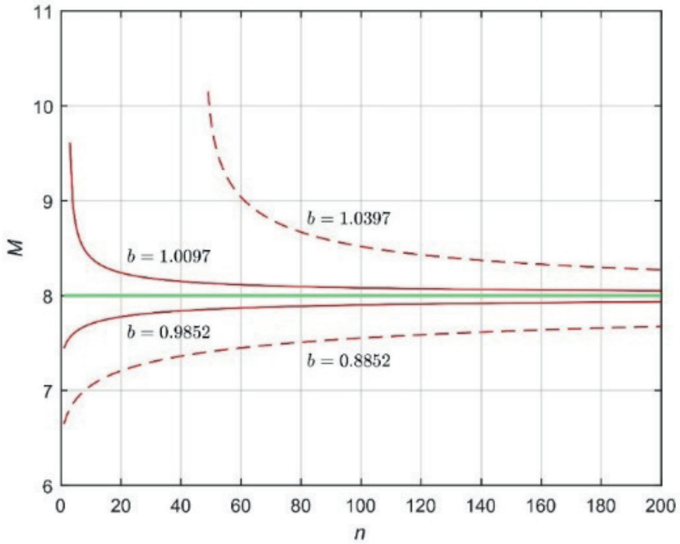


**Figure 4.** Convergence of the estimator $\hat{m}_{\max}$ .

## 9. The Gutenberg-Richter Distribution in the Discrete (binned) Case

Up to now, the analysis has focused on the continuous GR distribution. In practical applications, the events are grouped because of the limitations of the measurement and bounded computing. As a consequence, all values from the interval $m \in [x - 0.05, x + 0.05[$ are replaced by the value $x$ (or rounded to $x$). It leads to the phenomenon contrary to the theoretical results.

It was shown above that the EVC estimator is ordered: $m_{(n)} \leq m_{(n+1)}$ for all $n = 1, \ldots, N$. Mathematically speaking, this is a monotonically increasing sequence $\{m_{(n)}\}$. The next theorem states the convergence of this monotone sequence (Rudin, 1987):

**Theorem**. Suppose $\{m_{(n)}\}$ is monotonic. Then $\{m_{(n)}\}$ converges if and only if it is bounded.

In the case of earthquakes, the maximum of the magnitudes must be bounded because otherwise we would have, for example, an earthquake with infinite energy or endless moving of the fault. This means that the EVC estimator converges to some limit $m_{\max}$. Since the Pisarenko estimator is a maximum event in the catalogue, thus $\hat{E}(M_N) = m_{(N)}$, we conclude that $\hat{E}(M_N) = m_{(N)} \rightarrow m_{\max} = \hat{E}(M_{(\infty)})$, just like it was demonstrated above. This means that there exists some $N = 1, 2, \ldots$ such that so that $m_{\max} - m_{(N)} < \varepsilon, \quad \varepsilon > 0$. Using (6) and the Pisarenko estimator, we can write

$$0 \le m_{\max} - E\left(M_N\right) = m_{\max} - m_{(N)} = \frac{1}{\beta} f_N^{KS-1}\left(\mathrm{X}_{(N)}\right) < \varepsilon.$$

Let's consider $\beta = 1\log(10)$, $m_{\min} = 4.95$, $m_{\max} = 8.05$ and $\varepsilon = 0.1$, we find $N \ge 4415$ that $m_{\max}$ and $m_{(4415)}$ could belong to the same bin. Thus, it is expected that $m_{(4415)} = 8.0$. If we have three subcatalogues of size $N \ge 4415$, it is expected that $m_{(4415),1} = m_{(4415),2} = m_{(4415),3} = 8.0$ indicating $\beta = -\infty$ even we should not reach the minus infinity in the theoretical sense (Haarala, 2021). Because of the limitations of the measurement and bounded computing (given the discretization of the continuous data to the bins of equal size) we can even find a negative infinity value for $b$ in the real catalogues.

## 10. Examples

*a) Can the b-value entail a minus infinity in real case?*

Let's consider the volume bounded by the coordinates S28.3$^0$, S26.9$^0$, W70.5$^0$, W65.0$^0$ and depths between 0 km and 30 km. The superficial area of this volume is marked with a red line in Figure 5.

**Table 7.** Subcatalogue for the events on figure 5. *(Data: ISC, 2023)*

| # | EventID | Time | Latitude | Longitude | Depth | Auth. | M | Type | Auth. |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1884411 | 2001/06/13 04:10:47.19 | -27,1775 | -66,4524 | 13,3 | ISC | 5,0 | MW | GCMT |
| 2 | 624494650 | 2022/06/21 08:10:00.80 | -28,1100 | -69,1100 | 22,5 | GCMT | 5,1 | MW | GCMT |
| 3 | 602709234 | 2013/03/19 11:41:55.94 | -27,8545 | -69,4714 | 9,8 | ISC | 5,1 | MW | SJA |
| 4 | 603231530 | 2013/07/13 21:36:12.60 | -26,9684 | -66,8994 | 15,7 | ISC | 5,2 | MW | GCMT |
| 5 | 14240651 | 2010/01/19 17:28:15.45 | -27,5764 | -65,8767 | 27,0 | ISC | 5,2 | MW | GCMT |
| 6 | 1734914 | 2000/04/30 05:31:26.23 | -27,0000 | -65,9970 | 24,8 | ISC | 5,2 | MW | GCMT |
| 7 | 687573 | 1978/01/13 01:21:00.84 | -27,5747 | -65,7771 | 23,4 | ISC | 5,2 | mb | NEIS |

The subcatalogue was created by using primarily location information from the ISC (International Seismological Centre). If location data had not been available from the ISC, then it would have been taken from the author who has published the magnitude data. The magnitude has been selected in the order in which the first is accepted at the moment magnitude of GCMT (The Global CMT Project, Columbia University), if still not available, then another author who has published the moment magnitude (here only SJA – INPRES, Instituto Nacional de Prevención Sísmica) would have been selected. If there had been no reported moment magnitude, then the magnitude would have been taken from the NEIS or NEIC catalogue of the USGS (United States Geological Survey). The events $m \geq 5.0$ of this volume are listed in Table 7 above as a subcatalogue.

Even though Table 7 shows only the seven largest magnitudes of the volume, it can be seen that four of them are maximums and equivalent. Mathematically speaking, it is enough to give a realistic example of the possible catalogue to show that the b-value can reach minus infinity in the real case.

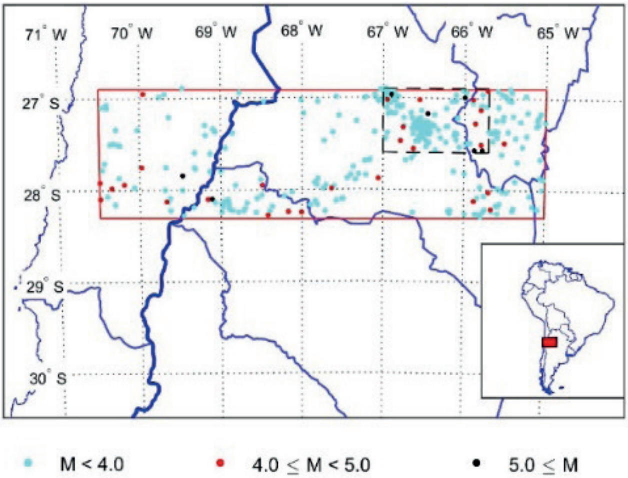


**Figure 5.** An example catalogue. *(Data: ISC, 2023)*

The volume (red area in Figure 5) may represent a seismogenic zone. The dashed black line presents a subvolume bounded by the coordinates S27.6$^0$, S26.9$^0$, W67.0$^0$, W65.7$^0$ and depths between 0 km and 30 km. Events number 2 and 3 do not belong to this smaller subvolume, but both zones get $b = -\infty$. If we select any other zone within the red area, which includes the black one, $b = -\infty$ is reached, which demonstrates that there are many chances to face negative infinity case.

Previous analysis (Haarala, 2021) stated that if the expected values are on the same horizontal line, then we have either $m_{\max} = m_{\min}$ or $\beta = \pm\infty$. Knowing how the EV curve behaves depending on $\beta$ (Haarala and Vermeulen, 2022), the EVC can be expressed in the case of $\beta = -\infty$ ($m_{\min} < m_{\max}$) as

$$f(\eta) = \begin{cases} m_{\min}, & \eta = 0, \\ m_{\max}, & 0 < \eta \leq \infty \end{cases}$$

It's impossible to say anything about the minimum in this case, so the estimates can be set as $\hat{m}_{\max} = \hat{m}_{\min} = m_{(N)}$, thus

$$f(\eta) = m_{\max}, \quad 0 \leq \eta \leq \infty.$$

The majority of the magnitudes in Table 7 are the moment magnitude type, except one, which is body magnitude. Considering the original data from the bulletin of ISC (ISC, 2023), event number 7 had three reported magnitudes; without taking into account the homogenization of magnitude units, in our case the report of NEIS was adopted, and with that we got the negative b-value.

| **Magnitude** | | **Err** | **Nsta** | **Author** | **OrigID** |
|---|---|---|---|---|---|
| mb | 5.2 | | | NEIS | 1521873 |
| MSZ | 4.2 | | | NEIS | 1521873 |
| mb | 5.3 | 0.3 | 16 | ISC | 1521874 |

If we had chosen the magnitude of ISC, there would have been no negative infinity for the $b$-value because there would have been only a single maximum value 5.3. This gives an idea that the set of maximums $\{5.2, 5.2, 5.2, 5.2\}$ could be replaced by the set. $\{5.15, 5.20, 5.20, 5.25\}$ in the calculus to avoid the problem of negative infinity, because it should not happen as the theory states (Haarala, 2021); but it happens because of the measurements and the limitations of the calculus. It is important to remember that those highest events in which ones fall into the interval $[5.15, 5.25[$ are not exactly 5.2. They are rounded.

*b) Estimates of the algebraic solution*

Let's consider the border area between Argentina and Bolivia bounded by the coordinates S24.0$^0$, S21.0$^0$, W66.0$^0$, W62.5$^0$ and depths between 0 km and 30 km. The events are shown in Figure 6. Table 8 lists these events with magnitude 4.0 or greater using the same criteria concerning the type than the events in Table 7.

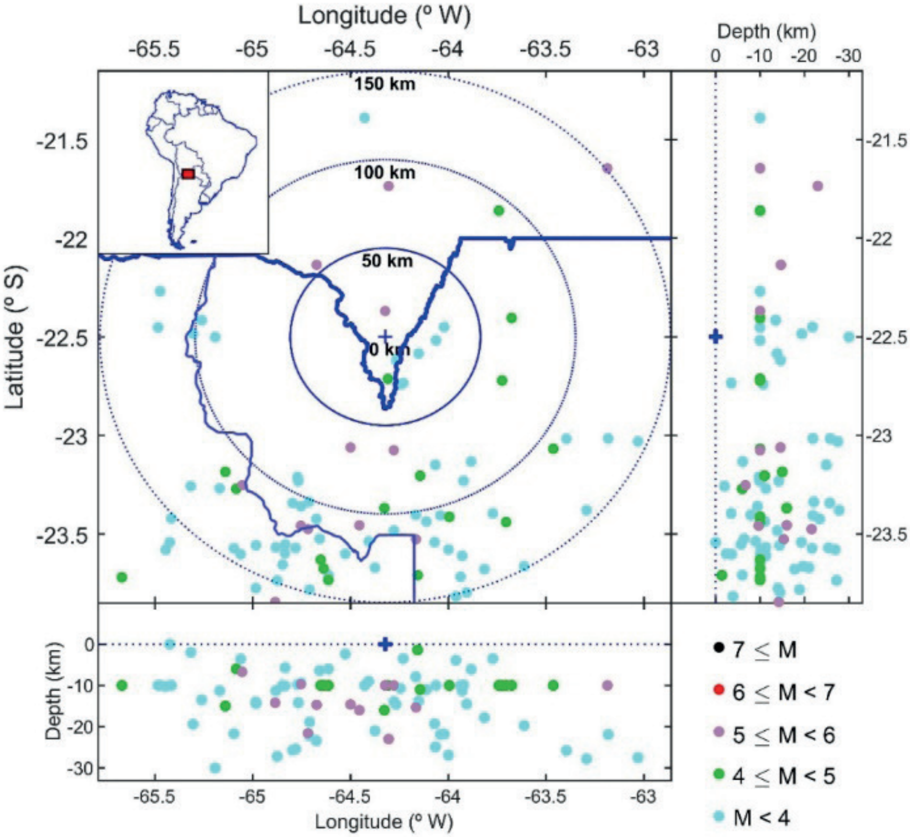


**Figure 6.** Argentinian and Bolivian border region *(Data: ISC, 2023)*

**Table 8.** Events on the Argentinian and Bolivian border region *(Data: ISC, 2023)*

| # | EventID | Time | Latitude | Longitude | Depth | Auth. | M | Type | Auth. |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 604743683 | 2013/04/11 11:02:33,00 | -23,7410 | -64,3860 | 10,0 | SJA | 4,0 | MW | SJA |
| 2 | 607673504 | 2015/07/24 12:41:13,49 | -23,7336 | -64,6132 | 10,0 | ISC | 4,0 | MW | SJA |
| 3 | 617459834 | 2018/11/12 10:32:48,44 | -23,4134 | -63,9947 | 10,0 | ISC | 4,0 | MW | SJA |
| 4 | 604744066 | 2013/04/27 22:33:28,20 | -23,7200 | -65,6690 | 10,0 | SJA | 4,1 | MW | SJA |
| 5 | 608060256 | 2015/11/30 09:01:38,44 | -23,5572 | -64,6042 | 10,0 | ISC | 4,1 | MW | SJA |
| 6 | 7315744 | 2004/03/22 06:23:25,13 | -23,2619 | -64,4878 | 16,5 | ISC | 4,2 | mb | NEIC |
| 7 | 601863935 | 2012/09/21 02:15:28,54 | -23,7099 | -64,1569 | 1,4 | ISC | 4,2 | MW | SJA |
| 8 | 612529869 | 2018/08/03 05:57:16,64 | -23,2049 | -64,1438 | 11,0 | ISC | 4,2 | MW | SJA |
| 9 | 618140070 | 2020/04/22 16:27:30,76 | -23,6492 | -64,6217 | 26,9 | ISC | 4,2 | MW | SJA |
| 10 | 620928864 | 2021/07/19 09:44:48,94 | -23,2892 | -64,5205 | 10,0 | ISC | 4,2 | MW | SJA |
| 11 | 621828931 | 2022/01/30 02:44:53,20 | -23,2710 | -65,0840 | 6,0 | SJA | 4,2 | MW | SJA |
| 12 | 602643 | 1982/03/21 08:44:25,84 | -22,4044 | -63,6757 | 10,0 | ISC | 4,3 | mb | NEIS |
| 13 | 601867657 | 2012/11/01 12:02:00,22 | -23,9076 | -65,0637 | 10,0 | ISC | 4,3 | MW | SJA |
| 14 | 619648365 | 2021/01/13 06:27:10,97 | -23,7105 | -64,7381 | 10,0 | ISC | 4,3 | MW | SJA |
| 15 | 601837043 | 2010/12/31 18:24:24,40 | -23,9130 | -64,4880 | 22,9 | SJA | 4,4 | MW | SJA |
| 16 | 601848169 | 2011/10/06 12:13:55,20 | -22,7220 | -63,7250 | 10,0 | SJA | 4,4 | MW | SJA |
| 17 | 602780765 | 2013/04/11 09:15:47,18 | -23,6778 | -64,6408 | 10,0 | ISC | 4,4 | MW | SJA |
| 18 | 617810119 | 2020/03/22 01:13:22,67 | -23,6965 | -64,5594 | 13,1 | ISC | 4,5 | MW | SJA |
| 19 | 620573088 | 2021/06/29 00:31:54,00 | -23,8850 | -64,3886 | 6,6 | ISC | 4,5 | MW | SJA |
| 20 | 625301594 | 2022/11/23 18:37:02,65 | -23,6309 | -64,6525 | 10,0 | NEIC | 4,5 | mb | NEIC |
| 21 | 608097349 | 2015/09/08 03:03:14,60 | -22,7134 | -64,3089 | 10,0 | ISC | 4,6 | mb | NEIC |
| 22 | 1037023 | 1997/07/15 03:53:55,58 | -23,4401 | -63,7044 | 10,0 | ISC | 4,7 | mb | NEIC |
| 23 | 621841713 | 2022/02/03 09:35:46,30 | -23,1850 | -65,1400 | 15,0 | SJA | 4,7 | MW | SJA |
| 24 | 600130085 | 2011/12/31 16:15:09,72 | -23,3688 | -64,3268 | 16,0 | ISC | 4,8 | MW | SJA |
| 25 | 608458285 | 2013/09/22 13:39:22,47 | -21,8602 | -63,7415 | 10,0 | ISC | 4,8 | MW | SJA |
| 26 | 617577801 | 2020/02/27 01:25:25,92 | -23,6753 | -64,6374 | 10,0 | ISC | 4,8 | MW | GCMT |
| 27 | 618119367 | 2020/04/16 18:17:43,84 | -23,0682 | -63,4630 | 10,0 | ISC | 4,8 | MW | SJA |
| 28 | 608079043 | 2015/12/05 02:10:00,26 | -23,5076 | -64,6859 | 12,5 | ISC | 4,9 | MW | SJA |
| 29 | 851194 | 1966/01/24 21:07:41,37 | -23,5270 | -64,1652 | 15,3 | ISC | 5,0 | mb | USCGS |
| 30 | 614973348 | 2019/02/23 16:34:08,55 | -21,6448 | -63,1862 | 10,0 | ISC | 5,0 | MW | GCMT |
| 31 | 1893856 | 2001/06/25 09:57:01,22 | -21,7348 | -64,3052 | 23,0 | ISC | 5,1 | mb | NEIC |
| 32 | 7487935 | 2005/03/31 21:52:30,92 | -23,5342 | -64,3427 | 9,5 | ISC | 5,1 | MW | GCMT |
| 33 | 614911565 | 2019/02/17 13:22:07,56 | -23,4580 | -64,7543 | 9,7 | ISC | 5,1 | MW | GCMT |
| 34 | 611829837 | 2016/05/22 12:20:42,58 | -22,3693 | -64,3226 | 10,0 | ISC | 5,2 | MW | GCMT |
| 35 | 200649 | 1993/10/02 00:06:03,69 | -23,9763 | -64,4469 | 25,9 | ISC | 5,3 | MW | GCMT |
| 36 | 549589 | 1984/07/24 22:23:45,55 | -23,0753 | -64,2785 | 10,0 | ISC | 5,3 | mb | NEIS |
| 37 | 1658685 | 1999/12/30 14:49:56,98 | -23,8449 | -64,8851 | 14,2 | ISC | 5,3 | MW | GCMT |
| 38 | 13917676 | 2009/11/06 08:49:53,72 | -23,4561 | -64,4546 | 16,0 | ISC | 5,4 | MW | GCMT |
| 39 | 745417 | 1974/07/01 16:51:53,59 | -22,1352 | -64,6732 | 14,7 | ISC | 5,5 | mb | NEIS |
| 40 | 608060607 | 2015/11/29 18:56:22,17 | -23,4521 | -64,5633 | 10,0 | ISC | 5,5 | MW | GCMT |
| 41 | 7315685 | 2004/03/22 04:22:57,71 | -23,0611 | -64,4999 | 14,6 | ISC | 5,7 | MW | GCMT |
| 42 | 619561623 | 2020/11/29 16:40:43,66 | -23,2530 | -65,0546 | 6,7 | ISC | 5,7 | MW | GCMT |
| 43 | 608059513 | 2015/11/29 18:52:51,42 | -23,4755 | -64,7173 | 21,6 | ISC | 5,8 | MW | GCMT |

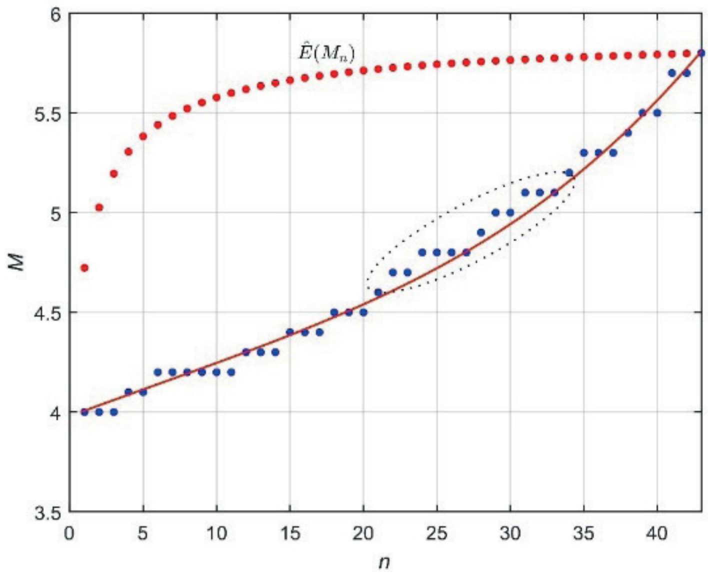


**Figure 7.** Ordered events and their EV estimates

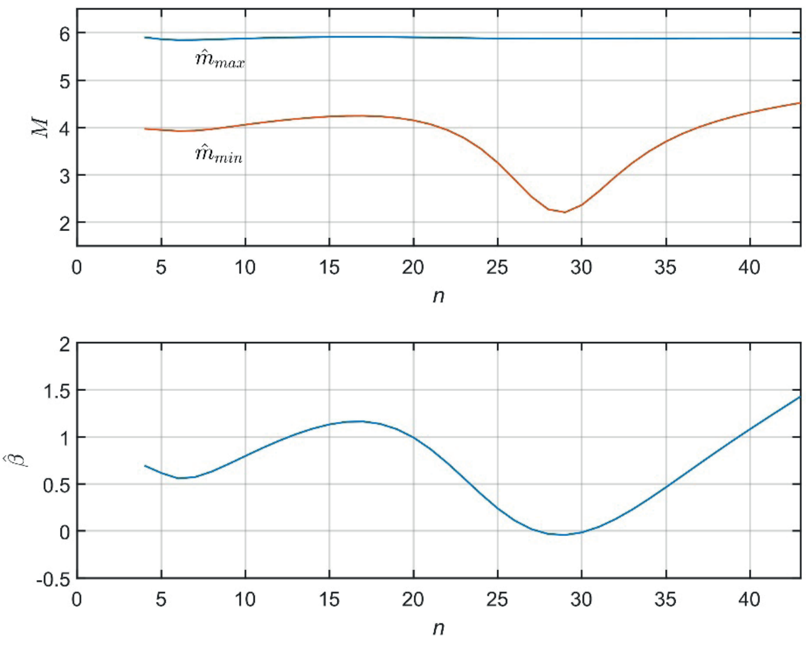


**Figure 8.** Estimates for the catalogue of the table 8

Figure 7 shows the ordered events (blue dots) and their EVC estimates (red dots). The red line is drawn to show why $\beta$ decreases (Figure 8) and even gets negative values ($\min(\beta_n) = -0.0427$). The estimates of the maximum magnitude (earthquake) are quite stable to get, taking values of $\min(m_{\max}(n)) = 5.84$, $\max(m_{\max}(n)) = 5.91$ and $\overline{m}_{\max}(n) = 5.89$.

The negative $\beta$ indicates that the events inside the dotted ellipsoid of Figure 7 are stronger than what was expected.

Given that the data was truncated at 4 and $n$ was set $n = 4, \dots, 43$, the solution tried to estimate $m_{\min}$ as 4. The variation of $\beta$ and $m_{\min}$ indicates the variation of the magnitude of the events around the theoretical model of the events (red line in Figure 8).

## 11. Conclusion

We have demonstrated that there exists an algebraic solution to estimate $\beta$, $m_{\max}$ and $m_{\min}$. Since it is based on the four EVC estimators, there exists a possibility that the solution will fail sometimes. In order to avoid this problem, the solution should use only three estimators.

Another important result was to demonstrate that the negative $\beta$ is not only of theoretical interest, but is necessary to consider its existence in the real catalogues when we use the generalized estimators (not within the framework of classical theory), including the negative infinity for the $\beta$; the reason for this is the binned continuous events.

The theory we have developed over time and from where the foundations of this work come from, is thought to be worthy for hazard assessment studies, especially in seismic zones with little history. The analysis also shows the rich information that the real event catalogue provides to access reliable parameters in probabilistic seismic hazard studies.

With the idea of getting rid of the need to “guess” necessary values to apply to existing models, the development of the theory led to generalize the GR distribution function to consider the non-positive cases of $\beta$.

## Appendix A

We gave in formula (10) the expected value

$$\overline{m}_n = \binom{N}{n}^{-1} \sum_{p=n}^{N} \binom{p-1}{n-1} m_{(p)} .$$

We shouldn't use the ready functions for the binomial coefficient. Quite simply, we can get a form to use in the numerical calculus

$$\begin{aligned} \overline{m}_n &= \sum_{p=n}^{N} \frac{\dfrac{(p-1)!}{(n-1)!(p-n)!}}{\dfrac{N!}{n!(N-n)!}} m_{(p)} \\ &= n \sum_{p=n}^{N} \frac{1}{p} \prod_{k=0}^{n-1} \frac{p-k}{N-k} m_{(p)} \end{aligned}$$

These factors can be represented as

$$\begin{pmatrix} \overline{m}_{(n_1)} \\ \overline{m}_{(n_2)} \\ \overline{m}_{(n_3)} \\ \overline{m}_{(n_4)} \end{pmatrix} = \begin{pmatrix} n_1 \frac{1}{p_1} \frac{p_1-0}{n_4-0} & n_1 \frac{1}{p_2} \frac{p_2-0}{n_4-0} & n_1 \frac{1}{p_3} \frac{p_3-0}{n_4-0} & n_1 \frac{1}{p_4} \frac{p_4-0}{n_4-0} \\ 0 & n_2 \frac{1}{p_2} \frac{p_2-0}{n_4-0} \frac{p_2-1}{n_4-1} & n_2 \frac{1}{p_3} \frac{p_3-0}{n_4-0} \frac{p_3-1}{n_4-1} & n_2 \frac{1}{p_4} \frac{p_4-0}{n_4-0} \frac{p_4-1}{n_4-1} \\ 0 & 0 & n_3 \frac{1}{p_3} \frac{p_3-0}{n_4-0} \frac{p_3-1}{n_4-1} \frac{p_3-2}{n_4-2} & n_3 \frac{1}{p_4} \frac{p_4-0}{n_4-0} \frac{p_4-1}{n_4-1} \frac{p_4-2}{n_4-2} \\ 0 & 0 & 0 & n_4 \frac{1}{p_4} \frac{p_4-0}{n_4-0} \frac{p_4-1}{n_4-1} \frac{p_4-2}{n_4-2} \frac{p_4-3}{n_4-3} \end{pmatrix} \begin{pmatrix} m_{(p_1)} \\ m_{(p_2)} \\ m_{(p_3)} \\ m_{(p_4)} \end{pmatrix}.$$

Assuming that $D = \left(m_{p_1}, \ldots, m_{p_k}\right)^T$ is a column array of the magnitudes and let the $n$ to be an index array of the size of the array $D$. The MATLAB code can be written as

```
D  = sort(D(:));
n  = n(:)';
Mn = zeros(numel(n),1);

factor = ones(size(n))/n(end);
ind    = true(size(factor));
```

```
Mn(1)  = n(1) * (factor * D);
ind(1) = false;


for k = 2:numel(n)

   factor(ind) = factor(ind) .* ((n(ind)-k+1)/(n(end)-k+1));
   Mn(k)       = n(k) * (factor(ind) * D(ind));
   ind(k)      = false;

end
```

Let's consider another idea:

$$
\begin{aligned}
\bar{m}_n &= \sum_{p=n}^{N} \frac{n}{p} \frac{\prod_{k=0}^{n-1}(p-k)}{\prod_{k=0}^{n-1}(N-k)} m_{(p)} = \sum_{p=n}^{N} \frac{n}{p} \frac{\prod_{k=p-n+1}^{p} k}{\prod_{k=N-n+1}^{N} k} m_{(p)} \\
&= \sum_{p=n}^{N} \frac{n}{p} \frac{\prod_{k=p-n+1}^{p} k \prod_{k=p+1}^{N-n} k}{\prod_{k=N-n+1}^{N} k \prod_{k=p+1}^{N-n} k} m_{(p)} = \sum_{p=n}^{N} \frac{n}{p} \frac{\prod_{k=p-n+1}^{N-n} k}{\prod_{k=p+1}^{N} k} m_{(p)} \\
&= \sum_{p=n}^{N} \frac{n}{p} \frac{\prod_{k=p+1}^{N} (k-n)}{\prod_{k=p+1}^{N} k} m_{(p)} = \sum_{p=n}^{N} \frac{1}{p} \left[ n \prod_{k=p+1}^{N} \left(1 - \frac{n}{k}\right) \right] m_{(p)}.
\end{aligned}
$$

Note that $\prod_{k=N+1}^{N}(\bullet) = 1$. The matrix of the factors is now

$$
\begin{pmatrix} \bar{m}_{(n_1)} \\ \bar{m}_{(n_2)} \\ \bar{m}_{(n_3)} \\ \bar{m}_{(n_4)} \end{pmatrix} =
\begin{pmatrix}
\frac{1}{p_1} n_1 \left(1-\frac{n_1}{n_4}\right)\left(1-\frac{n_1}{n_3}\right)\left(1-\frac{n_1}{n_2}\right) & \frac{1}{p_2} n_1 \left(1-\frac{n_1}{n_4}\right)\left(1-\frac{n_1}{n_3}\right) & \frac{1}{p_3} n_1 \left(1-\frac{n_1}{n_4}\right) & \frac{1}{p_4} n_1 \\
0 & \frac{1}{p_2} n_2 \left(1-\frac{n_2}{n_4}\right)\left(1-\frac{n_2}{n_3}\right) & \frac{1}{p_3} n_2 \left(1-\frac{n_2}{n_4}\right) & \frac{1}{p_4} n_2 \\
0 & 0 & \frac{1}{p_3} n_3 \left(1-\frac{n_3}{n_4}\right) & \frac{1}{p_4} n_3 \\
0 & 0 & 0 & \frac{1}{p_4} n_4
\end{pmatrix}
\begin{pmatrix} m_{(p_1)} \\ m_{(p_2)} \\ m_{(p_3)} \\ m_{(p_4)} \end{pmatrix}.
$$

As above, the MATLAB code can be written as

```
D   = sort(D(:));
n   = n(:);

factor    = n;
ind       = true(size(factor));

Mn        = factor * (D(end)/n(end));
ind(end) = false;


for k = 1:numel(n)-1

   factor(ind) = factor(ind) .* (1-n(ind)/n(end-k+1));
   Mn(ind)     = Mn(ind) + factor(ind) * (D(end-k)/n(end-k));
   ind(end-k)  = false;

end
```

In any case, this version is faster than the first idea. The difference in evaluation time between these codes comes out, when $n \gg 10000$.

## Appendix B

Following Part I (Haarala and Orosco, 2016a), it is

$$E\left(M_\eta\right) = m_{\max} - \int_{m_{\min}}^{m_{\max}} \left( \frac{1-\exp\left[-\beta\left(m-m_{\min}\right)\right]}{1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]} \right)^{\eta} dm.$$

Integrating

$$-\frac{1}{\beta n}\frac{\partial}{\partial m}\left(1-\exp\left[-\beta\left(m-m_{\min}\right)\right]\right)^{\eta} = \left(1-\exp\left[-\beta\left(m-m_{\min}\right)\right]\right)^{\eta}$$
$$-\left(1-\exp\left[-\beta\left(m-m_{\min}\right)\right]\right)^{\eta-1},$$

it is

$$\int_{m_{\min}}^{m_{\max}} \left(1-\exp\left[-\beta\left(m-m_{\min}\right)\right]\right)^{\eta} dm = -\frac{1}{\beta\eta}\left(1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]\right)^{\eta}$$
$$+\int_{m_{\min}}^{m_{\max}} \left(1-\exp\left[-\beta\left(m-m_{\min}\right)\right]\right)^{\eta-1} dm.$$

After dividing by $\left(1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]\right)^{\eta}$, it yields

$$\int_{m_{\min}}^{m_{\max}} \left( \frac{1-\exp\left[-\beta\left(m-m_{\min}\right)\right]}{1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]} \right)^{\eta} dm = \frac{\displaystyle\int_{m_{\min}}^{m_{\max}} \left( \frac{1-\exp\left[-\beta\left(m-m_{\min}\right)\right]}{1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]} \right)^{\eta-1} dm}{1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]} - \frac{1}{\beta\eta}. \tag{30}$$

Replacing

$$\int_{m_{\min}}^{m_{\max}} \left( \frac{1-\exp\left[-\beta\left(m-m_{\min}\right)\right]}{1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]} \right)^{\eta} dm = m_{\max} - E\left(M_\eta\right)$$

in equation (30) and multiplying by $\beta$, we get the recurrence formula

$$\beta\left(m_{\max}-E\left(M_\eta\right)\right) = \frac{\beta\left(m_{\max}-E\left(M_{\eta-1}\right)\right)}{1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]} - \frac{1}{\eta}.$$

This is a general proof of the recurrence formula presented in Part III (Haarala, 2021) and it holds for all $\beta \in \mathbb{R}, \eta \in \mathbb{R}_+$.

The case of $n = 1, 2, \ldots$ can be written as

$$\frac{\beta\left(m_{\max} - E\left(M_{n-1}\right)\right)}{\beta\left(m_{\max} - E\left(M_n\right)\right) + \dfrac{1}{n}} = 1 - \exp\left[-\beta\left(m_{\max} - m_{\min}\right)\right],$$

$$\frac{\beta\left(m_{\max} - E\left(M_{n-2}\right)\right)}{\beta\left(m_{\max} - E\left(M_{n-1}\right)\right) + \dfrac{1}{n-1}} = 1 - \exp\left[-\beta\left(m_{\max} - m_{\min}\right)\right].$$

Solving the equation

$$\frac{\left(m_{\max} - E\left(M_{n-1}\right)\right)}{\left(m_{\max} - E\left(M_n\right)\right) + \dfrac{1}{\beta n}} = \frac{\left(m_{\max} - E\left(M_{n-2}\right)\right)}{\left(m_{\max} - E\left(M_{n-1}\right)\right) + \dfrac{1}{\beta(n-1)}},$$

it gives the final result for the $m_{\max}$ as

$$m_{\max} = \frac{(n-1)E\left(M_{n-2}\right)\left[\beta n E\left(M_n\right) - 1\right] - nE\left(M_{n-1}\right)\left[\beta(n-1)E\left(M_{n-1}\right) - 1\right]}{\beta n(n-1)\left[E\left(M_{n-2}\right) - 2E\left(M_{n-1}\right) + E\left(M_n\right)\right] + 1}.$$

A similar way to remove $\beta$, it is to set

$$m_{\max} = \frac{\beta n(n-1)\left[E\left(M_{n-2}\right)E\left(M_n\right) - E^2\left(M_{n-1}\right)\right] + nE\left(M_{n-1}\right) - (n-1)E\left(M_{n-2}\right)}{\beta n(n-1)\left[E\left(M_{n-2}\right) - 2E\left(M_{n-1}\right) + E\left(M_n\right)\right] + 1}$$

$$m_{\max} = \frac{\beta(n-1)(n-2)\left[E\left(M_{n-3}\right)E\left(M_{n-1}\right) - E^2\left(M_{n-2}\right)\right] + (n-1)E\left(M_{n-2}\right) - (n-2)E\left(M_{n-3}\right)}{\beta(n-1)(n-2)\left[E\left(M_{n-3}\right) - 2E\left(M_{n-2}\right) + E\left(M_{n-1}\right)\right] + 1}$$

where it can be found

$$\beta_1 = \frac{1}{(n-1)\left(E\left(M_{n-1}\right) - E\left(M_{n-2}\right)\right)},$$

$$\beta_2 = -\frac{(n-2)E\left(M_{n-3}\right) - 2(n-1)E\left(M_{n-2}\right) + nE\left(M_{n-1}\right)}{n(n-1)(n-2)\left(E^2\left(M_{n-2}\right) + E^2\left(M_{n-1}\right) + E\left(M_{n-3}\right)\left(E\left(M_n\right) - E\left(M_{n-1}\right)\right) - E\left(M_{n-2}\right)\left(E\left(M_{n-1}\right) + E\left(M_n\right)\right)\right)}.$$

The root $\beta_1$ is not a correct solution because it is always positive, since $E\left(M_{n-1}\right) > E\left(M_{n-2}\right)$ and $n \geq 4$.

In the end, the minimum can be written as

$$m_{\min} = m_{\max} + \frac{1}{\beta}\log\left[1 - \frac{\beta\left(m_{\max} - E\left(M_{n-1}\right)\right)}{\beta\left(m_{\max} - E\left(M_{n}\right)\right) + \frac{1}{n}}\right].$$

The algebraic solution is valid when

$$\frac{\beta\left(m_{\max} - E\left(M_{n-1}\right)\right)}{\beta\left(m_{\max} - E\left(M_{n}\right)\right) + \frac{1}{n}} \leq 1 \qquad \Leftrightarrow \qquad n\beta\left(E\left(M_{n}\right) - E\left(M_{n-1}\right)\right) \leq 1.$$

In case $n\beta\left(E\left(M_{n}\right) - E\left(M_{n-1}\right)\right) = 1$, it is $m_{\min} = -\infty$.

## References


Abramowitz, M., and I. A. Stegun (1972). *Handbook of mathematical functions*, 10th ed., Dover Publ., New York.

Gutenberg, B., and C. F. Richter (1944). Frequency of earthquakes in California. *Bull. Seism. Soc. Am.*, *34*, 185-188.

Haarala, M. and L. Orosco (2016a). Analysis of Gutenberg-Richter b-value and $m_{max}$. Part I: Exact solution of Kijko-Sellevoll estimator $m_{max}$, *Cuadernos de Ingeniería. Nueva Serie (9)*, 51-77. https://revistas.ucasal.edu.ar/index.php/CI/article/view/145

Haarala, M. and L. Orosco (2016b). Analysis of Gutenberg-Richter b-value and $m_{max}$. Part II: Estimators for *b*-value and exact variance, *Cuadernos de Ingeniería. Nueva Serie (9)*, 79-106. https://revistas.ucasal.edu.ar/index.php/CI/article/view/146

Haarala, M. (2021). Analysis of Gutenberg-Richter b-value and $m_{max}$. Part III: Non-positive Gutenberg-Richter *b*-value, *Cuadernos de Ingeniería. Nueva Serie 13(XIII)*, 45-84.
https://doi.org/10.53794/ci.v13iXIII.348

Haarala, M. and Vermaulen, P. (2022). The quartile functions for the Generalized Gutenberg-Richter distribution. *Cuadernos De Ingeniería, 14(IV)*, 91-101.
https://doi.org/10.53794/ci.v14iIV.363

Ishimoto, M., and K. Iida (1939). Observations of earthquakes registered with the microseismograph constructed recently. *Bull. Earthquake Res. Inst.*, *17*, 443–478.

ISC, International Seismological Centre (2023), On-line Bulletin, (Last access to the data: 23.8.2023)
https://doi.org/10.31905/D808B830

Pisarenko, V. F., A. A. Lyubushin, V. B. Lysenko, and T. V. Golubieva (1996). Statistical estimation of seismic hazard parameters: Maximum possible magnitude and related parameters. *Bull. Seism. Soc. Am.*, *86*, 691–700.

Rudin, W. (1987). *Principles of Mathematical Analysis*, 3rd ed., 14th printing, McGraw-Hill, Singapore.

**Mika Haarala**

**Perfil académico y profesional:** MSc. in Mathematics, University of Joensuu (Finland). Researcher in Statistics, Applied Mathematics. Member of Seismic Risk Group of Interdisciplinary Studies in Engineering Institute (I.Es.I.Ing).

Correo electrónico: mika.haaralaorosco@gmail.com