Supplemental Material: Correcting exponentiality test for binned earthquake magnitudes

Angela Stallone, Ilaria Spassiani

S1: Non-binned, uniformly dithered magnitudes (purely theoretical)

Let us consider the random variable (r.v.) $M = X + Y$, where: $X \sim \mathcal{E}(\beta)$ and $Y \sim \mathcal{U}(a, b)$ are respectively exponentially and uniformly distributed. The variable M represents the magnitude dithered with uniform noise Y. X is instead the r.v. distributed according to the Gutenberg-Richter (GR) relationship [Gutenberg and Richter, 1944].

The probability density distribution (PDF) of the sum is given by convolving the PDFs of the addends:

$$f_M(m) = (f_X \star f_Y)(m) = \int_{-\infty}^{\infty} f_X(m-z)\, f_Y(z)\, dz.$$

Since, by definition,

$$f_X(x) = \beta e^{-\beta(x-m_0)}, \qquad x \geq m_0, \tag{1}$$

and

$$f_Y(y) = \begin{cases} \frac{1}{b-a}, & \text{if} \quad a < y < b \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

the PDF of the M variable is obtained as

$$f_M(m) = \int_{-\infty}^{\infty} \beta e^{-\beta(m-z-m_0)} \frac{1}{b-a} \mathbb{1}_{(a,b)}(z)\, dz.$$

where $\mathbb{1}_{(a,b)}(z)$ is the indicator function equal to 1 if $z \in (a, b)$, 0 otherwise. Then,

$$f_M(m) = \int_a^b \beta e^{-\beta(m-z-m_0)} \frac{1}{b-a} dz = \frac{e^{-\beta(m-b-m_0)}}{b-a} \left[1 - e^{-\beta(b-a)}\right]. \tag{3}$$

If we now set $b = -a = \frac{\Delta m}{2}$, that is, the noise is uniformly distributed in $(-\frac{\Delta m}{2}, \frac{\Delta m}{2})$, we get

$$f_M(m) = \frac{1}{\Delta m}\, e^{-\beta(m-\frac{\Delta m}{2}-m_0)} \left[1 - e^{-\beta\, \Delta m}\right]$$

which, except for the multiplicative constant $\frac{1}{\Delta m}$, is exactly the same as that obtained for the *non-binned, dithered* case. In the limit for $\Delta m \to 0$, recalling that $\lim_{x\to 0} \frac{s^{tx}-1}{x} = t \ln s$, the classical exponential PDF is eventually recorded.

If we set instead $b = 1 - a = 1$, that is, the noise is uniformly distributed in $(0, 1)$, from Eq. (3) we get

$$f_M(m) = e^{-\beta(x-1)} \left[1 - e^{-\beta}\right] = q\,(1-q)^{x-1}, \tag{4}$$

where we have set $x = m - m_0$ and $q = 1 - e^{-\beta}$. That is, the distribution is geometric with parameter (success probability) q. This proves equivalency between the case of non-binned, dithered magnitudes and non-dithered, binned magnitudes. Specifically, the first case with noise $Y \sim \mathcal{U}(0, 1)$ is exactly the same as the second case with magnitudes in the bin $[0\,,\,\Delta m = 1]$ non-centered.

S2: Check for the Residual Lifetime distribution $f_M(m)$ to be a PDF.

Let us consider the random variable $M = M_i + Y$, where M_i is geometrically distributed with parameter $1 - e^{-\beta\Delta m}$, while Y is uniformly distributed in $(-\frac{\Delta m}{2}, \frac{\Delta m}{2})$. The Residual Lifetime function obtained for M

$$f_M(m) = \begin{cases} \frac{1-e^{-\beta\Delta m}}{\Delta m}\, e^{-\beta\Delta m\,(i_m-1)}, & \text{if} \quad m \in \left(i_m - \frac{\Delta m}{2}, i_m + \frac{\Delta m}{2}\right), \quad i_M \in \mathbb{N}^+ \\ 0 & \text{otherwise,} \end{cases}$$

is actually a probability density function (PDF). In fact:

1. since the three quantities $1 - e^{-\beta\Delta m}$, Δm and $e^{-\beta\Delta m(i-1)}$ $\quad (i = 1, 2, ...)$ are all > 0, it follows that the function $f_M(m)$ is non-negative;
2. the total integral is 1, as we can explicitly compute:

$$\begin{aligned}\int_{-\infty}^{-\infty} f_M(m)dm &= \sum_{i=1}^{\infty}\int_{i-\frac{\Delta m}{2}}^{i+\frac{\Delta m}{2}} f_M(m)dm \\ &= \sum_{i=1}^{\infty}\frac{1-e^{-\beta\Delta m}}{\Delta m}\, e^{-\beta\Delta m\,(i-1)}\left[i + \frac{\Delta m}{2} - i + \frac{\Delta m}{2}\right] \\ &= \left(1 - e^{-\beta\Delta m}\right)\sum_{i=1}^{\infty} e^{-\beta\Delta m\,(i-1)} \\ &= \frac{1-e^{-\beta\Delta m}}{1-e^{-\beta\Delta m}} = 1,\end{aligned}$$

where we used that (geometric series) $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$ for $|x| < 1$.

S3: Remarks

S3.1: Binned magnitudes

We recall that the geometric distribution has the interpretation of the number of failures in a sequence of Bernoulli trials until the first success. If we consider a regime where the probability of success is very small, such that $p = \frac{\lambda}{N}$, in the large N limit it can be proved that the geometric distribution converges to the exponential distribution with parameter λ [e.g. Feller, 1957].

S3.2: Uniformly dithered (binned) magnitudes

To compute the limit, for $\Delta m \to 0$, of the function $f_M(m)$ defined in Eq. 10 in the main text, we can consider the property $\lim_{x\to 0}\frac{s^{tx}-1}{x} = t\,\ln s$, which gives

$$\lim_{\Delta m\to 0}\frac{1-e^{-\beta\Delta m}}{\Delta m} = \lim_{\Delta m\to 0}\beta\frac{e^{\beta\Delta m}-1}{\beta\Delta m} = \beta.$$

Now, recalling that $m - \frac{\Delta m}{2} - m_0 = \Delta m(i-1)$ (see Eqs. 1-4 in the main text), we can observe that the classical exponential PDF is recovered when the bin width Δm approaches 0. In this limiting case, the exponential and uniform mean points coincide within each Δm-bin used to group continuous magnitudes sampled from an exponential distribution. This result also follows directly from the memoryless property of exponentially distributed random variables [Marzocchi et al., 2019].

S3.3: Exact random noise to re-obtain the exponential distribution

Since Mi is defined starting from $i = 1$ (see Eq. 1 in the main text), while the support of Y is $[0, \Delta m)$, it could be more intuitive to select as non-zero term in the sum 13 in the main text the right extreme of the interval $[k\Delta m, (k+1)\Delta m)$, yielding $f_M(m) = \sum_{k=1}^{\infty} P(M_i = k+1) f_Y(m - (k+1)\Delta m)$, slightly more complicated calculations and obviously the same result as in Eq. 14 in the main text.

S4: Algorithm 1

The following procedure summarizes the numerical test used to assess how the rejection probability of the Lilliefors test varies with bin width Δm and catalog size when magnitudes are dithered with uniform noise (see Section 1.2.2 in the main text).

Algorithm 1 Lilliefors rejection probability as function of bin width and catalog size - Uniform noise

$M_{\min} = 1.0$
$\alpha = 0.1$
$b = 1$ ($\beta = b \cdot \ln(10)$)
$N_{\text{SIM}} = 1000$ ▷ number of simulated catalogs per (Δm, catalog size)
$N_{\text{NOISE}} = 100$ ▷ number of independent noise realizations per catalog
for each bin width Δm **do**
    for each catalog size n_{events} **do**
        counter $= 0$
        for $j = 1$ to N_{SIM} **do** ▷ Simulate one catalog
            for $i = 1$ to n_{events} **do** ▷ Generate binned magnitudes
                Draw $k_i \in \{0, 1, 2, \dots\}$ with $\Pr(k_i = k) = q(1-q)^k$, $q = 1 - e^{-\beta \Delta m}$
                Set $M_i = M_{\min} + k_i \cdot \Delta m$
            end for
            Initialize empty list P ▷ store p-values
            for $k = 1$ to N_{NOISE} **do** ▷ Dither magnitudes
                Draw noise $\varepsilon_i \sim \mathcal{U}(-\Delta m/2, \Delta m/2)$ for $i = 1, \dots, n_{\text{events}}$
                $M_i^{\text{dithered}} = M_i + \varepsilon_i$
                Shift magnitudes: $M_i^{\text{shifted}} = M_i^{\text{dithered}} - (M_{\min} - \Delta m/2)$
                Apply Lilliefors test: $p_k = \text{Lilliefors}(M^{\text{shifted}}, \text{dist} = \exp)$
                Append p_k to P
            end for
            Compute mean p-value: $\bar{p} = \text{mean}(P)$
            if $\bar{p} < \alpha$ **then** *counter* $\leftarrow$ *counter* $+ 1$
            end if
        end for
        Rejection probability: $R = 100 \cdot \frac{counter}{N_{\text{SIM}}}$
    end for
end for

S5: Algorithm 2

The following procedure summarizes the numerical test described in Section 1.3.1 of the main text. Here, the binned magnitudes are dithered by adding a truncated exponential random variable supported on $[0, \Delta m)$.

Algorithm 2 Lilliefors rejection probability as function of bin width and catalog size - Truncated exponential noise

$M_{\min} = 1.0$
$\alpha = 0.1$
$b = 1$ ($\beta = b \cdot \ln(10)$)
$N_{\text{SIM}} = 1000$ ▷ number of simulated catalogs per (Δm, catalog size)
$N_{\text{NOISE}} = 100$ ▷ number of independent noise realizations per catalog
for each bin width Δm **do**
  for each catalog size n_{events} **do**
    counter $= 0$
    for $j = 1$ to N_{SIM} **do** ▷ Simulate one catalog
      for $i = 1$ to n_{events} **do** ▷ Generate binned magnitudes
        Draw $k_i \in \{0, 1, 2, \ldots\}$ with $\Pr(k_i = k) = q(1-q)^k$, $q = 1 - e^{-\beta \Delta m}$
        Set $M_i = M_{\min} + k_i \cdot \Delta m$
      end for
      Initialize empty list P ▷ store p-values
      for $k = 1$ to N_{NOISE} **do** ▷ Dither magnitudes
        Draw noise

$$\varepsilon_i = -\frac{1}{\beta} \ln\Big(1 - U_i(1 - e^{-\beta \Delta m})\Big), \quad U_i \sim \mathcal{U}(0,1),\ i = 1, \ldots, n_{\text{events}}$$

        $M_i^{\text{dithered}} = M_i + \varepsilon_i$
        Shift magnitudes: $M_i^{\text{shifted}} = M_i^{\text{dithered}} - M_{\min}$
        Apply Lilliefors test: $p_k = \text{Lilliefors}(M^{\text{shifted}}, \text{dist} = \exp)$
        Append p_k to P
      end for
      Compute mean p-value: $\bar{p} = \text{mean}(P)$
      if $\bar{p} < \alpha$ **then** *counter* $\leftarrow$ *counter* $+ 1$
      end if
    end for
    Rejection probability: $R = 100 \cdot \frac{counter}{N_{\text{SIM}}}$
  end for
end for

References

William Feller. *An Introduction to Probability Theory and Its Applications, Volume 2.* Number v. 1-2 in An Introduction to Probability Theory and Its Applications. Wiley, 1957. ISBN 9780471257097. doi: 10.1137/1014119. URL `https://books.google.it/books?id=BsSwAAAAIAAJ`.

Beno Gutenberg and Charles F. Richter. Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, 34(4):185–188, 1944. doi: 10.1785/BSSA0340040185.

Warner Marzocchi, Ilaria Spassiani, Angela Stallone, and Matteo Taroni. How to be fooled searching for significant variations of the b-value. *Geophysical Journal International*, 220(3):1845–1856, 11 2019. ISSN 0956-540X. doi: 10.1093/gji/ggz541. URL `https://doi.org/10.1093/gji/ggz541`.