Records for the moving average of a time series

Claude Godrèche, Jean-Marc Luck

Records for the moving average of a time series

Claude Godrèche and Jean-Marc Luck

Institut de Physique Théorique, Université Paris-Saclay, CEA and CNRS, 91191 Gif-sur-Yvette, France

Abstract. We investigate how the statistics of extremes and records is affected when taking the moving average over a window of width p of a sequence of independent, identically distributed random variables. An asymptotic analysis of the general case, corroborated by exact results for three distributions (exponential, uniform, power-law with unit exponent), evidences a very robust dichotomy, irrespective of the window width, between superexponential and subexponential distributions. For superexponential distributions the statistics of records is asymptotically unchanged by taking the moving average, up to interesting distribution-dependent corrections to scaling. For subexponential distributions the probability of record breaking at late times is increased by a universal factor R_p, depending only on the window width.

E-mail: `claude.godreche@ipht.fr`,`jean-marc.luck@ipht.fr`

1. Introduction

When monitoring a time series, a feature which immediately attracts the attention of the observer is the sequence of record values, viz., the successive largest or smallest values in the series [1, 2, 3]. The first example which comes to mind are weather records, i.e., the extreme occurrences of weather phenomena such as the coldest or hottest days, the most rainy or windy days, and so on, for which studies abound (see [4, 5, 6, 7, 8, 9] and references therein). Other examples of records encountered in diverse complex physical systems are reviewed in [10], to which the reader is referred for a comprehensive list of references.

The simplest situation to analyse is when the data are samples of a sequence of independent, identically distributed (iid) random variables. In such an instance much is known on the statistics of records [1, 2, 3, 4, 11, 12], whose basics are easy to grasp. Consider a sequence of iid continuous random variables $X_1, X_2, \ldots$, with common distribution function $F(x) = \mathrm{Prob}(X < x)$ and density $f(x) = \mathrm{d}F(x)/\mathrm{d}x$. Throughout the following we assume that the X_i are positive. A record is said to occur at step n if X_n is larger than all previous variables, i.e., if

$$X_n > L_{n-1} = \max(X_1, X_2, \ldots, X_{n-1}),$$

where L_n denotes the largest X_i amongst the first n random variables. The probability of this event, or probability of record breaking,

$$Q_n = \mathrm{Prob}(X_n > L_{n-1}),$$

equals

$$Q_n = \frac{1}{n}, \tag{1.1}$$

as a consequence of the fact that the random variables X_i are exchangeable [2, 3]. The number M_n of records up to time n takes the values $1, \dots, n$ and can be expressed as the sum

$$M_n = I_1 + I_2 + \cdots + I_n, \tag{1.2}$$

where the indicator variable I_n is equal to 1 if X_n is a record and to 0 otherwise. Taking the average, we have $\langle I_i \rangle = Q_i = 1/i$, and so

$$\langle M_n \rangle = \sum_{i=1}^{n} \frac{1}{i} = H_n \approx \ln n + \gamma, \tag{1.3}$$

where H_n is the nth harmonic number and $\gamma = 0.577215\dots$ is Euler's constant. It is a simple matter to show that the indicator variables $I_1, I_2, \dots, I_n$ are statistically independent [2, 3, 10]. The distribution of M_n ensues from this fact by elementary considerations (see also section 7.4 below). The simple expression (1.1) of the probability of record breaking and the full distribution of M_n are universal, in the sense that they do not depend on the underlying distribution $f(x)$. From this standpoint the statistics of records for iid random variables exhibits a high degree of degeneracy. In contrast, the statistics of the extreme value L_n is distribution dependent, as is well known [13].

In the present work we investigate the statistics of records for sequences made of sums of p successive iid positive random variables, defined as follows. For $p = 2$,

$$Y_2 = X_1 + X_2, \quad Y_3 = X_2 + X_3, \dots, \quad Y_n = X_{n-1} + X_n, \ \dots, \tag{1.4}$$

for $p = 3$,

$$Y_3 = X_1 + X_2 + X_3, \dots, \quad Y_n = X_{n-2} + X_{n-1} + X_n, \ \dots, \tag{1.5}$$

or more generally,

$$Y_p = X_1 + \cdots + X_p, \dots, \quad Y_n = X_{n-p+1} + \cdots + X_n, \ \dots \tag{1.6}$$

The first terms of these sequences, which have not been written down explicitly, may be omitted in the analysis of records. For instance, in (1.4), $Y_1 = X_1$ is always smaller than Y_2. In (1.5), $Y_1 = X_1$ and $Y_2 = X_1 + X_2$ are always smaller than Y_3, and so on.

Up to a normalisation, each of these sequences can be seen as the *moving average* of the sequence of iid variables $X_1, X_2, \dots$, defined as the mean of the last p terms. For instance the moving average with $p = 2$ is

$$\frac{X_1 + X_2}{2}, \quad \frac{X_2 + X_3}{2}, \dots, \quad \frac{X_{n-1} + X_n}{2}, \ \dots$$

Taking the moving average is a well-known method to analyse time series, which is equivalent to making a convolution of the signal by a square window, thus smoothing the signal. For instance, instead of looking at the daily temperature at a given location, one can take the moving average over a period of one week, corresponding to choosing $p = 7$. The question posed here amounts therefore to knowing how records are affected by taking such an average. The normalisation by the factor p does not affect the outcome of the subsequent analysis.

Here the focus will be essentially on the particular case $p = 2$. Keeping the same notations as for the iid case, we shall primarily investigate the probability of record breaking,

$$Q_n = \mathrm{Prob}(Y_n > L_{n-1}), \tag{1.7}$$

where L_n denotes the largest Y_i amongst the first n ones,

$$L_n = \max(Y_1, Y_2, \ldots, Y_n), \tag{1.8}$$

and the mean number of records up to n,

$$\langle M_n \rangle = \sum_{i=2}^{n} Q_i, \tag{1.9}$$

where records are counted from the first complete sum Y_2 onwards. As we shall see, these quantities are now sensitive to the choice of the underlying distribution $f(x)$ of the parent random variables X_i. On the one hand, this does not come as a surprise since the new variables Y_i are no longer exchangeable, and the occurrences of records at various places are no longer independent. On the other hand, it is yet slightly paradoxical that the degeneracy induced by the exchangeability of the iid parent random variables X_i is now removed, revealing features of their common distribution, since by taking the moving average one could have expected a loss of information instead. We shall also study some features of the distribution of L_n.

In a nutshell, the main outcome of this work is as follows. We find that the product nQ_n has only two possible limits for $p = 2$, depending on the class of distribution $f(x)$, namely

$$nQ_n \to 1 \tag{1.10}$$

for superexponential distributions, that is, distributions either having a bounded support or falling off faster than any exponential, whereas

$$nQ_n \to \frac{3}{2} \tag{1.11}$$

for subexponential distributions, whose tails decrease more slowly than any exponential. The pure exponential distribution belongs to the first class, albeit marginally. Figure 1 shows a plot of nQ_n against $n \geq 4$ for all the examples of probability distributions $f(x)$ considered in the present paper (see table 1). Each dataset is the outcome of the numerical generation of 10^{10} sequences. The vertical arrow underlines that the dichotomy between (1.10) and (1.11) becomes more and more visible as n increases. The values of Q_n for $n = 2, 3, 4$ are universal, i.e., independent of the underlying distribution $f(x)$ (see section 2.6).

For higher values of the window width p, denoting the probability of record breaking by $Q_n^{(p)}$, (1.10) still holds for superexponential distributions, i.e.,

$$nQ_n^{(p)} \to 1, \tag{1.12}$$

while, for subexponential distributions, (1.11) becomes

$$nQ_n^{(p)} \to R_p, \tag{1.13}$$

where the R_p are universal rational numbers given by

$$R_p = \frac{3}{2}, \frac{15}{8}, \frac{35}{16}, \frac{315}{128}, \frac{693}{256}, \ldots \tag{1.14}$$

for $p = 2, 3, 4, 5, 6, \ldots$, and obtained by means of the Sparre Andersen theorem.

The setup of this paper is as follows. Sections 2 to 6 concern the case $p = 2$. In section 2 we present the general setting which will be used in all subsequent exact or asymptotic developments. The next three sections are devoted to exact analytical solutions of the problem for three distributions: the exponential distribution (section 3), the uniform distribution (section 4), and the power-law distribution with

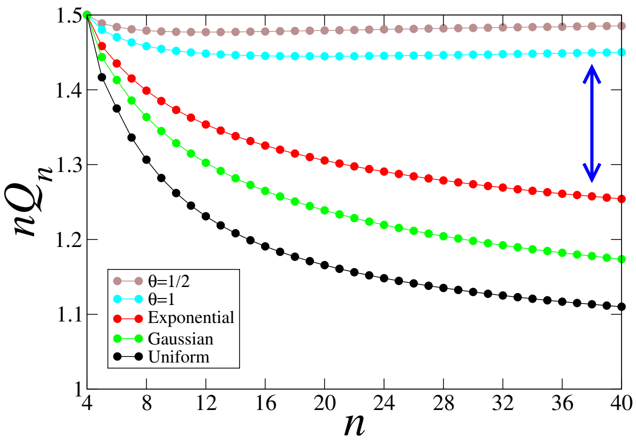

Figure 1. Plot of n times the probability of record breaking Q_n against $n \geq 4$ for several probability distributions $f(x)$. From top to bottom: power-law distributions with tail index $\theta = 1/2$ and $\theta = 1$, exponential, half-Gaussian, and uniform distributions (see table 1).

Distribution	$F(x)$	$f(x)$	support
Uniform	x	1	$0 < x < 1$
Exponential	$1 - \mathrm{e}^{-x}$	e^{-x}	$x > 0$
Half-Gaussian	$\mathrm{erf}\, x$	$\frac{2}{\sqrt{\pi}}\,\mathrm{e}^{-x^2}$	$x > 0$
Power-law ($\theta > 0$)	$1 - x^{-\theta}$	$\theta\, x^{-1-\theta}$	$x > 1$

Table 1. Distribution function $F(x)$, density $f(x)$ and support of the examples of probability distributions considered in this work.

index $\theta = 1$ (section 5). In order to compare the probability of record breaking to its universal value $Q_n = 1/n$ in the iid situation (see (1.1)), we set

$$nQ_n = 1 + \delta_n. \tag{1.15}$$

The exponential distribution appears as a marginal case where (1.10) holds, albeit with a logarithmic correction

$$\delta_n \approx \frac{1}{\ln n}.$$

For the uniform distribution δ_n falls off as

$$\delta_n \approx \sqrt{\frac{\pi}{8n}},$$

whereas (1.11) holds for the power-law distribution with $\theta = 1$. A heuristic asymptotic analysis of the general case is then performed in section 6, where the dichotomy between (1.10) and (1.11) is explained in simple terms, and an estimate for the relative correction δ_n is derived. Higher values of the window width p are considered in section 7 along the same line of thought. The overall picture, including the dichotomy between (1.10) and (1.11), remains unchanged. The non-trivial limit 3/2 in (1.11) is replaced by the p-dependent but otherwise universal limit R_p (1.14). Section 8 contains a brief discussion of our findings.

Let us finally mention that the statistics of persistent events for the sequence (1.4) has been studied in [14, 15, 16], for the case where the parent variables X_i have a symmetric distribution $f(x)$.

2. General setting

This section sets the basis of all subsequent exact or asymptotic developments. Hereafter and until the end of section 6 we focus our attention on the sequence (1.4) of sums of two terms. Higher values of the width p will be considered in section 7.

2.1. Recursive structure

We start by highlighting the recursive structure of the problem. The first two maxima are necessarily $L_1 = Y_1 = X_1$ and $L_2 = Y_2 = X_1 + X_2$. The next maxima obey the recursion

$$L_n = \max(L_{n-1}, Y_n) = \begin{cases} Y_n & \text{if } Y_n > L_{n-1}, \\ L_{n-1} & \text{if } Y_n < L_{n-1}. \end{cases} \tag{2.1}$$

This recursion should be understood as follows. Starting from the couple of random variables (L_{n-1}, X_{n-1}), one draws the random variable X_n, which is independent of L_{n-1} and X_{n-1}, and sets $Y_n = X_{n-1} + X_n$. This generates the new L_n, or alternatively the new couple (L_n, X_n):

$$(L_{n-1}, X_{n-1}) \overset{X_n}{\rightsquigarrow} (L_n, X_n).$$

In other words, at each time step n, the newly drawn random variable X_n acts as a noise on the dynamics of the couple (L_{n-1}, X_{n-1}). The value of L_n depends on the branch of the recursion, denoted respectively by **L** (for larger) and **S** (for smaller):

$$\begin{aligned} &(\mathbf{L}) : Y_n > L_{n-1} \Longrightarrow L_n = Y_n, \\ &(\mathbf{S}) : Y_n < L_{n-1} \Longrightarrow L_n = L_{n-1}. \end{aligned} \tag{2.2}$$

In the first case, Y_n is a record since it satisfies

$$Y_n > \max(Y_1, \ldots, Y_{n-1}).$$

This event occurs with probability Q_n (see (1.7)). Hereafter we make use of the recursion (2.1) to derive the key relations (2.14) and (2.15) for the functions $F_n(\ell, x)$ introduced in (2.3).

Let us mention that a similar, albeit simpler recursive scheme applies to the theory of records for iid random variables.

2.2. Basic quantities

Starting from the joint distribution function of the couple of random variables (L_n, X_n),

$$\text{Prob}(L_n < \ell, X_n < x),$$

and taking its derivative with respect to x, yields the quantity

$$F_n(\ell,x) = \partial_x \operatorname{Prob}(L_n < \ell, X_n < x),$$

which plays a central role in the present work. It is equivalently defined as

$$F_n(\ell,x)\mathrm{d}x = \operatorname{Prob}(L_n < \ell, x < X_n < x+\mathrm{d}x). \tag{2.3}$$

The underlying distribution $f(x)$ is recovered in the $\ell \to \infty$ limit:

$$F_n(\infty,x) = f(x).$$

By differentiating $F_n(\ell,x)$ with respect to ℓ, one gets the joint probability density of the couple (L_n, X_n):

$$f_n(\ell,x) = \partial_\ell F_n(\ell,x), \quad F_n(\ell,x) = \int_0^\ell \mathrm{d}\ell'\, f_n(\ell',x).$$

Conversely, integrating on the second variable restores

$$\operatorname{Prob}(L_n < \ell, X_n < x) = \int_0^x \mathrm{d}x'\, F_n(\ell,x').$$

In particular the distribution function of the maximum L_n is obtained when the integral runs over its full range (i.e., $x = \ell$):

$$\mathcal{F}_n(\ell) = \operatorname{Prob}(L_n < \ell) = \int_0^\ell \mathrm{d}x'\, F_n(\ell,x'). \tag{2.4}$$

Its derivative with respect to ℓ yields the density $f_{L_n}(\ell)$. The determination of the mean maximum ensues:

$$\langle L_n \rangle = \int_0^\infty \mathrm{d}\ell (1 - \mathcal{F}_n(\ell)). \tag{2.5}$$

Finally, the normalization of the joint density $f_n(\ell,x)$ implies

$$\int_0^\infty \mathrm{d}\ell \int_0^\ell \mathrm{d}x\, f_n(\ell,x) = \int_0^\infty \mathrm{d}x \int_x^\infty \mathrm{d}\ell\, f_n(\ell,x) = 1. \tag{2.6}$$

2.3. First values of n

The quantities defined above have explicit expressions for $n = 1$ and 2 in full generality.

For $n = 1$ we have

$$\operatorname{Prob}(L_1 < \ell, X_1 < x) = \operatorname{Prob}(X_1 < x) = F(x),$$

whenever $x < \ell$, since $L_1 = X_1$. Differentiating with respect to x gives

$$F_1(\ell,x) = f(x) \tag{2.7}$$

and

$$\mathcal{F}_1(\ell) = F(\ell).$$

For $n = 2$, knowing that $L_2 = Y_2 = X_1 + X_2$ allows one to compute

$$\operatorname{Prob}(L_2 < \ell, X_2 < x) = F(x)F(\ell - x) + \int_{\ell-x}^{\ell} \mathrm{d}x_1\, f(x_1)F(\ell - x_1),$$

from which $F_2(\ell,x)$ ensues by derivation with respect to x:

$$F_2(\ell,x) = f(x)F(\ell - x), \tag{2.8}$$

consistently with the definition (with informal notation)

$$\mathrm{Prob}(L_2 = X_1 + X_2 < \ell, X_2 = x) = \mathrm{Prob}(X_1 < \ell - x, X_2 = x).$$

Then, taking a derivative with respect to ℓ, we have

$$f_2(\ell, x) = f(x) f(\ell - x),$$

and finally

$$\mathcal{F}_2(\ell) = \int_0^\ell \mathrm{d}x\, f(x) F(\ell - x).$$

2.4. Recursion relation for the function $F_n(\ell, x)$

The recursion (2.1) implies

$$F_n(\ell, x) = f(x) \int \mathrm{d}\ell' \mathrm{d}x'\, f_{n-1}(\ell', x') \Theta(\ell - \max(\ell', x' + x)), \qquad (2.9)$$

where Θ denotes Heaviside function. The right-hand side of this equation decomposes into two contributions, associated to the two branches **L** and **S**,

$$F_n(\ell, x) = f(x) \int_{D_{\mathbf{L}}} \mathrm{d}\ell' \mathrm{d}x'\, f_{n-1}(\ell', x') + f(x) \int_{D_{\mathbf{S}}} \mathrm{d}\ell' \mathrm{d}x'\, f_{n-1}(\ell', x'), \qquad (2.10)$$

where the domains $D_{\mathbf{L}}$ and $D_{\mathbf{S}}$, depicted in figure 2, are respectively defined as

$$D_{\mathbf{L}} = \{\ell' < x + x' < \ell\},$$
$$D_{\mathbf{S}} = \{x + x' < \ell' < \ell\},$$

hence

$$\int_{D_{\mathbf{L}}} \mathrm{d}\ell' \mathrm{d}x'\, f_{n-1}(\ell', x') = \int_0^{\ell - x} \mathrm{d}x' \int_{x'}^{x + x'} \mathrm{d}\ell'\, f_{n-1}(\ell', x'), \qquad (2.11)$$

$$\int_{D_{\mathbf{S}}} \mathrm{d}\ell' \mathrm{d}x'\, f_{n-1}(\ell', x') = \int_0^{\ell - x} \mathrm{d}x' \int_{x + x'}^{\ell} \mathrm{d}\ell'\, f_{n-1}(\ell', x'). \qquad (2.12)$$

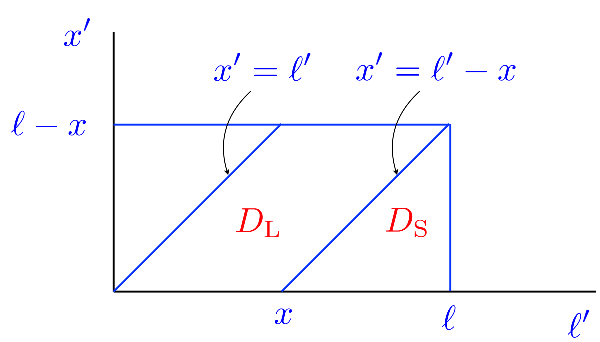

Figure 2. The domains $D_{\mathbf{L}}$ and $D_{\mathbf{S}}$ in the (ℓ', x') plane.

Adding these two contributions yields

$$F_n(\ell, x) = f(x) \int_0^{\ell - x} \mathrm{d}x' \int_{x'}^{\ell} \mathrm{d}\ell'\, f_{n-1}(\ell', x'),$$

which vanishes whenever its two arguments are equal ($n \geq 2$):

$$F_n(\ell, \ell) = 0. \tag{2.13}$$

We thus obtain the following recursion relation for the function $F_n(\ell, x)$:

$$F_n(\ell, x) = f(x) \int_0^{\ell - x} \mathrm{d}x' \, F_{n-1}(\ell, x'). \tag{2.14}$$

This equation and its differential form

$$F_n'(\ell, x) = \frac{f'(x)}{f(x)} F_n(\ell, x) - f(x) F_{n-1}(\ell, \ell - x), \tag{2.15}$$

obtained by differentiating (2.14) with respect to x ‡, are key formulas of this work and the starting points of many subsequent developments.

As a consequence of (2.14), we have ($n \geq 2$)

$$F_n(\ell, 0) = f(0) \mathcal{F}_{n-1}(\ell). \tag{2.16}$$

Finally, differentiating (2.10) with respect to ℓ yields

$$f_n(\ell, x) = \mathbf{L} f_{n-1}(\ell, x) + \mathbf{S} f_{n-1}(\ell, x), \tag{2.17}$$

with the notations

$$\begin{aligned} \mathbf{L} f_{n-1}(\ell, x) &= f(x) \int_{\ell - x}^{\ell} \mathrm{d}\ell' \, f_{n-1}(\ell', \ell - x) \\ &= f(x) F_{n-1}(\ell, \ell - x), \end{aligned} \tag{2.18}$$

$$\mathbf{S} f_{n-1}(\ell, x) = f(x) \int_0^{\ell - x} \mathrm{d}x' \, f_{n-1}(\ell, x'), \tag{2.19}$$

using (2.11) and (2.12). Alternatively, differentiating (2.14) with respect to ℓ gives

$$f_n(\ell, x) = f(x) F_{n-1}(\ell, \ell - x) + f(x) \int_0^{\ell - x} \mathrm{d}x' \, f_{n-1}(\ell, x'), \tag{2.20}$$

which is identical to (2.17).

2.5. Probability of record breaking

The probability of record breaking Q_n is the probability that the last variable is larger than all previous ones (see (1.7)),

$$Q_n = \mathrm{Prob}(Y_n > L_{n-1}).$$

This probability thus equals the weight of branch $\mathbf{L}$. For, recalling (2.6) and (2.17),

$$\begin{aligned} 1 &= \int_0^\infty \mathrm{d}\ell \int_0^\ell \mathrm{d}x \, f_n(\ell, x) \\ &= \int_0^\infty \mathrm{d}\ell \int_0^\ell \mathrm{d}x \, \left(\mathbf{L} f_{n-1}(\ell, x) + \mathbf{S} f_{n-1}(\ell, x) \right), \end{aligned}$$

where the two terms corresponding respectively to the weights of the two branches $\mathbf{L}$ and $\mathbf{S}$ are Q_n and $1 - Q_n$. So the expression of Q_n is ($n \geq 2$)

$$\begin{aligned} Q_n &= \int_0^\infty \mathrm{d}\ell \int_0^\ell \mathrm{d}x \, \mathbf{L} f_{n-1}(\ell, x) \\ &= \int_0^\infty \mathrm{d}\ell \int_0^\ell \mathrm{d}x \, f(x) F_{n-1}(\ell, \ell - x) \\ &= \int_0^\infty \mathrm{d}\ell \int_0^\ell \mathrm{d}x \, f(\ell - x) F_{n-1}(\ell, x). \end{aligned} \tag{2.21}$$

‡ Throughout the following, accents on functions denote their (partial) derivatives with respect to x.

2.6. Universal values of the probability of record breaking

The first few values of Q_n are universal, i.e., independent of the underlying distribution $f(x)$. For $n = 2$,

$$Q_2 = 1, \tag{2.22}$$

since $Y_2 = X_1 + X_2$ is always larger that $Y_1 = X_1$. This result can be recovered by inserting (2.7) into (2.21). For $n = 3$,

$$Q_3 = \frac{1}{2}, \tag{2.23}$$

since $Y_3 > Y_2$ is equivalent to $X_3 > X_1$, which holds with probability 1/2. This result can be recovered by inserting (2.8) into (2.21). It turns out that for $n = 4$, Q_n has also a universal value,

$$Q_4 = \frac{3}{8}, \tag{2.24}$$

irrespective of the distribution $f(x)$. This can be demonstrated by a simple application of the Sparre Andersen theorem [17, 18, 19]. This theorem states in particular that, for a sequence of iid variables Z_n with a continuous symmetric distribution, the probability that the first n partial sums are all positive,

$$P_n = \mathrm{Prob}(Z_1 > 0, Z_1 + Z_2 > 0, \ldots, Z_1 + Z_2 + \cdots + Z_n > 0),$$

is a universal rational number,

$$P_n = \frac{1}{2^{2n}}\binom{2n}{n} = \frac{(2n)!}{(2^n n!)^2} = 1,\ \frac{1}{2},\ \frac{3}{8},\ \frac{5}{16},\ \frac{35}{128},\ \frac{63}{256}, \ldots, \tag{2.25}$$

for $n = 0, 1, 2, 3, 4, 5, \ldots$, with generating function

$$\widetilde{P}(z) = \sum_{n \ge 0} z^n P_n = \frac{1}{\sqrt{1-z}}. \tag{2.26}$$

In the present case, by definition,

$$\begin{aligned}
Q_4 &= \mathrm{Prob}(Y_4 > Y_3, Y_4 > Y_2) \\
&= \mathrm{Prob}(X_4 - X_2 > 0, X_4 - X_2 + X_3 - X_1 > 0) \\
&= \mathrm{Prob}(Z_1 > 0, Z_1 + Z_2 > 0),
\end{aligned}$$

where the random variables $Z_1 = X_4 - X_2$ and $Z_2 = X_3 - X_1$ are iid and have a continuous symmetric distribution. Therefore the theorem applies and $Q_4 = P_2$, which is the result announced in (2.24). It would be cumbersome to recover this directly by means of (2.21).

The probability of record breaking Q_n is no longer universal for $n \ge 5$. It is indeed clear from figure 1 that already Q_5 depends on the underlying distribution $f(x)$.

It results from the foregoing that the first values of the mean number of records (see (1.9))

$$\langle M_2 \rangle = 1, \quad \langle M_3 \rangle = \frac{3}{2}, \quad \langle M_4 \rangle = \frac{15}{8},$$

are also universal.

In the forthcoming sections we apply the general formalism presented in this section to derive analytical solutions of the differential recursion (2.15) for the exponential distribution (section 3), the uniform distribution (section 4), and the power-law distribution with index $\theta = 1$ (section 5).

3. Exponential distribution

This section presents an exact solution of the problem for the case of exponentially distributed random variables X_i, with common density $f(x) = \mathrm{e}^{-x}$ and distribution function $F(x) = 1 - \mathrm{e}^{-x}$ (see table 1).

3.1. Differential equations

The exact solutions derived in this section and in the two subsequent ones rely on the differential equation (2.15), which reads, in the present case,

$$F_n'(\ell, x) + F_n(\ell, x) = -\mathrm{e}^{-x} F_{n-1}(\ell, \ell - x). \tag{3.1}$$

Differentiating once more yields ($n \geq 3$)

$$F_n''(\ell, x) + F_n'(\ell, x) + \mathrm{e}^{-\ell} F_{n-2}(\ell, x) = 0. \tag{3.2}$$

This is a recursive differential equation in the variable x, while ℓ plays the role of a parameter. Setting $x = 0$ in (3.1) gives ($n \geq 3$)

$$F_n(\ell, 0) + F_n'(\ell, 0) = 0, \tag{3.3}$$

where the interpretation of the first term is given in (2.16).

3.2. First values of n

For the first few values of n, we obtain

$$F_1(\ell, x) = \mathrm{e}^{-x}, \quad F_2(\ell, x) = \mathrm{e}^{-x} - \mathrm{e}^{-\ell},$$
$$F_3(\ell, x) = \mathrm{e}^{-x} - (\ell - x)\mathrm{e}^{-x-\ell} - \mathrm{e}^{-\ell}.$$

Inserting these expressions into (2.21), we recover the universal results for Q_2, Q_3 and Q_4 derived in section 2.6. Equations (2.13) and (3.3) are complemented by

$$\begin{aligned} &F_1(\ell, \ell) = \mathrm{e}^{-\ell}, \quad F_1(\ell, 0) + F_1'(\ell, 0) = 0, \\ &F_2(\ell, 0) + F_2'(\ell, 0) = -\mathrm{e}^{-\ell}. \end{aligned} \tag{3.4}$$

We have also

$$\mathcal{F}_1(\ell) = 1 - \mathrm{e}^{-\ell}, \quad \mathcal{F}_2(\ell) = 1 - (\ell + 1)\mathrm{e}^{-\ell},$$
$$\mathcal{F}_3(\ell) = 1 - \mathrm{e}^{-2\ell} - 2\ell\mathrm{e}^{-\ell}\ell.$$

Inserting these expressions into (2.5) yields $\langle L_1 \rangle = 1$, $\langle L_2 \rangle = 2$ and $\langle L_3 \rangle = 5/2$.

3.3. Generating function

In order to solve the recursive differential equation (3.2) for all values of n, we introduce the generating function

$$\widetilde{F}(z, \ell, x) = \sum_{n \geq 1} z^n F_n(\ell, x), \tag{3.5}$$

which satisfies (using (3.2))

$$\widetilde{F}''(z, \ell, x) + \widetilde{F}'(z, \ell, x) + z^2 \mathrm{e}^{-\ell} \widetilde{F}(z, \ell, x) = 0,$$

the solution of which is

$$\widetilde{F}(z, \ell, x) = A_+ \,\mathrm{e}^{a_+ x} + A_- \,\mathrm{e}^{a_- x}, \tag{3.6}$$

with

$$a_\pm = \frac{1 \pm w}{2}, \quad w = \sqrt{1 - 4z^2 \mathrm{e}^{-\ell}}.$$

The amplitudes $A_\pm$ are determined by the boundary conditions (see (3.4))

$$\widetilde{F}(z, \ell, 0) + \widetilde{F}'(z, \ell, 0) = -z^2 \mathrm{e}^{-\ell}, \quad \widetilde{F}(z, \ell, \ell) = z\mathrm{e}^{-\ell},$$

yielding

$$A_\pm = \pm z \frac{a_\pm \mathrm{e}^{-\ell/2} + z\mathrm{e}^{\pm w\ell/2 - \ell}}{w \cosh \frac{w\ell}{2} - \sinh \frac{w\ell}{2}}.$$

3.4. Probability of record breaking

Using (2.21), the generating function of the Q_n reads

$$\widetilde{Q}(z) = \sum_{n \geq 2} z^n Q_n = z \int_0^\infty \mathrm{d}\ell \, \mathrm{e}^{-\ell} I(z, \ell), \tag{3.7}$$

with

$$\begin{aligned} I(z, \ell) &= \int_0^\ell \mathrm{d}x \, \mathrm{e}^x \widetilde{F}(z, \ell, x) \\ &= A_+ \frac{\mathrm{e}^{(1-a_+)\ell} - 1}{1 - a_+} + A_- \frac{\mathrm{e}^{(1-a_-)\ell} - 1}{1 - a_-} = \frac{N(z, \ell)}{D(z, \ell)}, \end{aligned} \tag{3.8}$$

and

$$N(z, \ell) = 4z(1 - z)\mathrm{e}^{-\ell/2} + (1 - w^2 - 4z) \cosh \frac{w\ell}{2} + (2(1 + w^2)z + 1 - w^2) \frac{1}{w} \sinh \frac{w\ell}{2},$$

$$D(z, \ell) = (w^2 - 1) \left(\cosh \frac{w\ell}{2} - \frac{1}{w} \sinh \frac{w\ell}{2} \right).$$

The integral over ℓ in (3.7) cannot be carried out in closed form. By expanding $I(z, \ell)$ as a power series in z and integrating term by term with respect to ℓ, we obtain the values of the probability of record breaking Q_n and mean number of records $\langle M_n \rangle$ given in table 2 up to $n = 8$.

n	2	3	4	5	6	7	8
Q_n	1	$\frac{1}{2}$	$\frac{3}{8}$	$\frac{7}{24}$	$\frac{155}{648}$	$\frac{131}{648}$	$\frac{14503}{82944}$
$\langle M_n \rangle$	1	$\frac{3}{2}$	$\frac{15}{8}$	$\frac{13}{6}$	$\frac{1559}{648}$	$\frac{845}{324}$	$\frac{8549}{3072}$

Table 2. Exact values of the probability of record breaking Q_n and mean number of records $\langle M_n \rangle$ up to $n = 8$, for an exponential distribution. Expressions to the left of the double bar are universal.

The asymptotic decay of Q_n at large n can be derived as follows. Setting $z = \mathrm{e}^{-\varepsilon}$, (3.8) becomes

$$I(z, \ell) \approx \frac{\ell - 1}{\varepsilon + (\ell - 1)\mathrm{e}^{-\ell}},$$

in the relevant regime where ε and $\mathrm{e}^{-\ell}$ are simultaneously small. Inserting this into (3.7), and dealing with n as a continuous variable, we obtain the estimate

$$\widetilde{Q}(z) \approx \int_0^\infty \mathrm{d}n\, \mathrm{e}^{-n\varepsilon} Q_n \approx \int_1^\infty \mathrm{d}\ell\, \frac{(\ell-1)\mathrm{e}^{-\ell}}{\varepsilon + (\ell-1)\mathrm{e}^{-\ell}}.$$

Performing the inverse Laplace transform yields

$$Q_n \approx \int_1^\infty \mathrm{d}\ell\, (\ell-1)\, \exp\left(-\ell - n(\ell-1)\mathrm{e}^{-\ell}\right). \tag{3.9}$$

Setting

$$\lambda = \ln n \tag{3.10}$$

and changing integration variable from ℓ to μ such that $(\ell-1)\mathrm{e}^{-\ell} = \mathrm{e}^{-\mu}$, we obtain formally

$$nQ_n \approx \int_{-\infty}^\infty \mathrm{d}\mu\, \underbrace{\exp\left(\lambda - \mu - \mathrm{e}^{\lambda-\mu}\right)} \left(1 + \frac{1}{\ell(\mu) - 2}\right). \tag{3.11}$$

The expression underlined by the brace is the normalized Gumbel distribution with parameter λ. This distribution is peaked around $\mu = \lambda$. More precisely, considering the following average with respect to this distribution,

$$\int_{-\infty}^\infty \mathrm{d}\mu\, \mathrm{e}^{s\mu} \exp\left(\lambda - \mu - \mathrm{e}^{\lambda-\mu}\right) = \mathrm{e}^{s\lambda}\Gamma(1-s),$$

we obtain

$$\int_{-\infty}^\infty \mathrm{d}\mu\, \phi(\mu) \exp\left(\lambda - \mu - \mathrm{e}^{\lambda-\mu}\right) = \phi(\lambda) + \gamma\phi'(\lambda) + \left(\frac{\gamma^2}{2} + \frac{\pi^2}{12}\right)\phi''(\lambda) + \cdots,$$

for any slowly varying function $\phi(\mu)$, where γ is Euler's constant. Applying this to the function inside the large parentheses in (3.11), we obtain the expansion

$$nQ_n = 1 + \delta_n = 1 + \frac{1}{\lambda} - \frac{\nu - 2}{\lambda^2} + \frac{\nu^2 - 5\nu + 5 + \pi^2/6}{\lambda^3} + \cdots, \tag{3.12}$$

with the notation (3.10) and

$$\nu = \ln\lambda + \gamma = \ln\ln n + \gamma.$$

Omitting details, let us mention that a similar analysis yields the following expansion for the mean number of records up to time n:

$$\langle M_n \rangle = \lambda + \nu - 1 + \frac{\nu - 1}{\lambda} - \frac{\nu^2 - 4\nu + 3 + \pi^2/6}{2\lambda^2} + \cdots \tag{3.13}$$

Equations (3.12) and (3.13) give the first few terms of asymptotic expansions to all orders in $1/\lambda$. The ambiguity in the formal expression (3.11) originating in the pole at $\ell = 2$, i.e., $\mu = 1$, is indeed exponentially small in λ.

3.5. Mean value of the maximum

Using (2.5), (2.16) and (3.5), we obtain the generating function of the mean value $\langle L_n \rangle$ of the largest variable Y_i up to time n,

$$L(z) = \sum_{n\ge 1} z^n \langle L_n \rangle = \int_0^\infty \mathrm{d}\ell \left(\frac{1}{1-z} - \frac{1}{z}\widetilde{F}(z,\ell,0) \right).$$

The explicit expression (3.6) of the generating function $\widetilde{F}(z,\ell,x)$ implies

$$\frac{1}{1-z} - \frac{1}{z}\widetilde{F}(z,\ell,0) = \frac{z\,\mathrm{e}^{-\ell}}{1-z}\left(I(z,\ell)+1\right),$$

so that

$$L(z) = \frac{z+\widetilde{Q}(z)}{1-z},$$

and finally

$$\langle L_n \rangle = \langle M_n \rangle + 1. \tag{3.14}$$

This remarkable identity between mean values is a peculiarity of the exponential distribution. The first few values of $\langle L_n \rangle$ can therefore be read from table 2, whereas its asymptotic growth can be read from (3.13). Let us notice that a similar identity, i.e., $\langle L_n \rangle = \langle M_n \rangle = H_n$, holds for extremes and records of exponentially distributed iid random variables (see [20, 21] for a discussion of related matters).

4. Uniform distribution

The case where the random variables X_i are uniformly distributed on [0, 1], with common density $f(x) = 1$ and distribution function $F(x) = x$ for $0 < x < 1$ (see table 1), also lends itself to an exact solution of the problem.

4.1. Sectors

Here, the relevant part of the (ℓ, x) plane is the rectangle defined by $0 < \ell < 2$ and $0 < x < 1$. This region splits into four sectors (see figure 3):

$$\begin{cases} (1): & 1<\ell<2, \quad 0<x<\ell-1, \\ (2): & 1<\ell<2, \quad \ell-1<x<1, \\ (3): & 0<\ell<1, \quad 0<x<\ell, \\ (4): & 0<\ell<1, \quad \ell<x<1. \end{cases}$$

The functions $F_n(\ell, x)$ assume a priori different analytical forms in these four sectors. The recursion (2.14) reads

$$\begin{aligned} F_n^{(1)}(\ell,x) &= \int_0^{\ell-1} \mathrm{d}x'\,F_{n-1}^{(1)}(\ell,x') + \int_{\ell-1}^{1} \mathrm{d}x'\,F_{n-1}^{(2)}(\ell,x'), \\ F_n^{(2)}(\ell,x) &= \int_0^{\ell-1} \mathrm{d}x'\,F_{n-1}^{(1)}(\ell,x') + \int_{\ell-1}^{\ell-x} \mathrm{d}x'\,F_{n-1}^{(2)}(\ell,x'), \\ F_n^{(3)}(\ell,x) &= \int_0^{\ell-x} \mathrm{d}x'\,F_{n-1}^{(3)}(\ell,x'), \\ F_n^{(4)}(\ell,x) &= 0. \end{aligned} \tag{4.1}$$

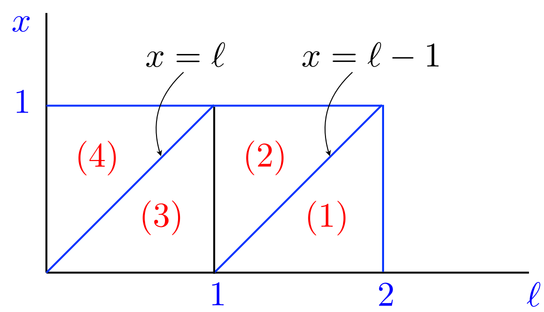

Figure 3. Four sectors in the (ℓ, x) plane for the case of a uniform distribution $f(x)$.

The first function $F_n^{(1)}(\ell, x)$ is independent of x, whereas the last one vanishes, so the information of interest is contained in sectors (2) and (3).

The probability of record breaking reads

$$Q_n = Q_n^{(2)} + Q_n^{(3)},$$

with

$$Q_n^{(2)} = \int_1^2 \mathrm{d}\ell \int_{\ell-1}^1 \mathrm{d}x\, F_{n-1}^{(2)}(\ell, x),$$

$$Q_n^{(3)} = \int_0^1 \mathrm{d}\ell \int_0^\ell \mathrm{d}x\, F_{n-1}^{(3)}(\ell, x).$$

Similarly, the mean value of the maximum reads

$$\langle L_n \rangle = 2 - I_n^{(2)} - I_n^{(3)},$$

with

$$I_n^{(2)} = \int_1^2 \mathrm{d}\ell\, F_{n+1}^{(2)}(\ell, \ell - 1),$$

$$I_n^{(3)} = \int_0^1 \mathrm{d}\ell\, F_{n+1}^{(3)}(\ell, 0).$$

Finally, differentiating (4.1) with respect to x, we obtain the following differential equations, valid in both sectors (2) and (3):

$$F_n'(\ell, x) = -F_{n-1}(\ell, \ell - x), \tag{4.2}$$

$$F_n''(\ell, x) = -F_{n-2}(\ell, x). \tag{4.3}$$

These equations, which can alternatively be read off from (2.15), will be instrumental hereafter. We thus obtain the following expressions for the first values of n:

$$F_1^{(1)}(\ell, x) = F_1^{(2)}(\ell, x) = F_1^{(3)}(\ell, x) = 1$$

for $n = 1$,

$$F_2^{(1)}(\ell, x) = 1,$$

$$F_2^{(2)}(\ell, x) = F_2^{(3)}(\ell, x) = \ell - x$$

for $n = 2$, and

$$F_3^{(1)}(\ell, x) = \frac{1}{2}(-2 + 4\ell - \ell^2),$$

$$F_3^{(2)}(\ell, x) = \frac{1}{2}(2\ell - 1 - x^2),$$

$$F_3^{(3)}(\ell, x) = \frac{1}{2}(\ell^2 - x^2)$$

for $n = 3$.

4.2. Analysis of sector (3)

The generating function

$$\widetilde{F}^{(3)}(z, \ell, x) = \sum_{n \geq 1} z^n F_n^{(3)}(\ell, x)$$

satisfies

$$\widetilde{F}'^{(3)}(z, \ell, x) = -z\widetilde{F}^{(3)}(z, \ell, \ell - x),$$

because of (4.2),

$$\widetilde{F}''^{(3)}(z, \ell, x) = -z^2\widetilde{F}^{(3)}(z, \ell, x),$$

because of (4.3), and

$$\widetilde{F}^{(3)}(z, \ell, \ell) = z, \quad \widetilde{F}'^{(3)}(z, \ell, 0) = -z^2, \tag{4.4}$$

because of (2.13). Hence

$$\widetilde{F}^{(3)}(z, \ell, x) = A \cos zx + B \sin zx,$$

where the amplitudes A and B, which depend a priori on z and ℓ, are determined by the boundary conditions (4.4). We thus obtain

$$\widetilde{F}^{(3)}(z, \ell, x) = z\,\frac{\cos z(\ell - x) - \sin zx}{1 - \sin z\ell}. \tag{4.5}$$

Hence

$$\begin{aligned}\widetilde{Q}^{(3)}(z) = \sum_{n \geq 2} z^n Q_n^{(3)} &= z \int_0^1 \mathrm{d}\ell \int_0^\ell \mathrm{d}x\, \widetilde{F}^{(3)}(z, \ell, x) \\ &= -z - \ln(1 - \sin z),\end{aligned} \tag{4.6}$$

and

$$\begin{aligned}\widetilde{I}^{(3)}(z) = \sum_{n \geq 1} z^n I_n^{(3)} &= -1 + \frac{1}{z} \int_0^1 \mathrm{d}\ell\, \widetilde{F}^{(3)}(z, \ell, 0) \\ &= -1 - \frac{1}{z} \ln(1 - \sin z) = \frac{\widetilde{Q}^{(3)}(z)}{z}.\end{aligned} \tag{4.7}$$

Relationship with Euler numbers. Consider n positive numbers $x_1, \ldots, x_n$ such that $x_i + x_{i+1} \leq 1$ for $1 \leq i \leq n - 1$. These conditions define a volume V_n for every integer n. The generating function of these numbers reads [22]

$$\begin{aligned}\widetilde{V}(z) = \sum_{n \geq 0} z^n V_n &= \frac{1}{\cos z} + \tan z \\ &= 1 + z + \frac{z^2}{2} + \frac{z^3}{3} + \frac{5}{24}z^4 + \cdots\end{aligned}$$

We have

$$V_n = \frac{E_n}{n!},$$

where $E_n = 1, 1, 1, 2, 5, 16, 61, \ldots$ are the Euler numbers, listed as sequence number A000111 in the On-Line Encyclopedia of Integer Sequences [23]. The volumes V_n are also simply related to the $Q^{(3)}(n)$, as we now show. Let us note (see (4.6)) that

$$\frac{\mathrm{d}\widetilde{Q}^{(3)}(z)}{\mathrm{d}z} = \widetilde{V}(z) - 1,$$

hence, for $n \geq 2$,

$$Q_n^{(3)} = \frac{V_{n-1}}{n} = \frac{E_{n-1}}{n!}.$$

The relationship between the two sequences $Q_n^{(3)}$ and V_n comes from the fact that

$$\begin{aligned}\int_0^\ell \mathrm{d}x' F_{n-1}^{(3)}(\ell, x') &= \mathrm{Prob}(L_{n-1} < \ell) \\ &= \mathrm{Prob}(Y_1 < \ell, \ldots, Y_{n-1} < \ell) = V_{n-1}\,\ell^{n-1},\end{aligned}$$

hence, integrating over ℓ,

$$Q_n^{(3)} = \int_0^1 \mathrm{d}\ell\, V_{n-1}\,\ell^{n-1} = \frac{V_{n-1}}{n}.$$

Let us remark that

$$\widetilde{F}^{(3)}(z, \ell, x) = z\left(\widetilde{V}(z\ell)\cos zx - \sin zx\right).$$

Finally, the generating function $\widetilde{V}(z)$ has a pole at $z = \pi/2$, with residue 2, and therefore

$$V_n \approx 2\left(\frac{2}{\pi}\right)^{n+1}, \quad Q_n^{(3)} \approx \frac{2}{n}\left(\frac{2}{\pi}\right)^n, \quad I_n^{(3)} \approx \frac{4}{\pi n}\left(\frac{2}{\pi}\right)^n.$$

4.3. Analysis of sector (2)

The generating function

$$\widetilde{F}^{(2)}(z, \ell, x) = \sum_{n \geq 1} z^n F_n^{(2)}(\ell, x),$$

satisfies

$$\widetilde{F}'^{(2)}(z, \ell, x) = -z\widetilde{F}^{(2)}(z, \ell, \ell - x),$$

because of (4.2),

$$\widetilde{F}''^{(2)}(z, \ell, x) = -z^2\widetilde{F}^{(2)}(z, \ell, x),$$

because of (4.3), and

$$\widetilde{F}^{(2)}(z, \ell, 1) = z\left(1 + (\ell - 1)\widetilde{F}^{(2)}(z, \ell, \ell - 1)\right),$$

as a consequence of (4.1), using the fact that $F_n^{(1)}(\ell, x)$ is independent of x.

We thus obtain, in analogy with (4.5)

$$\widetilde{F}^{(2)}(z, \ell, x) = z\,\frac{\cos z(\ell - x) - \sin zx}{\Delta(z, \ell)}, \tag{4.8}$$

with

$$\Delta(z,\ell) = z(\ell-1)(\sin z(\ell-1) - \cos z) + \cos z(\ell-1) - \sin z.$$

We have therefore

$$\begin{aligned}\widetilde{Q}^{(2)}(z) &= \sum_{n\ge 2} z^n Q_n^{(2)} = z\int_1^2 \mathrm{d}\ell \int_{\ell-1}^1 \mathrm{d}x\, \widetilde{F}^{(2)}(z,\ell,x) \\ &= z\int_1^2 \mathrm{d}\ell\, \frac{\cos z + \sin z - \cos z(\ell-1) - \sin z(\ell-1)}{\Delta(z,\ell)},\end{aligned} \tag{4.9}$$

and

$$\begin{aligned}\widetilde{I}^{(2)}(z) &= \sum_{n\ge 1} z^n I_n^{(2)} = -1 + \frac{1}{z}\int_1^2 \mathrm{d}\ell\, \widetilde{F}^{(2)}(z,\ell,\ell-1) \\ &= -1 + \frac{1}{z}\int_1^2 \mathrm{d}\ell\, \frac{\cos z - \sin z(\ell-1)}{\Delta(z,\ell)}.\end{aligned} \tag{4.10}$$

At variance with (4.6) and (4.7), the integrals over ℓ in (4.9) and (4.10) cannot be carried out in closed form.

4.4. Results

By expanding the integrands of (4.9) and (4.10) as power series in z, integrating over ℓ term by term, and adding up the contributions of (4.6) and (4.7), we obtain exact rational expressions for the probability of record breaking Q_n, the mean number of records $\langle M_n\rangle$ and the mean value of the maximum $\langle L_n\rangle$. These outcomes are given in table 3 up to $n = 8$.

n	2	3	4	5	6	7	8
Q_n	1	$\frac{1}{2}$	$\frac{3}{8}$	$\frac{17}{60}$	$\frac{11}{48}$	$\frac{481}{2520}$	$\frac{439}{2688}$
$\langle M_n\rangle$	1	$\frac{3}{2}$	$\frac{15}{8}$	$\frac{259}{120}$	$\frac{191}{180}$	$\frac{2599}{1008}$	$\frac{22109}{8064}$
$\langle L_n\rangle$	1	$\frac{7}{6}$	$\frac{77}{60}$	$\frac{49}{36}$	$\frac{511}{360}$	$\frac{3691}{2520}$	$\frac{272369}{181440}$

Table 3. Exact values of the probability of record breaking Q_n, mean number of records $\langle M_n\rangle$ and mean maximum $\langle L_n\rangle$ up to $n = 8$, for a uniform distribution.

The asymptotic behavior at large n of the various quantities of interest can be derived as follows. First of all, the contribution of sector (3) is exponentially small, and therefore entirely negligible. Setting again $z = \mathrm{e}^{-\varepsilon}$, the integrals entering (4.9) and (4.10) are dominated by a range of values of the difference $2-\ell$ that shrinks proportionally to $\sqrt{\varepsilon}$ as $\varepsilon \to 0$. Changing integration variable from ℓ to t such that $\ell = 2 - t\sqrt{\varepsilon}$, and keeping only terms which are singular in ε, we obtain

$$\begin{aligned}\widetilde{Q}^{(2)}(z) &= \ln\frac{1}{\varepsilon}\left(1 - \frac{\varepsilon}{3} + \cdots\right) - \pi\sqrt{\frac{\varepsilon}{2}}\left(1 - \frac{5\varepsilon}{9} + \cdots\right), \\ \widetilde{I}^{(2)}(z) &= \frac{2}{3}\ln\frac{1}{\varepsilon}\left(1 + \frac{38\varepsilon}{45} + \cdots\right) + \frac{\pi}{\sqrt{2\varepsilon}}\left(1 + \frac{5\varepsilon}{6} + \cdots\right),\end{aligned}$$

and so

$$
\begin{aligned}
nQ_n &= 1 + \delta_n \\
&= 1 + \frac{1}{3n} + \cdots + \sqrt{\frac{\pi}{8n}} \left(1 + \frac{5}{6n} + \cdots\right),
\end{aligned} \tag{4.11}
$$

and

$$
\langle L_n \rangle = 2\left(1 - \frac{1}{3n} + \cdots\right) - \sqrt{\frac{\pi}{2n}} \left(1 - \frac{5}{12n} + \cdots\right).
$$

Finally, omitting details, we obtain a similar asymptotic expansion for the mean number of records, i.e.,

$$
\langle M_n \rangle = \ln n + K + \frac{1}{6n} + \cdots - \sqrt{\frac{\pi}{2n}} \left(1 + \frac{1}{36n} + \cdots\right),
$$

where the finite part reads

$$
\begin{aligned}
K &= \gamma - 1 - \ln(2(1 - \sin 1)) + 2\int_1^2 \mathrm{d}\ell \left(\frac{1}{(2-\ell)\cot\frac{2-\ell}{2} - \ell} - \frac{1}{2-\ell}\right) \\
&= 1.092998\ldots
\end{aligned}
$$

5. Power-law distribution with index $\theta = 1$

The case where the random variables X_i have a power-law distribution with index $\theta = 1$, with common density $f(x) = 1/x^2$ and distribution function $F(x) = 1 - 1/x$ for $x > 1$ (see table 1), is our last example giving rise to an exact solution of the problem, although end results are somewhat less explicit than in the two previous cases. The distribution under consideration is marginal, in the sense that $\langle X \rangle$ is logarithmically divergent.

5.1. Differential equations

In the present case, the key equation (2.15) reads

$$
F_n'(\ell, x) = -\frac{2}{x} F_n(\ell, x) - \frac{1}{x^2} F_{n-1}(\ell, \ell - x)
$$

for $n \geq 2$, $\ell > 2$, and $1 < x < \ell - 1$. Setting

$$
F_n(\ell, x) = \frac{H_n(\ell, x)}{x^2(\ell - x)}, \tag{5.1}
$$

the new functions $H_n(\ell, x)$ obey the differential equation

$$
x(\ell - x)H_n'(\ell, x) + xH_n(\ell, x) = -H_{n-1}(\ell, \ell - x), \tag{5.2}
$$

with boundary condition $H_n(\ell, \ell - 1) = 0$, as well as

$$
x^2(\ell - x)^2 H_n''(\ell, x) = -H_{n-2}(\ell, x). \tag{5.3}
$$

We thus obtain

$$
\begin{aligned}
H_1(\ell, x) &= \ell - x, \\
H_2(\ell, x) &= \ell - x - 1, \\
H_3(\ell, x) &= \frac{(\ell - 1)(\ell - 1 - x)}{\ell} + \frac{\ell - x}{\ell^2} \ln \frac{x}{(\ell - 1)(\ell - x)}.
\end{aligned}
$$

5.2. Generating function

In order to solve the recursive differential equations (5.2), (5.3), we introduce the generating function

$$\widetilde{H}(z,\ell,x) = \sum_{n\ge 1} z^n H_n(\ell,x), \tag{5.4}$$

which obeys

$$x(\ell-x)\widetilde{H}'(\ell,x) + x\widetilde{H}(\ell,x) = -z\widetilde{H}(\ell,\ell-x), \tag{5.5}$$

with boundary condition

$$\widetilde{H}(\ell,\ell-1) = z, \tag{5.6}$$

as well as

$$x^2(\ell-x)^2\widetilde{H}''(\ell,x) = -z^2\widetilde{H}(\ell,x). \tag{5.7}$$

The general solution to (5.7) reads

$$\widetilde{H}(z,\ell,x) = A_+ x^{a_+}(\ell-x)^{a_-} + A_- x^{a_-}(\ell-x)^{a_+},$$

with

$$a_\pm = \frac{1\pm w}{2}, \quad w = \sqrt{1-\frac{4z^2}{\ell^2}}.$$

Notice the similarity with (3.6). The amplitudes $A_\pm$ are determined by (5.5) and (5.6), yielding

$$\widetilde{H}(z,\ell,x) = z\,\frac{\sqrt{a_-}\,x^{a_+}(\ell-x)^{a_-} - \sqrt{a_+}\,x^{a_-}(\ell-x)^{a_+}}{\sqrt{a_-}\,(\ell-1)^{a_+} - \sqrt{a_+}\,(\ell-1)^{a_-}}. \tag{5.8}$$

This result demonstrates that the functions $H_n(\ell,x)$ only involve integer powers of $\ln x$ and $\ln(\ell-x)$, besides rational functions.

5.3. Probability of record breaking

The generating function of the Q_n reads

$$\widetilde{Q}(z) = \sum_{n\ge 2} z^n Q_n = z\int_2^\infty \mathrm{d}\ell \int_1^{\ell-1} \frac{\mathrm{d}x}{x^2(\ell-x)^3}\,\widetilde{H}(z,\ell,x),$$

by virtue of (2.21), (5.1) and (5.4), where $\widetilde{H}(z,\ell,x)$ is given by (5.8). The integral over x can be carried out in closed form. We thus obtain

$$\widetilde{Q}(z) = \int_2^\infty \mathrm{d}\ell\, I(z,\ell), \tag{5.9}$$

with

$$I(z,\ell) = z^2\,\frac{\sqrt{a_-}\,I_+(z,\ell) - \sqrt{a_+}\,I_-(z,\ell)}{\sqrt{a_-}\,(\ell-1)^{a_+} - \sqrt{a_+}\,(\ell-1)^{a_-}}$$

and

$$\begin{aligned}
I_\pm(z,\ell) &= \int_1^{\ell-1} \mathrm{d}x\, \frac{x^{a_\pm}(\ell-x)^{a_\mp}}{x^2(\ell-x)^3} \\
&= \int_1^{\ell-1} \mathrm{d}x\, x^{-2+a_\pm}(\ell-x)^{-2-a_\pm} \\
&= \frac{1}{(1-a_\pm)\ell^3}\left((\ell-1)^{1-a_\pm} - (\ell-1)^{-(1-a_\pm)}\right) \\
&+ \frac{1}{(1+a_\pm)\ell^3}\left((\ell-1)^{1+a_\pm} - (\ell-1)^{-(1+a_\pm)}\right) \\
&+ \frac{2}{a_\pm\ell^3}\left((\ell-1)^{a_\pm} - (\ell-1)^{-a_\pm}\right).
\end{aligned}$$

As was the case for (3.7) and (4.9), the integrals over ℓ in (5.9) cannot be carried out analytically in closed form. By expanding the integrand in (5.9) as a power series in z and integrating term by term with respect to ℓ, we obtain the following values for the probability of record breaking Q_n, besides the universal ones derived in section 2.6:

$$\begin{aligned}
Q_5 &= \frac{5}{8} - \frac{\pi^2}{30} = 0.296013\ldots, \\
Q_6 &= \frac{61}{144} + \frac{14\pi^2}{135} - \zeta(3) = 0.245068\ldots, \\
Q_7 &= \frac{475}{252} - \frac{292\pi^2}{945} + \frac{8\zeta(3)}{7} = 0.209044\ldots,
\end{aligned}$$

and so on. In contrast with the two previous exactly solvable cases (see tables 2 and 3), here the non-universal Q_n are not rational, and they involve the values of Riemann zeta function at larger and larger positive integers.

The asymptotic behavior of Q_n at large n can be derived from (5.9) by setting again $z = \mathrm{e}^{-\varepsilon}$, and considering the regime where ε and $1/\ell$ are simultaneously small. To leading order, (5.9) reduces to

$$\widetilde{Q}(z) \approx \int_2^\infty \mathrm{d}\ell\, \frac{3}{2\ell(1+\varepsilon\ell)},$$

i.e., performing the inverse Laplace transform,

$$Q_n \approx \int_2^\infty \mathrm{d}\ell\, \frac{3}{2\ell^2}\,\mathrm{e}^{-n/\ell} \approx \frac{3}{2n},$$

up to negligible boundary terms. The above result is an explicit instance where (1.11) holds. A full asymptotic expansion of Q_n can be derived by keeping track of higher orders, yielding

$$nQ_n = 1 + \delta_n = \frac{3}{2} - \frac{2(\ln n + \gamma - 3)}{n} + \cdots \tag{5.10}$$

Finally, omitting details, we obtain a similar asymptotic expansion for the mean number of records, i.e.,

$$\langle M_n \rangle = \frac{3}{2}\,(\ln n + K) + \frac{2(\ln n + \gamma - 2)}{n} + \cdots,$$

where the finite part reads

$$K = \gamma - \ln 2 + \int_2^\infty \mathrm{d}\ell\left(\frac{2I(1,\ell)}{3} - \frac{1}{\ell}\right) = -0.387293\ldots$$

5.4. Distribution of the maximum

Here $\langle X\rangle$ is divergent, so that it makes no sense to evaluate $\langle L_n\rangle$. The full distribution of L_n should be considered instead. We have

$$\mathcal{F}_n(\ell) = \int_1^{\ell-1} \mathrm{d}x\, F_n(\ell, x) = F_{n+1}(\ell, 1) = \frac{H_{n+1}(\ell, 1)}{\ell - 1},$$

as a consequence of (2.4), (2.14) and (5.1). We thus obtain

$$\begin{aligned}
\mathcal{F}_1(\ell) &= 1 - \frac{1}{\ell-1}, \\
\mathcal{F}_2(\ell) &= 1 - \frac{2}{\ell} - \frac{2}{\ell^2}\ln(\ell-1), \\
\mathcal{F}_3(\ell) &= 1 - \frac{3(\ell-1)}{\ell^2} - \frac{1}{\ell^2(\ell-1)} - \frac{4(\ell-1)}{\ell^3}\ln(\ell-1).
\end{aligned}$$

The corresponding generating function reads

$$\begin{aligned}
\widetilde{\mathcal{F}}(z, \ell) &= \sum_{n\ge 0} z^n \mathcal{F}_n(\ell) = \frac{\widetilde{H}(z, \ell, 1)}{z(\ell-1)} \\
&= \frac{\sqrt{a_-}\,(\ell-1)^{-a_+} - \sqrt{a_+}\,(\ell-1)^{-a_-}}{\sqrt{a_-}\,(\ell-1)^{a_+} - \sqrt{a_+}\,(\ell-1)^{a_-}},
\end{aligned}$$

where the last expression is a consequence of (5.8).

The scaling behavior of the distribution of L_n at large n can be derived along the lines of the previous section. To leading order, we find the simple result

$$\mathcal{F}_n(\ell) \approx \mathrm{e}^{-n/\ell}. \tag{5.11}$$

In particular, the median value $L_n^\star$, such that $\mathcal{F}_n(L_n^\star) = \frac{1}{2}$, reads

$$L_n^\star \approx \frac{n}{\ln 2}.$$

The full asymptotic expansion of $\mathcal{F}_n(\ell)$ in the regime where n and ℓ are comparable reads

$$\mathcal{F}_n(\ell) = \mathrm{e}^{-n/\ell}\left(1 - \frac{n}{2\ell^2}(4\ln\ell - 1) + \cdots\right), \tag{5.12}$$

and so

$$L_n^\star = \frac{n}{\ln 2} + 2\ln\frac{n}{\ln 2} - \frac{1}{2} + \cdots$$

6. Asymptotic analysis of the general case

The probability of record breaking Q_n exhibits a great variety of asymptotic behaviors, depending on the underlying distribution $f(x)$. This is exemplified by the three exactly solvable cases studied in sections 3 to 5. In terms of the correction δ_n such that $nQ_n = 1 + \delta_n$ (see (1.15)), we have seen that

$$\delta_n \approx \frac{1}{\ln n} \tag{6.1}$$

for the exponential distribution (see (3.12)),

$$\delta_n \approx \sqrt{\frac{\pi}{8n}}$$

for the uniform distribution (see (4.11)), and

$$\delta_n \to \frac{1}{2}, \tag{6.2}$$

which is equivalent to (1.11), for the power-law distribution with $\theta = 1$ (see (5.10)).

This section is devoted to a heuristic but systematic analysis of the dependence of the asymptotic behavior of δ_n on the underlying distribution $f(x)$. It will turn out that the exponential distribution, where δ_n falls off logarithmically (see (6.1)), is a marginal case. For superexponential distributions, the analysis of sections 6.3 and 6.4 demonstrates that δ_n falls off to zero and yields a general asymptotic formula for δ_n (see (6.10)). For subexponential distributions, it will be shown in section 6.5 that nQ_n and δ_n go to the universal limits (1.11) and (6.2). This dichotomy will be extended to higher values of the window width p in section 7.

6.1. Cyclization of the sequence

The first step of the analysis consists in comparing the problem at hand with a cyclic variant of it. For the former, we have

$$Q_n = \mathrm{Prob}(Y_n > L_{n-1}),$$

with

$$L_n = \max(Y_2, \ldots, Y_n)$$

(see (1.7) and (1.8)). The cyclic variant of the problem is defined by introducing

$$Y_1^{\mathrm{cyclic}} = X_n + X_1.$$

The sequence $Y_1^{\mathrm{cyclic}}, Y_2, \ldots, Y_n$ thus obtained involves the basic variables $X_1, \ldots, X_n$ in a cyclically invariant fashion. It has therefore exchangeable entries, and so

$$Q_n^{\mathrm{cyclic}} = \mathrm{Prob}\left(Y_n > \max(Y_1^{\mathrm{cyclic}}, Y_2, \ldots, Y_{n-1})\right) = \frac{1}{n}.$$

Introducing the events

$$E = \{Y_n > Y_1^{\mathrm{cyclic}}\} = \{X_{n-1} > X_1\},$$
$$F = \{Y_n > L_{n-1}\} = \{L_n = Y_n\},$$

we have

$$\mathrm{Prob}(E \cap F) = Q_n^{\mathrm{cyclic}} = \frac{1}{n},$$
$$\mathrm{Prob}(F) = Q_n = \frac{1 + \delta_n}{n},$$

and so

$$\begin{aligned} \Delta_n &= \mathrm{Prob}(\bar{E}|F) = \mathrm{Prob}(X_1 > X_{n-1}|L_n = Y_n) \\ &= \frac{Q_n - Q_n^{\mathrm{cyclic}}}{Q_n} = \frac{\delta_n}{1 + \delta_n}. \end{aligned} \tag{6.3}$$

This equation gives a description of the difference $Q_n - Q_n^{\mathrm{cyclic}}$ in terms of a conditional probability, which will prove useful in the following.

6.2. Decoupled model

We now consider a decoupled variant of the original problem, whose main advantage is that the expression (6.3) can be given the explicit form (6.9), which will in turn yield the estimate (6.10) for the correction δ_n in appropriate situations.

The decoupled model is defined as follows. The random variables Y_i of the original problem are replaced by a sequence of iid random variables

$$\mathcal{Y}_i = X_i + X_i', \tag{6.4}$$

where X_i and X_i' are two independent replicas of the original random variables X_i with common density $f(x)$ and distribution function $F(x)$. The number of variables X is therefore doubled with respect to the original problem. The distribution function $F_2(y)$ and the density $f_2(y)$ of the variables $\mathcal{Y}_i$ thus read

$$\begin{aligned} F_2(y) &= \mathrm{Prob}(\mathcal{Y} < y) = \mathrm{Prob}(X + X' < y) = \int_0^y \mathrm{d}y' f_2(y'), \\ f_2(y) &= \int_0^y \mathrm{d}x\, f(x) f(y-x). \end{aligned} \tag{6.5}$$

In terms of the Laplace transform

$$\hat{f}(s) = \int_0^\infty \mathrm{d}x\, \mathrm{e}^{-sx}\, f(x),$$

this reads

$$\hat{f}_2(s) = \hat{f}(s)^2. \tag{6.6}$$

The conditional density of X given $X + X' = y$, denoted by $f(x|y)$, is equal to

$$f(x|y) = \frac{f(x)f(y-x)}{f_2(y)}. \tag{6.7}$$

The largest among the first n variables $\mathcal{Y}_i$, denoted by

$$\mathcal{Y}^* = X^* + X'^*,$$

has distribution function

$$F_{\mathcal{Y}^*}(y) = \mathrm{Prob}(\mathcal{Y}^* < y) = F_2(y)^n,$$

and density

$$f_{\mathcal{Y}^*}(y) = nF_2(y)^{n-1} f_2(y). \tag{6.8}$$

Using (6.7) and (6.8), the density of X^* is

$$f_{X^*}(x) = \int_0^\infty \mathrm{d}y\, f(x|y) f_{\mathcal{Y}^*}(y) = nf(x) \int_x^\infty \mathrm{d}y\, f(y-x) F_2(y)^{n-1}.$$

Within the setting of the decoupled model, the conditional probability Δ_n introduced in (6.3) therefore reads

$$\begin{aligned} \Delta_n &= \mathrm{Prob}(X > X^*) = \int_0^\infty \mathrm{d}x\, f_{X^*}(x) \bar{F}(x) \\ &= n \int_0^\infty \mathrm{d}y\, F_2(y)^{n-1} \int_0^y \mathrm{d}x\, \underbrace{f(x)f(y-x)}\, \bar{F}(x), \end{aligned} \tag{6.9}$$

with

$$\bar{F}(x) = \mathrm{Prob}(X > x) = 1 - F(x).$$

When n is large, the factor $F_2(y)^{n-1}$ in (6.9) selects large values of y, such that $\bar{F}_2(y)$ scales as $1/n$. These are the typical values of $\mathcal{Y}^*$. The product underlined by the brace, which already entered (6.5) and (6.7), describes to what extent the distribution of X is affected by the conditioning by such a large value y of the sum $\mathcal{Y} = X + X'$.

6.3. The key dichotomy

The dichotomy between the two limits (1.10) and (1.11) is now shown in general albeit non-rigorous terms to be dictated by the form of the tail of the underlying parent distribution $f(x)$ or, equivalently, by the analytic structure of its Laplace transform $\hat{f}(s)$.

- For *superexponential distributions*, i.e., distributions $f(x)$ either having a bounded support or falling off faster than any exponential, such as e.g. a half-Gaussian or any other compressed exponential, $\hat{f}(s)$ is an entire function, i.e., it is analytic in the whole complex s-plane. Then, as a general rule, $f_2(y)$ (see (6.5)) has a slower decay than $f(x)$. Furthermore, if the sum $\mathcal{Y} = X + X'$ is atypically large, then both X and X' are atypically large as well, with very high probability. As a consequence, the conditional probability Δ_n, as given by (6.9), falls off to zero for large n. Simplifying the latter expression, we thus obtain the following asymptotic estimate for δ_n:

$$\delta_n \approx n \int_0^\infty \mathrm{d}y\, \mathrm{e}^{-n\bar{F}_2(y)} \int_0^y \mathrm{d}x\, f(x) f(y-x) \bar{F}(x). \tag{6.10}$$

  We claim that this prediction becomes asymptotically exact for all superexponential distributions, in the sense that it correctly describes the decay of δ_n, to leading order for large n, in spite of its heuristic derivation using the decoupled model. The rationale behind this claim is that the difference between the original and the decoupled models, measured by the relative difference between Q_n and Q_n^{cyclic}, is consistently found to decay to zero, proportionally to the estimate (6.10) for δ_n.
- For *subexponential distributions*, i.e., distributions $f(x)$ which fall off smoothly enough and less rapidly than any exponential, such as e.g. a power law or a stretched exponential, $\hat{f}(s)$ has an isolated branch-point singularity at $s = 0$. The asymptotic equivalence of the tails,

$$\bar{F}_2(y) \approx 2\bar{F}(y) \quad (y \gg 1), \tag{6.11}$$

  can be derived by an inverse Laplace transform of (6.6), where the contour integral is dominated by the singularity of $\hat{f}(s)$ at $s = 0$. Equation (6.11) may be used as a mathematically rigorous definition of the class of subexponential distributions, following Chistyakov [24]. Its intuitive meaning is the following: if the sum $\mathcal{Y} = X + X'$ is very large, then one of the terms, either X or X'—hence the factor 2—is typical, i.e., distributed according to $f(x)$, while the other one is essentially equal to $\mathcal{Y}$. This behavior underlies the phenomenon of condensation for subexponential random variables conditioned by an atypical value of their sum (see [25] for a recent review and the references therein). As a consequence of (6.11), for subexponential distributions $f(x)$, the estimate (6.9) remains of order unity for large n. The decoupled model is therefore of little use to understand the original one. This situation will be investigated in section 6.5, where nQ_n and δ_n will be shown to admit the universal limits (1.11) and (6.2).

For *exponential distributions*, i.e., distributions $f(x)$ falling off either as a pure exponential $\mathrm{e}^{-\beta x}$, with $\beta > 0$, or as the product of such an exponential by a more slowly varying prefactor, such as e.g. a power of x, the leading (i.e., rightmost) singularity of $\hat{f}(s)$ is located on the negative real axis at $s = -\beta$. For our purpose, these distributions are marginal since they can lie on either sides of the dichotomy between (1.10) and (1.11) (see section 6.4.2).

6.4. Superexponential and (some) exponential distributions

The prediction (6.10) is now made explicit for a variety of superexponential and exponential distributions $f(x)$.

6.4.1. Pure exponential distribution. This is the distribution for which an exact solution has been presented in section 3. We have

$$f(x) = \bar{F}(x) = \mathrm{e}^{-x}, \quad f_2(y) = y\,\mathrm{e}^{-y}, \quad \bar{F}_2(y) = (y+1)\mathrm{e}^{-y}.$$

The estimate (6.10) therefore reads

$$\delta_n \approx n \int_0^\infty \mathrm{d}y\, \exp\left(-y - n(y+1)\mathrm{e}^{-y}\right). \tag{6.12}$$

This integral can be evaluated in analogy with (3.9). Setting $\lambda = \ln n$ (see (3.10)) and $(y+1)\mathrm{e}^{-y} = \mathrm{e}^{-\mu}$, we obtain

$$\delta_n \approx \int_{-\infty}^\infty \mathrm{d}\mu\, \exp\left(\lambda - \mu - \mathrm{e}^{\lambda-\mu}\right) \frac{1}{y(\mu)},$$

hence

$$\delta_n \approx \frac{1}{\lambda} - \frac{\ln\lambda + \gamma}{\lambda^2} + \cdots \tag{6.13}$$

A comparison with the exact expansion (3.12) shows that the estimate (6.10) is correct to leading order in this marginal case. The difference between the estimate (6.13) and the exact result is indeed subleading, since it scales as $2/\lambda^2$.

6.4.2. Exponential distribution modulated by a power law. We now consider distributions falling off as an exponential modulated by a power law, i.e.,

$$f(x) \approx \bar{F}(x) \approx A\, x^{a-1}\, \mathrm{e}^{-x} \quad (x \to \infty), \tag{6.14}$$

where a is arbitrary (positive or negative).

Let us consider first the case where $a > 0$. We have then

$$\hat{f}(s) \approx \frac{A\Gamma(a)}{(s+1)^a} \quad (s \to -1)$$

and

$$f_2(y) \approx \bar{F}_2(y) \approx B\, y^{2a-1}\, \mathrm{e}^{-y} \quad (y \to \infty), \tag{6.15}$$

with $B = (A\Gamma(a))^2/\Gamma(2a)$. Performing the integrals entering (6.10), we obtain

$$\delta_n \approx nA^3\Gamma(a) \int_0^\infty \mathrm{d}y\, \exp\left(-nBy^{2a-1}\mathrm{e}^{-y}\right) y^{a-1}\mathrm{e}^{-y}.$$

This integral can be evaluated in analogy with (3.9). Setting $\lambda = \ln(nB)$ and $y^{2a-1}\mathrm{e}^{-y} = \mathrm{e}^{-\mu}$, we obtain formally

$$\delta_n \approx \frac{A^3\Gamma(a)}{B} \int_{-\infty}^\infty \mathrm{d}\mu\, \exp\left(\lambda - \mu - \mathrm{e}^{\lambda-\mu}\right) \frac{1}{y(\mu)^a}.$$

To leading order, the identification $y(\mu) \approx \mu \approx \lambda$ yields the estimate

$$\delta_n \approx \frac{A\Gamma(2a)}{\Gamma(a)} \frac{1}{(\ln n)^a}. \tag{6.16}$$

We are thus led to claim that exponential distributions of the form (6.14) with $a > 0$, and presumably all exponential distributions such that $\hat{f}(s) \to +\infty$ as the leading singularity is approached from the right ($s \to -\beta^+$), belong to the superexponential side of the dichotomy, in the sense that (1.10) holds, and that (6.16) correctly predicts the decay of the correction δ_n. The logarithmically slow fall off of the latter expression confirms the marginal character of this class of exponential distributions.

On the contrary, if the exponent a entering (6.14) is negative, the above derivation already breaks down at the level of (6.15). Exponential distributions of the form (6.14) with $a < 0$, and presumably all exponential distributions such that $\hat{f}(s)$ remains bounded as $s \to -\beta^+$, therefore share with subexponential distributions the property that the estimate δ_n does not decay to zero, with the expected consequence that (1.11) should hold.

6.4.3. Distributions with bounded support and power-law singularity. We now consider the case where $f(x)$ is supported by the interval [0, 1] and has a power-law singularity at its upper edge, i.e.,

$$\begin{aligned} &\bar{F}(x) \approx A\varepsilon^a, \quad f(x) \approx aA\varepsilon^{a-1}, \\ &\bar{F}_2(y) \approx B\eta^{2a}, \quad f_2(y) \approx 2aB\eta^{2a-1}, \end{aligned} \tag{6.17}$$

with the notations $\varepsilon = 1 - x$, $\eta = 2 - y$. The exponent $a > 0$ and the amplitude $A > 0$ are arbitrary. We have $B = a(A\Gamma(a))^2/(2\Gamma(2a))$. Performing the integrals entering (6.10), we obtain a universal $1/\sqrt{n}$ decay for δ_n, irrespective of the exponent a, i.e.,

$$\delta_n \approx \frac{K(a)}{\sqrt{n}}, \tag{6.18}$$

where the amplitude $K(a)$ reads

$$K(a) = \frac{1}{\Gamma(a)^2\,\Gamma(3a)}\sqrt{\frac{\pi\Gamma(2a)^5}{2a}}. \tag{6.19}$$

The amplitude $K(a)$ is shown in figure 4. It has a local maximum at $K(0) = 3\sqrt{\pi}/8 = 0.664670\ldots$ and a local minimum at $K(1) = \sqrt{\pi/8} = 0.626657\ldots$ The latter value agrees with the exact result (4.11) for the uniform distribution. This provides another corroboration of our claim that the estimate (6.10) is correct to leading order. The exponential growth $K(a) \sim (32/27)^a$ of the amplitude at large a suggests that the $1/\sqrt{n}$ decay ceases to hold for distributions with an infinitely large exponent, i.e., with an essential singularity at their upper edge.

6.4.4. Distributions with bounded support and exponential singularity. We now consider the case where $f(x)$ is supported by the interval [0, 1] and has an exponentially small singularity at its upper edge, of the form

$$f(x) \sim \bar{F}(x) \sim \mathrm{e}^{-C/\varepsilon^b}, \tag{6.20}$$

with $b > 0$. Using the same notations as above, and working within exponential accuracy, we have

$$f_2 \sim \int_0^\eta \mathrm{d}\varepsilon\, \mathrm{e}^{-C(1/\varepsilon^b + 1/(\eta-\varepsilon)^b)},$$

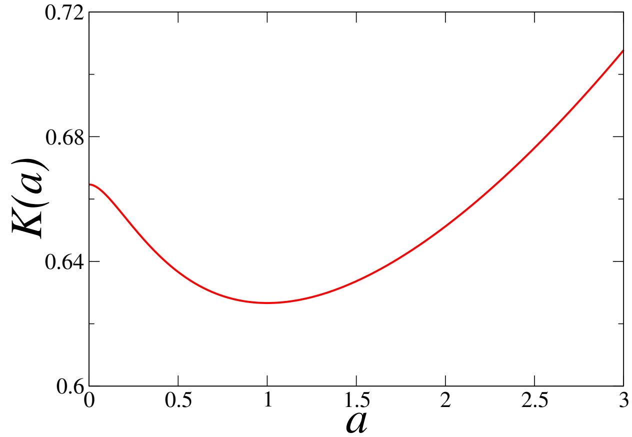

Figure 4. Amplitude $K(a)$ of the universal $1/\sqrt{n}$ decay (6.18) of the correction term δ_n, against the exponent a.

for small η, where the integral is dominated by a saddle point at $\varepsilon = \eta/2$, so that

$$f_2 \sim \bar{F}_2 \sim \mathrm{e}^{-2^{1+b}C/\eta^b}.$$

Similarly, the x-integral entering (6.10) is dominated by a saddle point at $\varepsilon = \tau\eta$, with $\tau = 1/(1 + 2^{-1/(b+1)})$, and so

$$\delta_n \sim \int_0^\infty \mathrm{d}\eta \, \exp\left(-\frac{2C}{\tau^{b+1}\eta^b} - n\,\mathrm{e}^{-2^{1+b}C/\eta^b}\right).$$

Using once more the saddle-point method, we obtain a power-law decay of the form

$$\delta_n \sim n^{-\omega_1(b)},$$

where the exponent

$$\omega_1(b) = \frac{(1 + 2^{-1/(b+1)})^{b+1}}{2^b} - 1 \qquad (6.21)$$

decreases monotonically as a function of b, from $\omega_1(0) = 1/2$ to $\omega_1(\infty) = \sqrt{2} - 1$.

6.4.5. Compressed exponential distributions. We now consider the case where $f(x)$ has a compressed exponential (or superexponential) tail extending up to infinity, of the form

$$f(x) \sim \bar{F}(x) \sim \mathrm{e}^{-Cx^c}, \qquad (6.22)$$

with $c > 1$. The analysis of this case is very similar to the previous one. We have

$$f_2(y) \sim \int_0^y \mathrm{d}x \, \mathrm{e}^{-C(x^c + (y-x)^c)}$$

for large y, where the integral is dominated by a saddle point at $x = y/2$, so that

$$f_2(y) \sim \bar{F}_2(y) \sim \mathrm{e}^{-2^{1-c}Cy^c}.$$

Similarly, the x-integral entering (6.10) is dominated by a saddle point at $x = \tau y$, with $\tau = 1/(1 + 2^{1/(c-1)})$, and so

$$\delta_n \sim \int_0^\infty \mathrm{d}y \, \exp\Big(-2C\tau^{c-1}y^c - n\,\mathrm{e}^{-2^{1-c}Cy^c}\Big).$$

We thus obtain a power-law decay of the form

$$\delta_n \sim n^{-\omega_2(c)},$$

where the exponent

$$\omega_2(c) = \frac{2^c}{(1 + 2^{1/(c-1)})^{c-1}} - 1 \tag{6.23}$$

increases monotonically as a function of c, from $\omega_2(c) \approx (c-1)\ln 2$ as $c \to 1$ to $\omega_2(\infty) = \sqrt{2} - 1$. In particular, for the half-Gaussian distribution ($c = 2$), we predict the decay exponent

$$\omega_{\text{Gaussian}} = \omega_2(2) = \frac{1}{3}. \tag{6.24}$$

As it turns out, the decay exponents $\omega_1(b)$ (see (6.21)) and $\omega_2(c)$ (see (6.23)) can be unified into a single function

$$\omega(\alpha) = 2^{(\alpha+1)/(2\alpha)} \left(1 + 2^{2\alpha/(1-\alpha)}\right)^{(\alpha-1)/(2\alpha)} - 1 \tag{6.25}$$

of a parameter α in the range $-1 < \alpha < 1$, as shown in figure 5. Distributions with a bounded support and an exponential singularity with index b correspond to $-1 < \alpha < 0$, whereas compressed exponential distributions with index c correspond to $0 < \alpha < 1$, with the identifications

$$b = -\frac{\alpha+1}{2\alpha}, \quad c = \frac{\alpha+1}{2\alpha}. \tag{6.26}$$

The exponent $\omega(\alpha)$ is a decreasing function from $\omega(-1) = \omega_1(0) = 1/2$ to $\omega(1) = \omega_2(1) = 0$, via the common limiting value $\omega(0) = \omega_1(\infty) = \omega_2(\infty) = \sqrt{2} - 1$, characteristic of distributions with a double exponential fall-off, either at the upper edge of a compact support or at infinity.

6.5. Subexponential distributions

We now consider subexponential distributions, whose tails decrease more slowly than any exponential. Our goal is to show that the correction δ_n goes to the universal limit (6.2), i.e., that Q_n falls off as

$$Q_n \approx \frac{3}{2n} \tag{6.27}$$

for large n. This result agrees to leading order with the expansion (5.10), ensuing from an exact solution for the power-law distribution with $\theta = 1$. It also agrees with the exact expression (6.33) of Q_n for finite n in the limiting situation of exponentially broad distributions.

The gist of the derivation of (6.27) consists in looking for a solution to the integral recursion (2.14) in an approximately factorized form, i.e.,

$$F_n(\ell, x) \approx K_n\, f(x)\, (1 - \varepsilon_n(\ell, x)). \tag{6.28}$$

The condition (2.13) yields

$$\varepsilon_n(\ell, \ell) = 1, \tag{6.29}$$

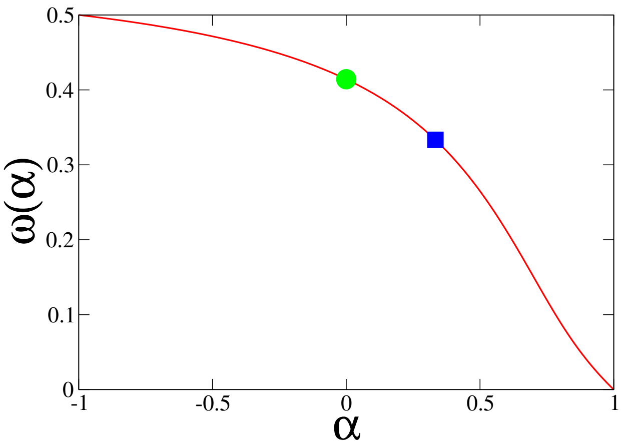

Figure 5. Exponent $\omega(\alpha)$ (see (6.25)) unifying the decay exponents $\omega_1(b)$ (see (6.21)) and $\omega_2(c)$ (see (6.23)) characteristic of distributions with exponential singularities. Green circular symbol: limiting value $\omega(0) = \sqrt{2} - 1$ characteristic of distributions with a double exponential fall-off. Blue square symbol: decay exponent (6.24) of the half-Gaussian distribution ($c = 2$, i.e., $\alpha = 1/3$).

whereas $\varepsilon_n(\ell, x)$ is assumed to be small in the regime of interest where n and ℓ are simultaneously large, with x being kept finite. We set

$$\varepsilon_n(\ell, 0) = 0, \tag{6.30}$$

fixing thus the prefactor K_n unambiguously. To leading order, the differential equation (2.15) yields

$$K_n \varepsilon_n'(\ell, x) \approx K_{n-1} f(\ell - x). \tag{6.31}$$

Equation (6.31), with boundary conditions (6.29) and (6.30), admits a similarity solution where $\varepsilon(\ell, x)$ and the ratio $q = K_n/K_{n-1}$ are independent of n, namely

$$\varepsilon(\ell, x) = \frac{F(\ell) - F(\ell - x)}{F(\ell)}, \quad q = F(\ell).$$

Whenever n and ℓ are simultaneously large, (6.28) simplifies to

$$F_n(\ell, x) \approx \mathrm{e}^{-n\bar{F}(\ell)} f(x) F(\ell - x),$$

so that (2.21) yields the estimate

$$Q_n \approx \int_0^\infty \mathrm{d}\ell\, \mathrm{e}^{-n\bar{F}(\ell)} \int_0^\ell \mathrm{d}x\, f(x) f(\ell - x) F(\ell - x). \tag{6.32}$$

The analysis of this expression for large n is somewhat similar to that of (6.9), performed in section 6.4. The exponential factor selects large values of ℓ, such that $\bar{F}(\ell)$ scales as $1/n$. These are the typical values of L_n. The subexponentiality of $f(x)$, in the intuitive sense explained below (6.11), suggests that the integral over the variable x in (6.32) is dominated by the vicinity of its endpoints, i.e., of the regimes where either x or the difference $\ell - x$ is kept finite. Adding these two contributions yields

$$\int_0^\ell \mathrm{d}x\, f(x)f(\ell-x)F(\ell-x) \approx \frac{3}{2}f(\ell),$$

for ℓ large. Inserting this estimate into (6.32) leads to the announced result (6.27).

The statistics of the number of records M_n for subexponential underlying distributions $f(x)$ will be investigated at the end of section 7.4.

6.6. Exponentially broad distributions

We now consider the limiting class of exponentially broad distributions, defined by setting

$$X = \mathrm{e}^{\Lambda T},$$

where Λ is parametrically large, whereas T has a fixed given distribution $g(t)$. Exponentially broad distributions play a part in the study of strongly disordered systems (see [26, 27] and the references therein). An explicit example is provided by the power-law distribution (see table 1) in the limit where the index θ goes to zero, with the identification $\Lambda = 1/\theta$ and $g(t) = \mathrm{e}^{-t}$. The overwhelming simplification brought by exponentially broad distributions in the $\Lambda \to \infty$ limit is that $X_1 < X_2$ is equivalent to $X_1 \ll X_2$. In other words, the distribution is so broad that, if two independent variables X_1 and X_2 are drawn from the latter, one is negligible with respect to the other with very high probability.

Considering exponentially broad distributions in the $\Lambda \to \infty$ limit is useful for our purpose in several regards. First, the exact probability of record breaking Q_n can be derived in this limit, even for finite n. Second, as we shall see, the derivation gives an insight on the clustering of records underlying the non-trivial limit (1.11). Third, this approach will be readily extended to higher values of the window width p in section 7, where other techniques are not available any more.

Within this setting, it is easy to derive the probability of record breaking Q_n. We recall that Q_n is the probability of having $Y_n > \max(Y_2, \ldots, Y_{n-1})$, with

$$Y_n = X_{n-1} + X_n, \quad Y_{n-1} = X_{n-2} + X_{n-1},$$

and so on. If the variables X are drawn from an exponentially broad distribution, only two events contribute to Q_n:

- The largest of the first n X-variables is X_n. This occurs with probability $1/n$. In the $\Lambda \to \infty$ limit, the variable Y_n is also larger than all previous ones with certainty. Hence the contribution $1/n$ to Q_n.
- The largest of the first n X-variables is X_{n-1}. This again occurs with probability $1/n$. The condition $Y_n > \max(Y_2, \ldots, Y_{n-1})$ reduces to $X_n > X_{n-2}$, so that the relative probability of that event is $1/2$. Hence the contribution $1/(2n)$ to Q_n.

As a consequence, the probability of record breaking is exactly given by

$$Q_n = \frac{3}{2n}, \tag{6.33}$$

for all $n \geq 3$ and all exponentially broad distributions in the $\Lambda \to \infty$ limit.

The formula (6.33) gives both the exact value of Q_n for exponentially broad distributions and its asymptotic decay law (see (1.11), (6.27)) for all distributions with a subexponential tail. The above derivation also demonstrates that the excess in the probability of record breaking (6.33) with respect to the iid situation is due to a clustering of records. The second event of the above list indeed yields two consecutive records. Finally, the data shown in figure 1 suggest that (6.33) provides an absolute upper bound for Q_n. It is indeed quite plausible that the quantity nQ_n plotted in figure 1 converges to the constant 3/2 from below in the $\theta \to 0$ limit, uniformly in n.

7. Extension to higher values of p

In this last section we consider sequences (1.6) obtained by taking the moving average of a sequence of iid variables X_i over an arbitrary finite window width $p \ge 2$. We shall mainly focus on the behavior of the probability of record breaking, that we now denote by

$$Q_n^{(p)} = \mathrm{Prob}(Y_n > L_{n-1}).$$

The recursive structure of the problem described in section 2 still holds true, however it becomes somewhat inefficient, as the number of variables is higher. The recursion equation generalizing (2.14) indeed involves a multiple integral over $p-1$ variables. In particular, no exact solution is available any more. In spite of this, we shall be able to extend to higher values of p most results of interest derived so far for $p = 2$.

7.1. Universal values of the probability of record breaking

The first few values of $Q_n^{(p)}$ are universal, i.e., independent of the underlying distribution $f(x)$. Their values can be derived along the lines of reasoning of section 2.6, using again the Sparre Andersen theorem. The first case of interest is $n = p$, where

$$Q_p^{(p)} = 1. \tag{7.1}$$

There is indeed always a record at $n = p$, as Y_p is the first complete sum of p terms. For $n = p + 1$, we have

$$Q_{p+1}^{(p)} = \mathrm{Prob}(X_{p+1} > X_1) = \frac{1}{2}. \tag{7.2}$$

For $n = p + 2$, we have

$$\begin{aligned} Q_{p+2}^{(p)} &= \mathrm{Prob}(X_{p+2} > X_2, X_{p+1} + X_{p+2} > X_1 + X_2) \\ &= \mathrm{Prob}(X_{p+2} - X_2 > 0, X_{p+2} - X_2 + X_{p+1} - X_1 > 0) \\ &= P_2 = \frac{3}{8}, \end{aligned} \tag{7.3}$$

using the same argument as in section 2.6 for the derivation of $Q_4 = P_2$ for $p = 2$. More generally, for $n = p + k$, with $1 \le k \le p$, we have

$$Q_{p+k}^{(p)} = P_k, \tag{7.4}$$

where the expression of P_k is given in (2.25).

The above formula generalizes the results of section 2.6 to an arbitrary window width $p \ge 2$. It exhausts the list of all universal values of the probability of record breaking. In other words, $Q_{2p+1}^{(p)}$ is the first non-universal one, just as Q_5 for $p = 2$.

7.2. Superexponential and (some) exponential distributions

The explanation given in section 6.3 of the key dichotomy between (1.10) and (1.11), based on the analytic structure of the Laplace transform $\hat{f}(s)$, is not limited to $p = 2$. Its consequences are therefore expected to hold irrespective of the window width p.

For superexponential distributions, as well as for some exponentially decaying distributions, we are therefore again led to compare the original problem to its cyclic variant and to introduce a decoupled model, where the random variables Y_i of the original problem are now replaced by a sequence of iid random variables

$$\mathcal{Y}_i = \underbrace{X_i + X_i' + X_i'' + \cdots}_{p \text{ replicas}},$$

generalizing (6.4). If the sum $\mathcal{Y}$ is atypically large, then all its terms are atypically large as well, with very high probability. We therefore predict a behavior of type (1.10), i.e.,

$$nQ_n^{(p)} = 1 + \delta_n^{(p)},$$

with the following estimate for the small relative correction $\delta_n^{(p)}$:

$$\delta_n^{(p)} \approx n \int_0^\infty \mathrm{d}y\, \mathrm{e}^{-n\bar{F}_p(y)} \int_0^y \mathrm{d}x\, f(x) f_{p-1}(y-x) \bar{F}(x), \tag{7.5}$$

which is a direct generalization of (6.10). We again claim that this prediction is asymptotically correct, to leading order for large n, whenever it decays to zero, i.e., essentially for all superexponential distributions.

The estimate (7.5) is now made explicit for a variety of distributions $f(x)$.

7.2.1. Pure exponential distribution. For an exponential distribution with density $f(x) = \mathrm{e}^{-x}$ and distribution function $F(x) = 1 - \mathrm{e}^{-x}$, we have

$$f_p(y) = \frac{y^{p-1}}{(p-1)!}\, \mathrm{e}^{-y}$$

as well as $\bar{F}_p(y) \approx f_p(y)$, to leading order for $y \gg 1$, and so (7.5) reads

$$\delta_n^{(p)} \approx n \int_0^\infty \mathrm{d}y\, \frac{y^{p-2}}{(p-2)!}\, \exp\left(-y - n\frac{y^{p-1}}{(p-1)!}\, \mathrm{e}^{-y}\right).$$

This integral can be evaluated along the lines of (3.9) and (6.12). Omitting details, we obtain to leading order

$$\delta_n^{(p)} \approx \frac{p-1}{\ln n}.$$

This estimate vanishes identically for $p = 1$ and coincides with (6.13) for $p = 2$. It demonstrates that the marginal character of the exponential distribution, with its logarithmic correction term, persists to all higher values of p.

7.2.2. Exponential distribution modulated by a power law. We now consider distributions falling off as an exponential modulated by a power law, i.e.,

$$f(x) \approx \bar{F}(x) \approx A\, x^{a-1}\, \mathrm{e}^{-x} \quad (x \to \infty). \tag{7.6}$$

Along the lines of section 6.4.2, let us consider first the case where $a > 0$. We have

$$f_p(y) \approx \bar{F}_p(y) \approx B_p\, y^{pa-1}\, \mathrm{e}^{-y} \quad (y \to \infty),$$

with $B_p = (A\Gamma(a))^p/\Gamma(pa)$. Performing the integrals entering (7.5), we are left with the estimate

$$\delta_n \approx \frac{A\Gamma(pa)}{\Gamma((p-1)a)}\,\frac{1}{(\ln n)^a}. \tag{7.7}$$

This formula is a direct generalization of (6.16). We thus conclude that exponential distributions of the form (7.6) with $a > 0$ belong to the superexponential side of the dichotomy, in the sense that (1.12) holds, with a correction falling off as (7.7). On the other hand, along the lines of section 6.4.2, we are led to claim that exponential distributions with $a < 0$ lead to (1.13), just as subexponential distributions.

7.2.3. Distributions with bounded support and power-law singularity. In the case where $f(x)$ is supported by the interval [0, 1] and has a power-law singularity of the form (6.17) at its upper edge, we have

$$\bar{F}_p(y) \approx B_p \eta^{pa}, \quad f_p(y) \approx paB_p\eta^{pa-1},$$

with $\eta = p - y$ and $B_p = a^{p-1}(A\Gamma(a))^p/(p\Gamma(pa))$. Performing the integrals entering (7.5), we obtain a power-law decay for $\delta_n^{(p)}$, i.e.,

$$\delta_n^{(p)} \approx \frac{K(p,a)}{n^{1/p}}, \tag{7.8}$$

where the exponent only depends on the width p, whereas the amplitude $K(p,a)$ reads

$$K(p,a) = \frac{\Gamma(1/p)\,\Gamma(2a)\,\Gamma(pa)^{1+1/p}}{(pa)^{1-1/p}\,\Gamma(a)^2\,\Gamma((p+1)a)}.$$

This result extends (6.19) to higher values of p. The amplitude $K(p,a)$ has a local maximum for $a = 0$, a local minimum for $a = 1$, and grows exponentially fast at large a. All these features hold irrespective of p, and survive in the formal $p \to \infty$ limit, i.e.,

$$K(\infty,a) = \frac{\mathrm{e}^{-a}\,\Gamma(2a)}{a\,\Gamma(a)^2}.$$

7.2.4. Distributions with exponential singularities. To close, we consider distributions with a bounded support and an exponentially small singularity at their upper edge, of the form (6.20), as well as compressed distributions with a superexponential tail extending up to infinity, of the form (6.22).

We again obtain a power-law decay for the correction $\delta_n^{(p)}$, with continuously varying decay exponents $\omega_1(p,b)$ and $\omega_2(p,c)$, which can be unified into a single monotonically decreasing function

$$\omega(p,\alpha) = 2\left(\frac{1+(p-1)2^{2\alpha/(1-\alpha)}}{p}\right)^{(\alpha-1)/(2\alpha)} - 1$$

of the parameter α in the range $-1 < \alpha < 1$, with the identifications (6.26). We have in particular $\omega(p,-1) = 1/p$, ensuring a smooth crossover with (7.8), $\omega(p,0) = \omega_1(p,\infty) = \omega_2(p,\infty) = 2^{1/p} - 1$ for the limiting situation of distributions with a double exponential fall-off, and $\omega(p,1/3) = 1/(2p-1)$, corresponding e.g. to the half-Gaussian distribution.

7.3. Exponentially broad distributions

For exponentially broad distributions in the $\Lambda \to \infty$ limit, the expression of $Q_n^{(p)}$ can be derived along the lines of section 6.6. We recall that $Q_n^{(p)}$ is the probability of having $Y_n > \max(Y_p, \dots, Y_{n-1})$, with

$$Y_n = X_{n-p+1} + \cdots + X_n, \quad Y_{n-1} = X_{n-p} + \cdots + X_{n-1},$$

and so on. If the X-variables are drawn from an exponentially broad distribution, only the following events contribute to $Q_n^{(p)}$:

- The largest of the first n X-variables is X_n. This occurs with probability $1/n$. In the $\Lambda \to \infty$ limit, the variable Y_n is also larger than all previous ones with certainty. Hence the contribution $1/n$ to $Q_n^{(p)}$.
- The largest of the first n X-variables is X_{n-1}. This again occurs with probability $1/n$. The condition $Y_n > \max(Y_2, \dots, Y_n)$ reduces to $X_n > X_{n-2}$, so that the relative probability of that event is $1/2$. Hence the contribution $1/(2n)$ to $Q_n^{(p)}$.
- The largest of the first n X-variables is X_{n-2}. This again occurs with probability $1/n$. The condition $Y_n > \max(Y_2, \dots, Y_n)$ reduces to

$$\begin{aligned} &X_n - X_{n-3} > 0, \\ &X_n - X_{n-3} + X_{n-1} - X_{n-4} > 0. \end{aligned}$$

  so that the relative probability of that event is $P_2 = 3/8$, again by virtue of the Sparre Andersen theorem. Hence the contribution P_2/n to $Q_n^{(p)}$, and so on.

Summing up the probabilities of the events listed above, we predict that the probability of record breaking is exactly given by

$$Q_n^{(p)} = \frac{R_p}{n}, \tag{7.9}$$

for all $p \geq 2$ and $n \geq 2p-1$ and all exponentially broad distributions in the $\Lambda \to \infty$ limit. The numerator of the above formula reads

$$R_p = \sum_{k=0}^{p-1} P_k,$$

where the integer k numbers the items of the above list and where the expression of P_k is given in (2.25). Equation (2.26) yields

$$\widetilde{R}(z) = \sum_{p \geq 1} z^p R_p = \frac{z}{1-z}\, \widetilde{P}(z) = \frac{z}{(1-z)^{3/2}}.$$

The R_p are therefore universal rational numbers given by

$$\begin{aligned} R_p &= \frac{(2p-1)!}{2^{2p-2}(p-1)!^2} = 2pP_p = (2p-1)P_{p-1} \\ &= 1,\ \frac{3}{2},\ \frac{15}{8},\ \frac{35}{16},\ \frac{315}{128},\ \frac{693}{256}, \dots, \end{aligned} \tag{7.10}$$

for $p = 1, 2, 3, 4, 5, 6, \dots$, and growing as

$$R_p \approx 2\sqrt{\frac{p}{\pi}}$$

at large p.

The formulas (7.4) and (7.9) overlap for two values of n, namely $2p-1$ and $2p$, for which they consistently predict

$$Q_{2p-1}^{(p)} = P_{p-1} = \frac{R_p}{2p-1}, \quad Q_{2p}^{(p)} = P_p = \frac{R_p}{2p}.$$

7.4. Subexponential distributions

Following the line of thought sketched in the very beginning of section 7.2, we are led to extend the dichotomy between (1.10) and (1.11) to higher values of p, and to predict the following asymptotic decay of the probability of record breaking at large n:

$$Q_n^{(p)} \approx \frac{R_p}{n}, \tag{7.11}$$

for all $p \geq 2$ and all subexponential distributions $f(x)$, where the amplitude R_p is predicted by the exact analysis of the limiting case of exponentially broad distributions (see section 7.3). The latter amplitude, given by (7.10), is therefore universal, in the sense that it only depends on the window width p.

The formula (7.9) therefore has the same status as (6.33). It gives the exact value of $Q_n^{(p)}$ for exponentially broad distributions in the $\Lambda \to \infty$ limit for all $n \geq 2p-1$. It is also expected to describe the asymptotic decay law of $Q_n^{(p)}$ for all subexponential distributions, and furthermore to provide an absolute upper bound for $Q_n^{(p)}$ for all $n \geq p$.

8. Discussion

This paper was devoted to the statistics of records for the moving average of a sequence of iid variables. Most results concern the case where the window width is $p = 2$. The main emphasis has been put on the probability of record breaking Q_n at time n, and on the distribution of the number of records M_n up to time n. In sections 3 to 5 we have given full analytical solutions of the problem for three particular parent distributions: exponential, uniform and power-law with $\theta = 1$. The exact results obtained there provide useful checks of the heuristic approach used in the asymptotic analysis of the general situation (section 6) and in its extension to higher values of p (section 7).

Quite serendipitously, the three distributions which have lent themselves to an exact analytical treatment are prototypical in several regards. First, each of them is a representative of one of the three universality classes of extreme value statistics: Weibull, Gumbel and Fréchet. Second, they are also representatives of the dichotomy, as regards the properties of records for the moving average, between superexponential distributions, where the product nQ_n tends to unity and the distribution of the number of records is asymptotically Poissonian, and subexponential distributions, where nQ_n admits the non-trivial universal limit $3/2$, or more generally R_p, and the distribution of the number of records exhibits novel universal clustering features. The uniform and power-law distributions are respectively typical of the superexponential and subexponential classes, whereas the exponential distribution is a representative of the exponential class, which is marginal and split on both sides of the dichotomy, as seen in section 6.4.2.

Our main results can be summarized in the sketchy representation of the realm of parent probability distributions shown in figure 6. The tail of the distribution is more and more heavy, i.e., the density $f(x)$ falls off more and more slowly, as one progresses from left to right. The red line in figure 6 represents the boundary of the dichotomy, with superexponential distributions to its left and subexponential distributions to its right, with the marginal class of exponential distributions sitting on the line itself. To the left of the red line, the product nQ_n tends to unity, just as for records of iid variables. Superexponential distributions can be classified according to the exponent ω describing the power-law decay $\delta_n \sim n^{-\omega}$ of the correction such that $nQ_n = 1 + \delta_n$. For distributions in the Weibull class, i.e., with a bounded support and a power-law singularity at its upper end, ω is constant and equal to $1/2$, and more generally $1/p$. For superexponential distributions in the Gumbel class, whose support is either bounded (Region I) or unbounded (Region II), the exponent $\omega(b)$ or $\omega(c)$ decreases from $1/2$ to 0, and more generally $\omega(p, b)$ or $\omega(p, c)$ decreases from $1/p$ to 0. To the right of the red line, the product nQ_n admits the non-trivial universal limit $3/2$, and more generally R_p. This limit holds both for subexponential distributions falling off faster than any power of x (Region III of the Gumbel class) and for those exhibiting a power-law tail (Fréchet class).

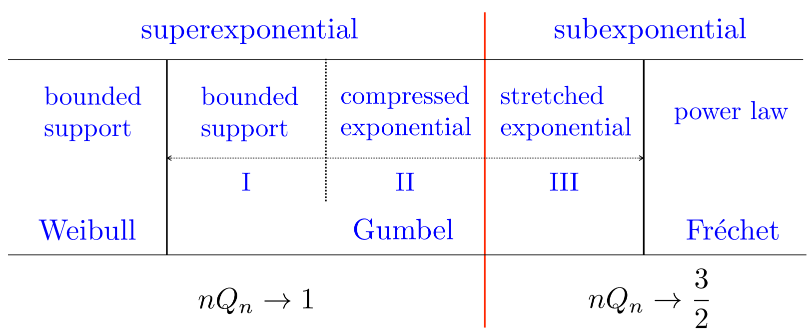

Figure 6. Sketchy representation of the realm of parent probability distributions (see text for details). The last line summarizes the results for the case $p = 2$.

The key dichotomy highlighted in the present work for the properties of records of the moving average, between (1.10) and (1.11) (or more generally between (1.12) and (1.13)), i.e., essentially between the subexponential and superexponential classes of distributions, appears as very robust. It is therefore expected to have far-reaching consequences on other quantities, besides the probability of record breaking Q_n and the number of records M_n. Consider the example of the distribution of the maximum L_n of the first n daughter Y-variables. The heuristic approach put forward in section 6 suggests that the distribution of L_n is close to that of the maximum of n iid X-variables for subexponential distributions, to the right of the red line, whereas it is close to that of the maximum of n iid variables of the form $Y = X + X'$ for superexponential distributions, to the left of the red line. This claim is corroborated by the exact or asymptotic expressions for the mean or median values of L_n derived in sections 3 to 5.

Let us close with a word on more general linear filters of the form

$$Y_n = \sum_{k \geq 0} K_k X_{n-k},$$

used e.g. in digital signal processing, transforming a sequence of iid random variables X_n to a filtered sequence Y_n, whose entries are clearly not iid any more. Many open questions of interest related to extremes and records in such filtered sequences could be addressed. It can be anticipated that the occurrences of records will exhibit some clustering, especially if the distribution $f(x)$ of the parent variables is broad enough, even though a clear-cut universal dichotomy is not to be expected in general.

Acknowledgments

CG is grateful to S Majumdar for arousing his interest in the statistics of records for sequences made of sums of successive iid random variables.

References

[1] Chandler K N 1952 J. Roy. Statist. Soc. Ser. B **14** 220
[2] Rényi A 1962 Ann. Sci. Univ. Clermont-Ferrand **8** 7
[3] Rényi A 1962 Proceedings Coll. Combinatorial Methods in Probability Theory (Math. Inst. Aarhus Univ., Aarhus, Denmark)
[4] Glick N 1978 Amer. Math. Monthly **85** 2
[5] Redner S and Petersen M R 2006 Phys. Rev. E **74** 061114
[6] Wergen G and Krug J 2010 Europhys. Lett. **92** 30008
[7] Coumou D and Rahmstorf S 2012 Nature Climate Change **2** 491
[8] Coumou D, Robinson A and Rahmstorf S 2013 Climatic Change **118** 771
[9] Wergen G, Hense A and Krug J 2014 Climate Dynamics **42** 1275
[10] Godrèche C, Majumdar S N and Schehr G 2017 J. Phys. A **50** 333001
[11] Arnold B C, Balakrishnan N and Nagaraja H N 1998 *Records* (New York: Wiley)
[12] Nevzorov V B 2004 *Records: Mathematical Theory* (Providence, RI: American Mathematical Society)
[13] Gnedenko B 1943 Ann. Math. **44** 423
[14] Majumdar S N and Dhar D 2001 Phys. Rev. E **64** 046123
[15] Majumdar S N 2002 Phys. Rev. E **65** 035104(R)
[16] Majumdar S N and Dean D S 2002 Phys. Rev. E **66** 041102
[17] Sparre Andersen E 1953 Math. Scand. **1** 263
[18] Sparre Andersen E 1954 Math. Scand. **2** 195
[19] Feller W 1971 *An Introduction to Probability Theory and its Applications* Vol 2 2nd ed (New York: Wiley)
[20] Bertin E 2005 Phys. Rev. Lett. **95** 170601
[21] Bertin E and Clusel M 2006 J. Phys. A **39** 7607
[22] Stanley R P 1986 Discrete Comput. Geom. **1** 9
[23] OEIS The On-line Encyclopedia of Integer Sequences `https://oeis.org`
[24] Chistyakov V P 1964 Theor. Probab. Appl. **9** 640
[25] Godrèche C 2019 J. Stat. Mech. 063207
[26] Tyč S and Halperin B I 1989 Phys. Rev. B **39** 877
[27] Le Doussal P 1989 Phys. Rev. B **39** 881