The distribution of maximal prime gaps in Cramér’s probabilistic model of primes

Alexei Kourbatov
JavaScripter.net/math
akourbatov@gmail.com

Abstract

In the framework of Cramér’s probabilistic model of primes, we explore the exact and asymptotic distributions of maximal prime gaps. We show that the Gumbel extreme value distribution \( \exp(-e^{-x}) \) is the limit law for maximal gaps between Cramér’s random “primes.” The result can be derived from a general theorem about intervals between discrete random events occurring with slowly varying probability monotonically decreasing to zero. A straightforward generalization extends the Gumbel limit law to maximal gaps between prime constellations in Cramér’s model.

1 Introduction

In this paper we apply extreme value theory to Cramér’s probabilistic model of primes \([4]\). But first let us say a few words about two mathematicians who pioneered these topics in mathematics. The Swedish mathematician Harald Cramér (1893–1985) made long-lasting contributions in statistics and number theory. His model of primes continues to serve as a heuristic tool leading to new insights in the distribution of primes. Cramér wrote:

In investigations concerning the asymptotic properties of arithmetic functions, it is often possible to make an interesting heuristic use of probability arguments. If, e.g., we are interested in the distribution of a given sequence \( S \) of integers, we then consider \( S \) as a member of an infinite class \( C \) of sequences, which may be concretely interpreted as the possible realizations of some game of chance. It is then in many cases possible to prove that, with a probability \( = 1 \), a certain relation \( R \) holds in \( C \), i.e. that in a definite mathematical sense “almost all” sequences of \( C \) satisfy \( R \). Of course we cannot in general conclude that \( R \) holds for the particular sequence \( S \), but results suggested in this way may sometimes afterwards be rigorously proved by other methods. \([4]\) p. 25]
It is difficult to ascertain whether Harald Cramér had ever met in person his contemporary, the German-American mathematician Emil Julius Gumbel (1891–1966), one of the founders of extreme value theory [8, 9] which is used today for describing phenomena in vastly diverse areas, ranging from actuarial science to hydrology to number theory. In Statistics of Extremes (1958) Gumbel observed:

... many engineers and practical statisticians ... are inclined to believe that, after all, nearly everything should be normal, and whatever turns out not to be so can be made normal by a logarithmic transformation. This is neither practical nor true. [9, p. 345]

In Les valeurs extrêmes des distributions statistiques (1935) Gumbel showed that extreme values taken from a sequence of i.i.d. random variables with an exponential distribution obey the double exponential limit law (now known as the Gumbel distribution) [8]. He reconfirmed the earlier result of Fisher and Tippett [4] that the same limit law also holds for extreme values of i.i.d. random variables with a normal distribution — and generalized it to a much wider class of initial distributions, the so called exponential type. To wit:

Pour les valeurs extrêmes ... on arrive à la distribution doublement exponentielle pourvu que la distribution initiale appartienne au type exponentiel ... Cette théorie est susceptible de nombreuses applications, puisque en particulier, la distribution de Gauss et la distribution exponentielle appartiennent au type exponentiel. [8, p. 154–155]

We will draw upon the work of both Cramér and Gumbel: In Section 4.2 we show that the limiting distribution of maximal prime gaps in Cramér’s probabilistic model is the Gumbel extreme value distribution. In the hindsight, the result is not very surprising: we know from computation and distribution fitting that the actual maximal prime gaps are indeed nicely approximated by the Gumbel distribution [12]. (What is somewhat surprising is that Cramér or Gumbel have not themselves published a similar result long ago, perhaps in the 1930s or 1940s. Did they possibly deem it obvious? We might never know.)

Once we establish the existence of a limit law for maximal prime gaps in Cramér’s model with \( n \) urns, additional questions naturally arise (they are to be addressed elsewhere):

- Is the Gumbel distribution, after proper rescaling, also an (almost sure) limit law for record prime gaps observed in a single infinite sequence of Cramér’s random primes? (Cf. [18] and [2, p. 193].)

- Is the Gumbel distribution, after proper rescaling, also a limit law for record gaps between true primes? (The latter question appears particularly difficult.)

The existence of the limiting Gumbel distribution for the (non-identically distributed) maximal gaps between Cramér’s random primes is also interesting in view of Mejzler’s theorem [10, p. 201]. For non-identically distributed independent random variables, Mejzler’s theorem states that the limiting distribution (if it exists at all) can be any distribution with a log-concave cdf. Thus, prime gaps in Cramér’s model give us an example of non-identically distributed random variables which nevertheless possess a limit law that is allowed in i.i.d. situations as well.
2 Notations and abbreviations

a.s. almost sure, almost surely
i.i.d. independent and identically distributed
cdf cumulative distribution function
pdf probability density function
\(Ex\) the expected value (mathematical expectation) of the random variable \(x\)
\(\text{Exp}(x; a)\) the exponential distribution cdf: \(\text{Exp}(x; a) = 1 - e^{-x/a}\)
\(\text{Gumbel}(x; a, \mu)\) the Gumbel distribution cdf: \(\text{Gumbel}(x; a, \mu) = 1 - e^{-e^{-x/M}}\)
\(a\) the scale parameter of exponential/Gumbel distributions, as applicable
\(\mu\) the location parameter (mode) of the Gumbel distribution
\(M\) the median of \(\text{Gumbel}(x; a, \mu)\): \(M = \mu - a \log \log 2 \approx \mu + 0.3665a\)
RP a random “prime” in the context of Cramér’s model (a white ball)
RC a random “composite” in the context of Cramér’s model (a black ball)
\(p_k\) the \(k\)-th prime; \(\{p_k\} = \{2, 3, 5, 7, 11, \ldots\}\)
\(P_k\) the \(k\)-th random “prime” (RP) in Cramér’s model
\(U_n\) the \(n\)-th urn producing RPs with probability \(\frac{1}{\log n}\) in Cramér’s model
\(R_n\) a random variable: the longest uninterrupted run of RCs \(\leq n\)
\(G_n\) a random variable: the maximal gap between RPs \(\leq n\); \(G_n := R_n + 1\)
\(\pi(x)\) the prime counting function, the total number of primes \(p_k \leq x\)
\(\Pi(x)\) the RP counting function, the total number of RPs \(P_k \leq x\)
\(\log x\) the natural logarithm of \(x\)
\(\text{li} x\) the logarithmic integral of \(x\): \(\text{li} x = \int_0^x \frac{dt}{\log t} = \int_2^x \frac{dt}{\log t} + 1.04516\ldots\)

2.1 Definitions of gaps and runs

Prime gaps are distances between two consecutive primes, \(p_k - p_{k-1}\) [15, 22, OEIS A001223]. In the context of Cramér’s model (see next section) “prime” gaps refer to distances between consecutive RPs, \(P_k - P_{k-1}\).

Maximal gaps between primes are usually defined as gaps that are strictly greater than all preceding gaps [15, 22, A005250]. However, for Cramér’s model, we will use the term maximal gap in the statistical sense defined below. Note that Cramér’s model does not guarantee that there are any “primes” \(P_k > 2\) at all. To make sure that maximal gaps are

---

1 Number theorists may use the terms maximal gaps and record gaps as synonyms [8], while statisticians make a distinction between maximal and record values [2, 14, 18]. Resnick [18, Theorem 8] shows that the distinction is quite profound: in the i.i.d. case, the limit laws for records and maxima cannot be the same. Nevertheless, for a wide class of sequences of non-identically distributed random variables (the \(F^\alpha\) model) the same extreme value distribution (e.g. Gumbel distribution) can be the limit law for both records and maxima [2, p.193]. Clearly, Cramér’s maximal “prime” gaps near, say, urn \(U_{10}\) are distributed not identically to those near urn \(U_{100}\). In Sect. 4.2 we show that the limiting distribution of Cramér’s maximal “prime” gaps is indeed the Gumbel distribution. Computational evidence suggests that the Gumbel distribution is also the a.s. limit law for record gaps in Cramér’s model; we will discuss the distribution of records elsewhere.
defined in all cases, we first define the longest run of random “composites” (RCs) \( \leq n \):

\[ R_n = \text{the longest run of consecutive RCs } \leq n \text{ (allowing runs of length 0)}. \]

We now define the maximal gap between RPs \( \leq n \) simply as

\[ G_n = R_n + 1. \]

Compare our definition of \( G_n \) to the definition of maximal prime gaps for true primes [15]:

\[ \text{maximal prime gap up to } n = \max_{p_k \leq n} (p_k - p_{k-1}). \]

Clearly, the same gap/run relation holds for true primes as well[2]

\[ \text{maximal prime gap} = 1 + \text{the longest run of composites below } n. \]

Of course, in the probabilistic model, both \( R_n \) and \( G_n \) are random variables. In what follows, we will investigate the distribution of maximal gaps \( G_n \).

3 Cramér’s probabilistic model of primes

Cramér’s probabilistic model of primes [4] is well known (and much criticized [7, 17]).

3.1 Setting up the model

Cramér [4] sets up the model of primes as follows:

With respect to the ordinary prime numbers, it is well known that, roughly speaking, we may say that the chance that a given integer \( n \) should be a prime is approximately \( \frac{1}{\log n} \). This suggests that by considering the following series of independent trials we should obtain sequences of integers presenting a certain analogy with the sequence of ordinary prime numbers \( p_n \).

Let \( U_1, U_2, U_3, \ldots \) be an infinite series of urns containing black and white balls, the chance of drawing a white ball from \( U_n \) being \( \frac{1}{\log n} \) for \( n > 2 \), while the composition of \( U_1 \) and \( U_2 \) may be arbitrarily chosen. We now assume that one ball is drawn from each urn, so that an infinite series of alternately black and white balls is obtained. If \( P_n \) denotes the number of the urn from which the \( n \)-th white ball in the series was drawn, the numbers \( P_1, P_2, \ldots \) will form an increasing sequence of integers, and we shall consider the class \( C \) of all possible sequences \( (P_n) \). Obviously the sequence \( S \) of ordinary prime numbers \( (p_n) \) belongs to this class.

We shall denote by \( \Pi(x) \) the number of those \( P_n \) which are \( \leq x \), thus forming an analogy to the ordinary notation \( \pi(x) \) for the number of primes \( p_n \leq x \). [4 pp. 25–26]

[2] with rare exceptions occurring towards the end of record prime gaps. Consider, e.g., the prime gap 113 . . . 127 and take \( n = 125 \). Then the largest prime gap below \( n \) is the gap of 8 (between 89 and 97), while the longest run of 12 composites \( \leq n \) occurs from 114 to 125.
Cramér’s model, as stated, is underdetermined: the content of urns $U_1$ and $U_2$ is arbitrary. To compute exact distributions of maximal gaps, we will assume that

- urn $U_1$ is empty — it produces neither “primes” nor “composites”;
- urn $U_2$ always produces white balls (i.e. the number 2 is certain to be “prime”).

### 3.2 Cramér’s results

Using his probabilistic model of primes, Cramér [4, pp. 26–28] obtained the following results:

- The expected value of $\Pi(x)$ is asymptotically equal to $\text{li} x$: $E\Pi(x) \sim \text{li} x$ as $x \to \infty$.
- With probability = 1, we have: $\limsup_{x \to \infty} \frac{|\Pi(x) - \text{li} x|}{\sqrt{2x \log \log x / \log x}} = 1$.
- With probability = 1, we have: $\limsup_{k \to \infty} \frac{P_{k+1} - P_k}{(\log P_k)^2} = 1$.

The latter result, restated for true primes $p_k$, constitutes the well-known Cramér’s conjecture. Computations of Oliveira e Silva, Herzog, and Pardi [16] have verified that

- $p_{k+1} - p_k < \log^2 P_k$ for all primes $p_k \notin \{2, 3, 7\}$ up to $4 \times 10^{18}$,
- $p_k - p_{k-1} < \log^2 P_k$ for all primes up to $4 \times 10^{18}$.

Thus the conjecture appears quite likely to be true.

### 4 Maximal gaps between Cramér’s random primes

#### 4.1 The exact distribution of maximal gaps

We continue where Cramér left off. To obtain exact distributions of maximal gaps between Cramér’s random primes, for now we will restrict ourselves to finite sets of $n$ consecutive urns $U_1, U_2, \ldots, U_n$. When our set of urns is small, we can compute the exact distributions of maximal gaps by hand, even without a computer.

For example, for $n = 3$ we have only three urns: $U_1, U_2, U_3$. Of these, only $U_3$ produces random results:

- a white ball (RP) with probability $q_3 = \frac{1}{\log 3} \approx 0.91$, or
- a black ball (RC) with probability $1 - q_3 = 1 - \frac{1}{\log 3} \approx 0.09$.

Thus, for the longest run of RCs, $R_3$, we have $R_3 = 0$ with probability 0.91, and $R_3 = 1$ with probability 0.09. Consequently, for the maximal gap between RPs, we have $G_3 = 1$ with probability 0.91, and $G_3 = 2$ with probability 0.09. (Recall that, by definition, $G_n = R_n + 1$.)

One can visualize the exact distributions of maximal gaps in the form of histograms. With the help of a computer, we can find the exact distributions up to, say, $n = 40$ urns. Figure 1 shows the respective computer-generated distributions of maximal gaps.
Figure 1: Histograms of the exact distributions of maximal prime gaps in Cramer’s model with $n$ urns. (The histograms have been computed using exact formulas for $n \leq 40$, and a Monte-Carlo algorithm for larger $n$.) The curves represent Gumbel distributions (pdf) with scale $a = n/\text{li } n$ and mode $\mu = n \log(\text{li } n)/\text{li } n$. 
4.2 The limiting distribution of maximal gaps

In Figure 1, the exact distributions (histograms) of maximal gaps between RPs appear to approach the pdf curves of the Gumbel distribution with scale $\alpha_n = n/\ln n$ and mode $\mu = \alpha_n \log \ln n$. We can restate this observation in a more precise form:

**Theorem 1.** In Cramér’s model with $n$ urns, the Gumbel distribution $\exp(-e^{-z})$ is the limiting distribution of maximal gaps $G_n$ between RPs: there exist $a_n > 0$ and $b_n$ such that

$$\lim_{n \to \infty} P(G_n \leq x \equiv a_n z + b_n) = \exp(-e^{-z}), \quad \text{where} \quad a_n \sim \frac{n}{\ln n}, \quad b_n \sim \alpha_n \log \ln n.$$ 

Equivalently, for $R_n$ (the longest runs of RCs) we have

$$\lim_{n \to \infty} P(R_n \leq a_n z + b_n) = \exp(-e^{-z}).$$

We will sketch two proofs of Theorem 1. The first proof will use the following lemmas.

**Lemma of Common Median.** Suppose two Gumbel distributions have a common median and different scales $a \pm \varepsilon$, where $0 < \varepsilon < \frac{3}{2}$. Then the cdfs of these Gumbel distributions differ by no more than $\delta \frac{\varepsilon}{a}$.

**Lemma of Common Scale.** Suppose two Gumbel distributions have a common scale $a \geq 1$ and medians $M \pm \delta$. Then these Gumbel cdfs differ by no more than $\delta \frac{\varepsilon}{a}$.

Denote by $F_n(x)$ the cdf of the exact distribution of maximal gaps in Cramér’s model with $n$ urns $U_1,\ldots,U_n$ ($n \geq 2$), as defined in Section 3.1. Denote by $I_n$ ($n \geq 10$) the largest interval of the $x$ axis such that for all $x \in I_n$ we have $F_n(x) \in [(\log n)^{-1}, 1 - (\log n)^{-1}]$. (One can show that $I_n \subset [\log n, \log^2 n]$.)

**Squeeze Lemma.** Let $M_n$ be the median of $F_n(x)$. Then $F_n(x)$ is squeezed in the area bounded by the Gumbel distribution cdfs with medians $M = M_n \pm \delta$ and scales $a = n/\ln n \pm \varepsilon$:

$$\min_{M = M_n \pm \delta} 2^{-e^{-\frac{x-M}{a}}} \leq F_n(x) \leq \max_{M = M_n \pm \delta} 2^{-e^{-\frac{x-M}{a}}} \quad \text{for} \quad n \geq 10, \quad x \in I_n,$$

where one can take $\varepsilon = 3/2$, $\delta = 1$. (See Fig. 2. The sequence $\{M_n\}$ is OEIS A235492 [22].)

Hereafter we often use the following expressions giving the Gumbel distribution cdf in terms of its scale $a$, mode $\mu$ and median $M$:

$$\text{Gumbel}(x; a, \mu) = e^{-e^{-\frac{x-\mu}{a}}} = 2^{-e^{-\frac{\mu-M}{a}}}, \quad \text{where} \quad M = \mu - a \log \log 2 \approx \mu + 0.3665a.$$ 

4.2.1 First proof of Theorem 1

Suppose the number of urns $n$ is large. We will examine two cases: (a) $x \in I_n$; (b) $x \not\in I_n$.

**Case (a)** First, we consider $F_n(x)$ for $x \in I_n$, i.e., $(\log n)^{-1} \leq F_n(x) \leq 1 - (\log n)^{-1}$.

Let us estimate $M_n$, the median of $F_n(x)$. As before, let $a_n = n/\ln n$ and $\varepsilon = 3/2$. We can approximate Cramér’s proportions of white balls in urns using two different procedures:
(i) set the white-to-black-balls ratio in all urns to $\frac{1}{\alpha_n + \varepsilon}$, then increase the percentages of white balls in all urns or, alternatively,

(ii) set the white-to-black-balls ratio in all urns to $\frac{1}{\alpha_n - \varepsilon}$, then reduce the percentages of white balls in all but a small subset of urns.

Observe that increasing the percentage of white balls pushes the median of maximal gaps to the left, while reducing this percentage pushes the median to the right. Therefore, $M_n$ must be somewhere between the medians of $(\text{Exp}(x; \alpha_n - \varepsilon))^{\text{lin}}$ and $(\text{Exp}(x; \alpha_n + \varepsilon))^{\text{lin}}$:

$$\text{median } (\text{Exp}(x; \alpha_n - \varepsilon))^{\text{lin}} \leq M_n \leq \text{median } (\text{Exp}(x; \alpha_n + \varepsilon))^{\text{lin}}.$$

Since $\alpha_n = O(\log n)$ while $\varepsilon = O(1)$, the above lower and upper bounds are asymptotic to each other and to the median of $(\text{Exp}(x; \alpha_n))^{\text{lin}}$, therefore we must have

$$M_n \sim \text{median } (\text{Exp}(x; \alpha_n))^{\text{lin}} \text{ as } n \to \infty.$$

That is to say, as $n \to \infty$, the median $M_n$ of $F_n(x)$ must be asymptotic to the median of the limiting distribution of maxima of $|\text{li } n|$ i.i.d. random variables with the exponential distribution $\text{Exp}(x; \alpha_n)$. But this limiting distribution is precisely the Gumbel distribution $\text{Gumbel}(x; \alpha_n, \alpha_n \log \text{li } n)$ [6, 8, 11], so

$$M_n \sim \text{median } (\text{Gumbel}(x; \alpha_n, \alpha_n \log \text{li } n)) \approx \alpha_n \log \text{li } n + 0.3665 \alpha_n.$$

On the other hand, it follows from Lemmas that $F_n(x)$ is squeezed in the $\Delta_n$-neighborhood of the Gumbel cdf $2^{-e^{-\frac{x-M_n}{\alpha_n}}}$, where $\varepsilon$ and $\delta$ are defined as in Squeeze Lemma,

$$\Delta_n = \frac{\varepsilon + \delta}{\alpha_n}, \quad \text{and} \quad \alpha_n = \frac{n}{\text{li } n} \sim \log n \to \infty \text{ as } n \to \infty.$$

Therefore, we have $\Delta_n \to 0$ as $n \to \infty$.

To sum it up: the median $M_n$ is asymptotic to the median of $\text{Gumbel}(x; \alpha_n, \alpha_n \log \text{li } n)$, while the limiting shape of $F_n(x)$ is dictated by the fact that $F_n(x)$ is in $\Delta_n$-neighborhood of the Gumbel cdf $2^{-e^{-\frac{x-M_n}{\alpha_n}}}$, with $\lim_{n \to \infty} \Delta_n = 0$. This completes the proof for case (a).

**Case (b)** Now consider $x \notin I_n$: either $F_n(x) < (\log n)^{-1}$ or $F_n(x) > 1 - (\log n)^{-1}$. Taking into account the monotonicity of cdfs, we can conclude from Lemmas that, for large $n$,

$$2^{-e^{-\frac{x-M_n}{\alpha_n}}} < \frac{\varepsilon + \delta}{\alpha_n} + (\log n)^{-1} \quad \text{if} \quad F_n(x) < (\log n)^{-1},$$

$$1 - 2^{-e^{-\frac{x-M_n}{\alpha_n}}} < \frac{\varepsilon + \delta}{\alpha_n} + (\log n)^{-1} \quad \text{if} \quad 1 - F_n(x) < (\log n)^{-1},$$

where $\alpha_n = \frac{n}{\text{li } n} \sim \log n$. Therefore, $|F_n(x) - 2^{-e^{-\frac{x-M_n}{\alpha_n}}}| = O((\log n)^{-1}) \to 0$ as $n \to \infty$. \(\square\)
4.2.2 Second proof of Theorem 1

We begin by proving the theorem for longest runs $R_n$ of “composites”. In Cramér’s model, urns $U_n$ ($n \geq 3$) produce white balls with a monotonically decreasing, slowly varying probability $1/\log n \to 0$ as $n \to \infty$. We observe that the more general Theorem 2 (see Appendix) is applicable to our situation. Theorem 2 tells us that, as $n \to \infty$, the limiting distribution of longest runs exists; it is the Gumbel distribution with the scale and mode parameters determined by $E \Pi(n)$, the expected total number of RPs $\leq n$:

$$
scale \ a_n \sim \frac{n}{E \Pi(n)} = \frac{n}{\operatorname{li} n + O(1)} \sim \frac{n}{\operatorname{li} n}$$

$$
mode \ b_n \sim \frac{n \log E \Pi(n)}{E \Pi(n)} = \frac{n}{\operatorname{li} n + O(1)} \log(\operatorname{li} n + O(1)) \sim \frac{n}{\operatorname{li} n} \log \operatorname{li} n.
$$

Here we have used the fact that, for $n \geq 3$, the expected total number of RPs $\leq n$ is

$$
E \Pi(n) = 1 + \sum_{k=3}^{n} \frac{1}{\log k} = \operatorname{li} n + O(1) \quad \text{as } n \to \infty.
$$

Thus we can use the above scale and mode as rescaling parameters $a_n$ and $b_n$, to obtain the standard Gumbel distribution $\exp(-e^{-z})$. Note that $1 \ll a_n \ll b_n$ as $n \to \infty$; so the distribution rescaling formula $z = \frac{x - b_n}{a_n}$ will produce approximately equal values of $z$, no matter whether we are rescaling the longest runs $R_n$ or maximal gaps $G_n \equiv R_n + 1$. □
4.3 Properties of the distribution of maximal gaps

Let us look at the properties of the distribution of maximal gaps. We can readily see that the exact distribution of maximal gaps is discrete and bounded, while the limiting Gumbel distribution is continuous, smooth, and unbounded. Below we discuss two properties that are common to the exact and asymptotic distributions of maximal gaps: log-concavity and unimodality.

4.3.1 Log-concavity

A function $f(x)$ is log-concave if $\log f(\lambda x + (1 - \lambda)y) \geq \lambda \log f(x) + (1 - \lambda) \log f(y)$ for all $0 \leq \lambda \leq 1$ and for all $x, y$. A discrete sequence $\{s_k\}$ is log-concave if $s_k^2 \geq s_{k-1}s_{k+1}$ for each middle term $s_k$.

**Log-concavity of the limiting distribution of maximal gaps.** It is well known (and easy to check given the formulas for the distribution’s pdf and cdf) that the Gumbel distribution pdf and cdf are log-concave. Note that, in general, if $f(x)$ is a continuous distribution pdf and $F(x)$ is the corresponding cdf, then the following implications are true:

$$f(x) \text{ is log-concave } \Rightarrow F(x) \text{ is log-concave},$$

$$f(x) \text{ is log-concave } \iff 1 - F(x) \text{ is log-concave}.$$  

**Log-concavity of exact distributions of maximal gaps.** A direct computational check shows that all the exact distribution functions we have computed in Section 4.1 are also log-concave. (However, as expected, the log-concavity is not preserved if we use a faster Monte-Carlo algorithm rather than the exact formula for computing the finite distributions.)

4.3.2 Unimodality

The Gumbel distribution $Gumbel(x; a, \mu)$ is unimodal for any $\mu$ and any $a > 0$: it has a unique mode (most probable value), namely, $\mu$. What about the exact distribution of maximal gaps between RPs in Cramér’s model with $n$ urns? By Theorem 1, the Gumbel distribution is the limiting distribution of maximal gaps between RPs in Cramér’s model; therefore it is reasonable to expect that, for large $n$, in Cramér’s model with $n$ urns the exact distribution of maximal gaps between RPs is also unimodal. On the other hand, for moderate values of $n$ one can check by direct computation that each of the exact distributions of maximal gaps between RPs is unimodal. Here we will state without a formal proof the following

**Unimodality Lemma.** In Cramér’s model with $n$ urns, the exact distribution of maximal gaps between RPs is unimodal for each $n > 1$. 

10
5 Appendix: Maximal intervals between random events occurring with slowly varying probability $\ell(t) \rightarrow 0$

In this appendix we give a general theorem whose particular case ($\ell(t) = 1/\log t$ for $t \geq 3$) has been used in Section 4.2.2. We say that a function $\ell(t) > 0$ is slowly varying if it is defined for positive $t$, and $\lim_{t \to \infty} \frac{\ell(\lambda t)}{\ell(t)} = 1$ for any fixed $\lambda > 0$ [10, p. 362].

**Theorem 2.** Consider biased coins with tails probability $\ell(k)$ at the $k$-th toss, $0 < \ell(k) < 1$, where $\ell(t)$ is a smooth, slowly varying, monotonically decreasing function, and $\lim_{t \to \infty} \ell(t) = 0$.

Then, after a large number $n$ of tosses, the asymptotic distribution of the longest runs of heads $R_n$ is the Gumbel distribution: there exist $a_n > 0$ and $b_n$ such that

$$\lim_{n \to \infty} P(R_n \leq x \equiv a_n z + b_n) = \exp(-e^{-z}), \quad \text{where} \quad a_n \sim \frac{n}{E\Pi(n)}, \quad b_n \sim \frac{n \log E\Pi(n)}{E\Pi(n)}.$$  

Here $E\Pi(n)$ is the mathematical expectation of the total number of tails $\Pi(n)$ observed during the first $n$ tosses:

$$E\Pi(n) = \sum_{k=1}^{n} \ell(k).$$

It may be surprising that the asymptotic distribution does exist at all for discrete events occurring with a slowly varying probability $\ell(t) \rightarrow 0$. In contrast, for a biased coin with a constant positive probability of tails, the limiting distribution of the longest run of heads does not exist [20, p. 203].

In the lemmas below, we assume the conditions of Theorem 2 and take $\lambda$ to be arbitrarily large ($\lambda \gg 1$). The lemmas are easy to prove using the theory of regularly varying functions [10, p. 361–367].

**Lemma 3.** $\lim_{t \to \infty} \Pi(t) = \infty$ a.s., while $\lim_{t \to \infty} \frac{E\Pi(t)}{t} = 0$.

**Lemma 4.** $\lim_{t \to \infty} E(\Pi(\lambda t) - \Pi(t)) = \infty$.

**Lemma 5.** $\lim_{t \to \infty} (E\Pi(\lambda t)/E\Pi(t)) = \lambda$.

**Lemma 6.** $\lim_{t \to \infty} \frac{E(\Pi(\lambda t) - \Pi(t))}{E\Pi(\lambda t)} = \frac{\lambda - 1}{\lambda}$.

**Proof of Theorem 2.** Let $k$ denote the ordinal number of a coin toss, so the $k$-th toss has the tails probability $\ell(k)$. Take $\lambda \gg 1$ and consider the sequence $S(t, \lambda t)$ of consecutive tosses with $t < k < \lambda t$. For toss sequences $S(t, \lambda t)$ with larger and larger $t$, we see that:

(i) $\ell(k)$ becomes nearly constant: $\ell(\lambda t) \lesssim \ell(k) \lesssim \ell(t)$ (by monotonicity + slow variation);
(ii) the expected total number of tails is $E(\Pi(\lambda t) - \Pi(t)) \to \infty$ as $t \to \infty$ (by Lemma 4).
As $t$ grows larger, the outcome of the toss sequence $S(t, \lambda t)$ becomes indistinguishable from an equally long sequence of tosses of a constant-bias coin whose tails probability $q$ is

$$q = \frac{E(\Pi(\lambda t) - \Pi(t))}{\lambda t - t} \approx \frac{E\Pi(\lambda t)}{\lambda t} \approx \frac{E\Pi(n)}{n},$$

where we set $n = \lfloor \lambda t \rfloor$.

In the latter setup with a constant-bias coin, head runs are modeled by i.i.d. geometric random variables [20]. The expected total number $m$ of head runs is

$$m = q(\lambda t - t) = E(\Pi(\lambda t) - \Pi(t)) \approx \frac{\lambda - 1}{\lambda} E\Pi(n), \quad \text{with } n = \lfloor \lambda t \rfloor, \quad (1)$$

In turn, these $m$ geometric i.i.d. variables are highly accurately approximated by exponential i.i.d. variables [1] whose common cdf is $\text{Exp}(\xi; a) = 1 - e^{-\xi/a}$, with the mean $a$ given by

$$a = \frac{\lambda t - t}{E(\Pi(\lambda t) - \Pi(t))} \approx \frac{n}{E\Pi(n)}, \quad \text{(here and above we have used Lemma 6).}$$

The largest value $L$ of $m$ exponential i.i.d. random variables, with the cdf $1 - e^{-\xi/a}$, has the limiting Gumbel distribution [6, 8, 11]:

$$\lim_{m \to \infty} P(L \leq az + a \log m) = \exp(-e^{-z}).$$

However, it is well known that for the geometrically distributed runs of heads with a constant-bias coin (i.e. the tails probability $q$ = const) the longest runs $R$ do not have a limiting distribution [1, 20]:

$$P(R \leq az + a \log m) = \exp(-e^{-z}) + O(q), \quad (2)$$

with the geometric-to-exponential approximation error $O(q)$ preventing the convergence of the exact longest run distributions to the Gumbel distribution. Nevertheless, in our original setup with slowly varying bias of the coin, instead of a constant $q$ we have

$$q_n \sim a_n^{-1} \sim n E\Pi(n) \to 0 \quad \text{as } n \to \infty. \quad (3)$$

For the sequence $S(t, \lambda t)$ to contain the absolute longest run of heads $R_n$ up to $n = \lfloor \lambda t \rfloor$, we must take $\lambda$ very large, so $\frac{\lambda - 1}{\lambda} \to 1$, and from equation (1) we have $\log m \approx \log E\Pi(n)$. Now eqs. (2) and (3) yield the limiting distribution of $R_n$ under the theorem conditions:

$$\lim_{n \to \infty} P(R_n \leq a_n z + b_n) = \exp(-e^{-z}), \quad \text{where } a_n \sim \frac{n}{E\Pi(n)}, \quad b_n \sim \frac{n \log E\Pi(n)}{E\Pi(n)}.$$ 

Thus the convergence to the limiting Gumbel distribution is restored. The convergence rate clearly depends on the actual choice of the slowly varying function $\ell(t) \to 0$. \hfill \Box 

**Remarks.** Theorem 2 is general enough to describe maximal gaps/runs for “primes” and “prime constellations” in Cramér’s model. However, comparing it to Mladenović’s result [13, Thm. 1], one might look for a further generalization. For example, what kind of a limiting distribution (if any) would we get if in Theorem 2 we replace the slowly varying probability $\ell(t) \to 0$ by some regularly varying probability $q(t) \to 0$, $q(t) \in \text{RV}_x$ with, say, $-1 \leq \alpha \leq 0$? Here $\alpha$ is the index of regular variation; for definitions of $\alpha$ and $\text{RV}_x$ see, e.g., [18, p. 362]. (Note that $\alpha = 0$ corresponds to slowly varying functions.)

12
References

[1] C. W. Anderson, Extreme value theory for a class of discrete distributions with applications to some stochastic processes. J. Appl. Prob. 7 (1970), 99–113.

[2] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja, Records, Wiley, 1998.

[3] C. Caldwell, The gaps between primes. Section 3: Table and graph of record gaps. Available at http://primes.utm.edu/notes/gaps.html#table

[4] H. Cramér, On the order of magnitude of the difference between consecutive prime numbers. Acta Arith. 2 (1936), 23–46.

[5] R. A. Fisher and L. H. C. Tippett, Limiting forms of the frequency distribution of the largest and smallest member of a sample, Math. Proc. Cambridge Philos. Soc. 24 (1928), 180–190.

[6] B. V. Gnedenko, Sur la distribution limite du terme maximum d’une série aléatoire. Ann. Math. 44 (1943), 423–453. English translation: On the limiting distribution of the maximum term in a random series, in Breakthroughs in Statistics, Volume 1: Foundations and Basic Theory. Springer, 1993, pp. 185–225.

[7] A. Granville, Harald Cramér and the distribution of prime numbers. Scand. Actuar. J. 1 (1995), 12–28.

[8] E. J. Gumbel, Les valeurs extrêmes des distributions statistiques. Annales de l’institut Henri Poincaré 5 (1935), 115–158.

[9] E. J. Gumbel, Statistics of Extremes, Columbia University Press, 1958. Dover, 2004.

[10] L. de Haan and A. Ferreira, Extreme Value Theory, Springer, 2006.

[11] W. J. Hall and J. A. Wellner, The rate of convergence in law for the maximum of an exponential sample. Statist. Neerlandica 33 (1979), 151–154.

[12] A. Kourbatov, Maximal gaps between prime k-tuples: a statistical approach. Journal of Integer Sequences, 16 (2013), Article 13.5.2. http://arxiv.org/abs/1301.2242

[13] P. Mladenović, Limit theorems for the maximum terms of a sequence of random variables with marginal geometric distributions. Extremes 2:4 (1999), 405–419.

[14] V. B. Nevzorov, Records: Mathematical Theory, AMS Translations Series 194 (2001).

[15] T. R. Nicely, First occurrence prime gaps, preprint, 2013. Available at http://www.trnicely.net/gaps/gaplist.html
[16] T. Oliveira e Silva, S. Herzog, and S. Pardi, Empirical verification of the even Goldbach conjecture, and computation of prime gaps, up to $4 \cdot 10^{18}$, Math. Comp. 83 (to appear).

[17] J. Pintz, Cramér vs Cramér: On Cramér’s probabilistic model of primes. Funct. Approx. Comment. Math. 37 (2007), 361–376.

[18] S. I. Resnick, Record values and maxima. The Annals of Probability 1 (1973), 650–662.

[19] L. Rodriguez and C. Rivera, Conjecture 66. Gaps between consecutive twin prime pairs, 2009. Available at http://www.primepuzzles.net/conjectures/conj066.htm.

[20] M. F. Schilling, The longest run of heads. College Math. J. 21 (1990), 196–207.

[21] D. Shanks, On maximal gaps between successive primes. Math. Comp. 18 (1964), 646–651.

[22] N. J. A. Sloane, The On-Line Encyclopedia of Integer Sequences, 2012. Available at http://oeis.org.

[23] M. Wolf, Some heuristics on the gaps between consecutive primes, preprint, 2011. Available at http://arxiv.org/abs/1102.0481.

2010 Mathematics Subject Classification: Primary 11N05; Secondary 60G70, 62E20.

Keywords: distribution of primes, prime gap, prime k-tuple, prime constellation, extreme value statistics, Cramér conjecture, Gumbel distribution, asymptotic distribution.

(Concerned with OEIS sequences A005250, A113274, A113404, A200503, A201051, A201062, A201073, A201251, A201596, A201598, A202281, A202361, A235402, A235492.)