The American Statistician

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/utas20

First-Passage Times for Random Partial Sums: Yadrenko's Model for *e* and Beyond

Joel E. Cohen

To cite this article: Joel E. Cohen (2024) First-Passage Times for Random Partial Sums: Yadrenko's Model for *e* and Beyond, The American Statistician, 78:1, 111-114, DOI: 10.1080/00031305.2023.2244542

To link to this article: https://doi.org/10.1080/00031305.2023.2244542

Published online: 10 Aug 2023.

THE AMERICAN STATISTICIAN
2024, VOL. 78, NO. 1, 111–114: Short Technical Note
https://doi.org/10.1080/00031305.2023.2244542

Taylor & Francis
Taylor & Francis Group

SHORT TECHNICAL NOTE

First-Passage Times for Random Partial Sums: Yadrenko's Model for *e* and Beyond

Joel E. Cohen[a,b,c]

[a]The Rockefeller University, New York, NY; [b]Columbia University, New York, NY; [c]University of Chicago, Chicago, IL

ABSTRACT
M. I. Yadrenko discovered that the expectation of the minimum number N_1 of independent and identically distributed uniform random variables on (0, 1) that have to be added to exceed 1 is e. For any threshold $a > 0$, K. G. Russell found the distribution, mean, and variance of the minimum number N_a of independent and identically distributed uniform random summands required to exceed a. Here we calculate the distribution and moments of N_a when the summands obey the negative exponential and Lévy distributions. The Lévy distribution has infinite mean. We compare these results with the results of Yadrenko and Russell for uniform random summands to see how the expected first-passage time $E(N_a), a > 0$, and other moments of N_a depend on the distribution of the summand.

ARTICLE HISTORY
Received May 2023
Accepted July 2023

KEYWORDS
Exponential distribution; Lévy distribution; Life length; M. I. Yadrenko; Napier's number *e*; Negative exponential distribution; Reliability theory; Uniform distribution

1. Introduction

Gnedenko (1978, p. 194, Ex. 22) stated a beautiful problem that he attributed to Myhailo I. Yadrenko: Let $U_1, U_2, \ldots$ be independent random variables uniformly distributed on $(0, 1)$. Let $S_n := U_1 + \cdots + U_n$ and let N_1 be the least n such that $S_n > 1$. Prove that $E(N_1) = e$.

Russell (1983) proved this result and extended it by finding, for any threshold $a > 0$, the distribution of the minimum value N_a of n such that $S_n > a$. We think of the random variable N_a as a first-passage time, the first time n for S_n to pass the threshold a. N_a has varied interpretations in probability theory and applications outside of probability theory (Russell 1983, pp. 175–177), including length of life in reliability theory and mathematical demography.

Russell found expressions "that are not readily simplified" in general for the mean and variance of N_a, but that simplify nicely in some special cases. For example, $E(N_1) = e$, as Yadrenko observed, and $Var(N_1) = e(3 - e)$; $E(N_2) = e^2 - e$, $Var(N_2) = -e^4 + 2e^3 + 4e^2 - 5e$. Russell (1991) used the simulation of e by Yadrenko's model to illustrate the power of the method of antithetic variables to estimate e with reduced variance.

Here we calculate the distribution and moments of N_a when the summands obey the negative exponential (Feller 1971, pp. 1, 8–21) and Lévy (Feller 1971, p. 173) distributions. We compare these results with the known results (Gnedenko 1978; Russell 1983) when the summands are uniformly distributed on (0,1). The point of the comparison is to see how the expected first-passage time $E(N_a), a > 0$, and other moments of N_a depend on the distribution of the summand X_i in $S_n := \sum_{i=1}^{n} X_i$. When each summand X_i is uniformly distributed on $(0, 1)$, the variance/mean^2 (which is the squared coefficient of variation) of X_i is $(1/12)/(1/2)^2 = 1/3 < 1$. By contrast, the variance/mean^2 of the negative exponential distribution is always 1, regardless of the distribution's mean. The Lévy distribution (Feller 1971, p. 173) has infinite mean and infinite variance, so its variance/mean^2 is not defined, but its (sample variance)/(sample mean)^2 becomes arbitrarily large as the sample size increases. We shall see how these dramatic differences in (sample variance)/(sample mean)^2 and in the shapes of the uniform, negative exponential, and Lévy distributions affect the first-passage time N_a.

Section 2 states the results in three theorems: when the summand X_i in $S_n := \sum_{i=1}^{n} X_i$ is a nonnegative random variable with a general continuous probability density function; when it has a negative exponential distribution; and when it has the Lévy distribution. Section 3 proves the three theorems. Section 4 discusses the results and sketches the application of the same methods to any strictly stable distribution with tail index $\alpha \in (0, 1)$ (Feller 1971, p. 170).

2. Results

Let $X \overset{d}{=} Y$ mean that the cumulative distribution function of the random variable X equals the cumulative distribution function of the random variable Y. Let $\mathbb{N} := \{1, 2, \ldots\}$ be the set of positive integers or natural numbers. Let $a > 0$. Let $X, X_i, i \in \mathbb{N}$ be independent and identically distributed with continuous probability density function f and cumulative distribution function F such that $\Pr(X \in [0, \infty)) = 1$, $\Pr(X = 0) < 1$. For $n \in \mathbb{N}$, define $S_n := X_1 + \cdots + X_n$. Define $S_0 := 0$ with probability 1, so that $1 = \Pr(S_0 < a)$. Define $N_a := \min\{n \in \mathbb{N} \mid S_n > a\}$. N_a is the first-passage time to exceed threshold a.

CONTACT Joel E. Cohen cohen@rockefeller.edu The Rockefeller University, 1230 York Avenue, Box 20, New York, NY 10065.

Theorem 1. With the above background and definitions, the probability density function of N_a is related to the cumulative distribution function of S_n by

$$\Pr(N_a = n) = \Pr(S_{n-1} < a) - \Pr(S_n < a), \qquad n \in \mathbb{N}. \quad (1)$$

For $k \in \mathbb{N}$, the moments $E(N_a^k)$ are related to the cumulative distribution function $\Pr(S_n < a)$ of S_n by

$$E(N_a^k) = \sum_{n=0}^{\infty} [(n+1)^k - n^k] \Pr(S_n < a). \quad (2)$$

Consequently $E(N_a) = 1 + \sum_{n=1}^{\infty} \Pr(S_n < a)$ and $E(N_a^2) = 1 + \sum_{n=1}^{\infty} (2n+1) \Pr(S_n < a)$ and $\mathrm{var}(N_a^2) = E(N_a^2) - [E(N_a)]^2$.

For any $\lambda > 0$, let $\mathrm{Exp}(\lambda)$ be a real-valued random variable with probability density function $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$, and $f(x) = 0$, $x < 0$. Such a random variable is said to be negative exponential (or simply exponential) with rate parameter λ. The cumulative distribution function of $\mathrm{Exp}(\lambda)$ is $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$. $\mathrm{Exp}(\lambda)$ has expectation $1/\lambda$ and variance $1/\lambda^2$. Let $X_i, i \in \mathbb{N}$, be independent and identically distributed random variables $\stackrel{d}{=} \mathrm{Exp}(\lambda)$, $\lambda > 0$. For $n \in \mathbb{N}$, the probability density function of S_n is $f(x; n, \lambda) = x^{n-1}\lambda^n e^{-\lambda x}/(n-1)!$. The cumulative distribution function of S_n is

$$F(x; n, \lambda) = \int_0^x f(x; n, \lambda) dx = 1 - e^{-\lambda x} \sum_{i=0}^{n-1} \frac{(\lambda x)^i}{i!}. \quad (3)$$

S_n has the Erlang distribution, which is the special case of the gamma distribution given by a sum of n independent copies of $\mathrm{Exp}(\lambda)$. Thus $f(x; 1, \lambda) = \lambda e^{-\lambda x}$, the probability density function of $\mathrm{Exp}(\lambda)$, and $F(x; 1, \lambda) = 1 - e^{-\lambda x}$, the cumulative distribution function of $\mathrm{Exp}(\lambda)$.

Theorem 2. With the above background and definitions, for any $\lambda > 0$, $a > 0$, $n \in \mathbb{N}$,

$$\begin{aligned} \Pr(N_a = n) &= e^{-\lambda a} \frac{(\lambda a)^{n-1}}{(n-1)!}, \\ E(N_a) &= 1 + \lambda a, \\ E(N_a^2) &= (\lambda a)^2 + 3\lambda a + 1, \\ \mathrm{var}(N_a) &= \lambda a, \\ E(N_a^3) &= (\lambda a)^3 + 6(\lambda a)^2 + 7(\lambda a) + 1, \\ E(N_a^4) &= (\lambda a)^4 + 10(\lambda a)^3 + 25(\lambda a)^2 + 15(\lambda a) + 1. \end{aligned} \quad (4)$$

Let Φ be the cumulative distribution function of the standard normal distribution $\mathcal{N}(0,1)$ with mean 0 and variance 1. The Lévy distribution with scale parameter $c > 0$, which we denote Lévy(c), has cumulative distribution function $F(x) = 2(1 - \Phi(\sqrt{c/x}))$ and probability density function $f(x) = \sqrt{c/(2\pi)} \exp[-c/(2x)]x^{-3/2}$ for $x \geq 0$. Both the cumulative distribution function and the probability density function equal zero if $x < 0$. As these formulas indicate, Lévy(c) $\stackrel{d}{=} c/[\mathcal{N}(0,1)]^2$. The population mean and the population variance of Lévy(c) are infinite. If X and $X_i, i \in \mathbb{N}$, are independent and identically distributed random variables $\stackrel{d}{=}$ Lévy(c), $c > 0$, then for all $n \in \mathbb{N}$, $S_n := X_1 + \cdots + X_n \stackrel{d}{=} n^2 X$.

Theorem 3. With the above background and definitions, for all $k, n \in \mathbb{N}$, $a > 0$,

$$\Pr(S_n < a) = 2[1 - \Phi(n\sqrt{c/a})], \quad (5)$$

$$E(N_a^k) = 2 \sum_{n=0}^{\infty} [(n+1)^k - n^k][1 - \Phi(n\sqrt{c/a})]. \quad (6)$$

Consequently $E(N_a) = 2\sum_{n=0}^{\infty}[1 - \Phi(n\sqrt{c/a})]$ and $E(N_a^2) = 2\sum_{n=0}^{\infty}(2n+1)[1 - \Phi(n\sqrt{c/a})]$ and $\mathrm{var}(N_a) = E(N_a^2) - [E(N_a)]^2$.

3. Proofs

Proof of Theorem 1. Because f is continuous, $\Pr(S_n < a) + \Pr(S_n > a) = 1$. Therefore, for all $n \in \mathbb{N}$,

$$\begin{aligned} \Pr(S_{n-1} < a) &= \Pr(S_{n-1} < a \cap S_n < a) \\ &\quad + \Pr(S_{n-1} < a \cap S_n > a) \\ &= \Pr(S_n < a) + \Pr(N_a = n), \end{aligned}$$

which can be rearranged to become (1).

Next, for any $k \in \mathbb{N}$, using the definition of the kth moment for the first equality, and (1) for the third equality, and $1 = \Pr(S_0 < a)$ for the fourth equality,

$$\begin{aligned} E(N_a^k) &:= \sum_{n=1}^{\infty} n^k \Pr(N_a = n) \\ &= 1^k \cdot \Pr(N_a = 1) + 2^k \cdot \Pr(N_a = 2) \\ &\quad + 3^k \cdot \Pr(N_a = 3) + \cdots \\ &= \Pr(S_0 < a) - \Pr(S_1 < a) \\ &\quad + 2^k[\Pr(S_1 < a) - \Pr(S_2 < a)] \\ &\quad + 3^k[\Pr(S_2 < a) - \Pr(S_3 < a)] + \cdots \\ &= 1 + \sum_{n=1}^{\infty} [(n+1)^k - n^k] \Pr(S_n < a) \\ &= \sum_{n=0}^{\infty} [(n+1)^k - n^k] \Pr(S_n < a) \end{aligned} \quad (7)$$

since $0^k = 0$ if $k \in \mathbb{N}$. The expression for $E(N_a^2)$ uses $(n+1)^2 - n^2 = 2n + 1$, $n \in \mathbb{N}$. □

Proof of Theorem 2. From (1) and (3),

$$\begin{aligned} \Pr(N_a = n) &= \Pr(S_{n-1} < a) - \Pr(S_n < a) \\ &= F(a; n-1, \lambda) - F(a; n, \lambda) \\ &= e^{-\lambda a} \left(\sum_{i=0}^{n-1} \frac{(\lambda a)^i}{i!} - \sum_{i=0}^{n-2} \frac{(\lambda a)^i}{i!} \right) \\ &= e^{-\lambda a} \frac{(\lambda a)^{n-1}}{(n-1)!}. \end{aligned}$$

Hence

$$
\begin{aligned}
E(N_a) &:= \sum_{k=0}^{\infty} k \cdot \Pr(N_a = k) \\
&= 0 + e^{-\lambda a} + e^{-\lambda a} \sum_{k=2}^{\infty} k \cdot \frac{(\lambda a)^{k-1}}{(k-1)!} \\
&= e^{-\lambda a} \sum_{k=0}^{\infty} k \cdot \frac{(\lambda a)^{k-1}}{(k-1)!} = e^{-\lambda a} \sum_{j=0}^{\infty} (j+1) \cdot \frac{(\lambda a)^j}{j!} \\
&= e^{-\lambda a} \left(\sum_{j=0}^{\infty} j \cdot \frac{(\lambda a)^j}{j!} + \sum_{j=0}^{\infty} \frac{(\lambda a)^j}{j!} \right) \\
&= e^{-\lambda a} (\lambda a e^{\lambda a} + e^{\lambda a}) \\
&= 1 + \lambda a.
\end{aligned}
$$

Further,

$$
\begin{aligned}
E(N_a^2) &:= \sum_{k=0}^{\infty} k^2 \cdot \Pr(N_a = k) = e^{-\lambda a} \sum_{k=0}^{\infty} k^2 \cdot \frac{(\lambda a)^{k-1}}{(k-1)!} \\
&= (\lambda a)^2 + 3\lambda a + 1, \\
\mathrm{var}(N_a) &= E(N_a^2) - [E(N_a)]^2 = \lambda a.
\end{aligned}
$$

The proofs for the third and fourth raw moments require lengthy algebra, which we omit. □

Proof of Theorem 3. Let X and $X_i, i \in \mathbb{N}$, be independent and identically distributed random variables $\stackrel{d}{=}$ Lévy(c), $c > 0$. The Lévy distribution satisfies $S_n := X_1 + \cdots + X_n \stackrel{d}{=} n^2 X$. Hence $\Pr(S_n < a) = \Pr(n^2 X < a) = \Pr(X < an^{-2}) = F(an^{-2}) = 2[1 - \Phi(n\sqrt{c/a})]$, which is (5). Hence for any $k \in \mathbb{N}$, (2) specializes to (6). For $k = 1, 2$, (6) simplifies to $E(N_a) = 2\sum_{n=0}^{\infty}[1 - \Phi(n\sqrt{c/a})]$ and $E(N_a^2) = 2\sum_{n=0}^{\infty}[2n + 1][1 - \Phi(n\sqrt{c/a})]$. □

4. Discussion

4.1. Negative Exponential Summands

The result $E(N_a) = 1 + \lambda a$ in (4) has a nice intuitive interpretation. Since $\lambda a = a/E[\mathrm{Exp}(\lambda)]$, the mean minimum number of summands $E(N_a)$ required to exceed a is just one more than the number of exponential means contained in a. In this light, Yadrenko's result that $E(N_1) = e$ for summands uniformly distributed on (0, 1) is surprising, because the threshold in this case $a = 1$ is exactly two uniform means (of 1/2 each), and one more than two would be three, whereas in fact only $e < 3$ uniform summands are required, on average, to exceed 1.

The moments of N_a can be written in terms of hypergeometric functions $F(a; b; z)$, using the notation of Abramowitz and Stegun (1964, chap. 15), as

$$
\begin{aligned}
E(N_a) &= e^{-\lambda a} \cdot F(2; 1; \lambda a), \\
E(N_a^2) &= e^{-\lambda a} \cdot F(2\ 2; 1\ 1; \lambda a), \\
E(N_a^3) &= e^{-\lambda a} \cdot F(2,\ 2,\ 2; 1,\ 1,\ 1; \lambda a), \\
E(N_a^4) &= e^{-\lambda a} \cdot F(2,\ 2,\ 2,\ 2; 1,\ 1,\ 1,\ 1; \lambda a).
\end{aligned}
$$

For further comparison with the results of Yadrenko and Russell (1983), set $a = 1$. Since the summands U_i uniformly distributed on $(0, 1)$ have mean 1/2, suppose the exponentially distributed summands also have mean $1/\lambda = 1/2$ or $\lambda = 2$. Then for uniformly distributed summands, $E(N_1) = e$, $\mathrm{var}(N_1) = e(3 - e) \approx 0.7658$ as noted above, while for exponentially distributed summands with the same mean, $E(N_1) = 3 > e$, $\mathrm{var}(N_1) = 2 > e(3 - e)$. With $a = 1$, to get $E(N_1) = e$ for exponential summands, we must have $\lambda = e - 1$ and hence the mean $1/\lambda$ of the exponential summands must be $1/(e - 1) \approx 0.5820 > 1/2$.

4.2. Lévy Summands

We approximated numerically $E(N_a)$ and $\mathrm{var}(N_a)$ for each combination (a, c) of threshold $a = 1, 2, 3$ and scale parameter $c = 1/16, 1/8, 1/4, 1/2, 1, 2, 4, 8$. In evaluating numerically the summations on the right side of (6), we dropped all terms with $n > 12$ because $1 - \Phi(8) \approx 6.6613 \cdot 10^{-16}$ and the finite accuracy of the software (The Mathworks, Inc. 2023) returns $1 - \Phi(n) = 0$, $n = 9, 10, 11, 12$.

For comparison with these numerical approximations, for each combination (a, c), we simulated 10^6 independent and identically distributed copies of Lévy(c) by means of $c/[\mathcal{N}(0, 1)]^2$ and summed successive terms until the sum exceeded a, recording the number N_a of terms in the sum; then we started the sum again with the next simulated value. From the recorded values of N_a, we computed the sample mean $\bar{N}_a$ of N_a and the sample variance $s^2(N_a)$ of N_a. Apart from a possible small effect at the end of the simulation if there were not enough summands for the last sum to exceed a, the number of simulations of N_a produced by this procedure equaled $10^6/\bar{N}_a$ plus or minus sampling variation.

For the base case $a = c = 1$, in these simulations, $\bar{N}_1 = 1.3650$. Numerically, from (6), $E(N_1) = 1.3656$, different by less than 0.001. Not surprisingly, both estimates of the mean are smaller than $E(N_1) = e$ in Yadrenko's case of uniform summands. The number of simulated values of N_1 was 732,625, and $10^6/\bar{N}_1 = 10^6/1.3650 = 732,601$. While $s^2(N_1) = 0.33378$ from the simulations, $\mathrm{var}(N_1) = 0.33411$ numerically, also different by less than 0.001. Both estimates of the variance are smaller than the variance $\mathrm{var}(N_1) = e(3 - e) \approx 0.7658$ of the first-passage times of uniform summands. The greater c is, the smaller the expectation $E(N_1)$ is for Lévy(c) summands. According to the numerical calculations, for Lévy(1/8), $E(N_1) = 2.7803 > e > E(N_1) = 2.1292$ for Lévy(1/4).

4.3. Strictly Stable Summands

The strictly stable distributions with tail index $\alpha \in (0, 1)$ (Feller 1971; Nolan 2020) include the Lévy(1) distribution as the special case when $\alpha = 1/2$. Let X and $X_i, i \in \mathbb{N}$, be independent and identically distributed strictly stable random variables with $\alpha \in (0, 1)$ and cumulative distribution function $F(\cdot)$. X has a continuous probability density function and takes values in $[0, \infty)$. Then, by definition of a strictly stable distribution, for all $n \in \mathbb{N}$, $S_n := X_1 + \cdots + X_n \stackrel{d}{=} n^{1/\alpha} X$. Therefore, $\Pr(S_n < a) = \Pr(n^{1/\alpha} X < a) = \Pr(X < an^{-1/\alpha}) = F(an^{-1/\alpha})$. Hence, for any $k \in \mathbb{N}$,

$$
E(N_a^k) = \sum_{n=0}^{\infty} [(n+1)^k - n^k] F(an^{-1/\alpha}). \tag{8}
$$

In particular, $E(N_a) = \sum_{n=0}^{\infty} F(an^{-1/\alpha})$ and $E(N_a^2) = \sum_{n=0}^{\infty}[2n + 1]F(an^{-1/\alpha})$ and $\mathrm{var}(N_a) = E(N_a^2) - [E(N_a)]^2$. Unfortunately, apart from the special case when $\alpha = 1/2$, no explicit expression for the cumulative distribution function $F(\cdot)$ is known, so (8) can be evaluated numerically but apparently not analytically.

Acknowledgments

The author thanks Roseanne Benjamin for help during this work; and a referee, the Associate Editor, and the Editor, for generous comments and constructive suggestions.

Disclosure Statement

The author reports there are no competing interests to declare.

References

Abramowitz, M., and Stegun, I. A. (1964), *Handbook of Mathematical Functions* (Vol. 256), Applied Mathematics Series 55. Washington, DC: U.S. National Bureau of Standards, Government Printing Office. [113]

Feller, W. (1971), *An Introduction to Probability Theory and Its Applications* (Vol. 2), New York: Wiley. [111,113]

Gnedenko, B. V. (1978), *The Theory of Probability*, translated by George Yankovsky, Moscow: Mir Publishers. [111]

Nolan, J. P. (2020), *Univariate Stable Distributions: Models for Heavy Tailed Data*, Cham: Springer. DOI:10.1007/978-3-030-52915-4 [113]

Russell, K. G. (1983), "On the Number of Uniform Random Variables Which Must Be Added To Exceed a Given Level," *Journal of Applied Probability*, 20, 172–177. [111,113]

——— (1991), "Estimating the Value of *e* by Simulation," *American Statistician*, 45, 66–68. [111]

The Mathworks, Inc. (2023), "Matlab, version R2023a," Natick, MA. Available at *https://www.mathworks.com/* [113]