

Deformation of power law in the double Pareto distribution using uniformly distributed observation time

Ken Yamamoto

Faculty of Science, University of the Ryukyus, Nishihara, Okinawa 903-0213, Japan

Takashi Bando, Hirokazu Yanagawa

*Production Engineering Department, SCM Division, Measurement Business Group,
ANRITSU CORPORATION, Atsugi, Kanagawa 243-8555, Japan*

Yoshihiro Yamazaki

School of Advanced Science and Engineering, Waseda University, Shinjuku, Tokyo 169-8555, Japan

The double Pareto distribution is a heavy-tailed distribution with a power-law tail, that is generated via geometric Brownian motion with an exponentially distributed observation time. In this study, we examine a modified model wherein the exponential distribution of the observation time is replaced with a continuous uniform distribution. The probability density, complementary cumulative distribution, and moments of this model are exactly calculated. Furthermore, the validity of the analytical calculations is discussed in comparison with numerical simulations of stochastic processes.

I. INTRODUCTION

Stochastic models have substantially contributed to the analysis of fluctuating or noisy systems [1, 2]. Furthermore, simple stochastic processes have been proved to adequately describe real phenomena, and the connection between stochastic processes and resultant probability distributions have been established theoretically. A probability distribution having a subexponential tail is referred to as a heavy-tailed distribution [3], which provides theoretical support for statistical physics such as critical phenomena [4], anomalous diffusion [5], and long-range memories [6]. Additionally, heavy-tailed distributions are crucial for complex systems such as social [7, 8] and biological [9] systems. Lognormal and power-law distributions are the heavy-tailed distributions focused upon in this study.

A random variable X is said to follow the lognormal distribution if the logarithm of X is normally distributed [10]. The probability density function (PDF) of the lognormal distribution is expressed as

$$f_{\text{LN}}(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma x}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right),$$

where μ and σ are the mean and standard deviation of $\ln X$, respectively. Its complementary cumulative distribution function (CCDF) is expressed as

$$F_{\text{LN}}(x; \mu, \sigma^2) = \int_x^\infty f_{\text{LN}}(y; \mu, \sigma^2) dy = \frac{1}{2} \operatorname{erfc}\left(\frac{\ln x - \mu}{\sqrt{2}\sigma}\right),$$

where

$$\operatorname{erfc}(z) := \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-u^2} du \quad (1)$$

is the complementary error function [11]. Furthermore, the k th moment of the lognormal random variable X is computed as

$$E[X^k] = \exp\left(k\mu + \frac{k^2}{2}\sigma^2\right). \quad (2)$$

Specifically, the mean and variance are

$$E[X] = \exp\left(\mu + \frac{\sigma^2}{2}\right), \quad V[X] = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1),$$

respectively.

The lognormal distribution occurs in the multiplicative stochastic process, which is given by

$$X_n = M_n X_{n-1}, \quad (3)$$

where the initial value X_0 is a positive constant, and M_1, M_2, \dots are random variables distributed independently and identically. The distribution of X_n for sufficiently large n can be approximated via a lognormal distribution, owing to the central limit theorem. The process (3) has been used as a simplified model for the X-ray burst [12] and the growth of organisms [13, 14]; this model was originally analyzed by Kolmogorov [15].

As Eq. (3) has a simple form, various additional effects have been applied. By modifying Eq. (3), the lognormal distribution can change qualitatively to other distributions. For example, power-law distributions are obtained by introducing additive noise [16], reset event [17], random stopping [18], and temporal cumulative sum [19, 20]. The introduction of a lower bound yields the power-law distribution [21], and a related model introducing the sample-dependent lower bound yields a heavy-tailed but not power-law distribution [22].

Geometric Brownian motion [23] can yield a lognormal distribution, and is expressed as the stochastic differential equation

$$dS_t = \mu S_t dt + \sigma S_t dB_t, \quad (4)$$

where B_t indicates the Brownian motion, μ is a real constant, and σ is a positive constant. For simplicity, we assume that the initial value S_0 is constant. By using Itô's lemma [23], the value of S_t at a given time $t = T$ is lognormally distributed, whose PDF is $f_{\text{LN}}(x; \ln S_0 + \tilde{\mu}T, \sigma^2 T)$, where $\tilde{\mu} = \mu - \sigma^2/2$. The geometric Brownian motion is applied to the Black-Scholes equation in mathematical finance [24].

When the observation time T of the geometric Brownian motion (4) is changed to an exponential random variable, S_T does not follow the lognormal distribution. Instead, the double Pareto (DP) distribution is observed [25, 26]. The PDF of S_T can be expressed as

$$f_{\text{DP}}(x) = \int_0^\infty \lambda e^{-\lambda t} f_{\text{LN}}(x; \ln S_0 + \tilde{\mu}t, \sigma^2 t) dt = \int_0^\infty \frac{\lambda e^{-\lambda t}}{\sqrt{2\pi t} \sigma x} \exp\left(-\frac{(\ln x - \ln S_0 - \tilde{\mu}t)^2}{2\sigma^2 t}\right) dt,$$

where $\lambda e^{-\lambda t}$ is the PDF of the exponential distribution with mean $1/\lambda$. This integral represents the mixture of the time-dependent lognormal size distribution with the exponential distribution as a weight.

To calculate this integral, $u = \sqrt{t}$ is introduced and the formula [11]

$$\int_0^\infty \exp\left(-a^2 u^2 - \frac{b^2}{u^2}\right) du = \frac{\sqrt{\pi}}{2a} e^{-2ab} \quad (5)$$

can be used. Finally, we obtain

$$f_{\text{DP}}(x) = \begin{cases} \frac{\lambda}{S_0 \sqrt{\tilde{\mu}^2 + 2\sigma^2 \lambda}} \left(\frac{x}{S_0}\right)^{-\alpha-1} & x \geq S_0, \\ \frac{\lambda}{S_0 \sqrt{\tilde{\mu}^2 + 2\sigma^2 \lambda}} \left(\frac{x}{S_0}\right)^{\beta-1} & x < S_0, \end{cases} \quad (6)$$

where

$$\alpha = \frac{-\tilde{\mu} + \sqrt{\tilde{\mu}^2 + 2\sigma^2 \lambda}}{\sigma^2}, \quad \beta = \frac{\tilde{\mu} + \sqrt{\tilde{\mu}^2 + 2\sigma^2 \lambda}}{\sigma^2}.$$

The term ‘‘double Pareto’’ originates from the property that $f_{\text{DP}}(x)$ has two different power-law exponents depending on whether $x \geq S_0$ or $x < S_0$. The DP distribution has been observed in various phenomena, such as income [27] and microblog posting interval [28].

A natural generalization of the DP distribution involves replacing the exponential distribution of T by other distributions. Upon replacing $\lambda e^{-\lambda t}$ with a general PDF $g(t)$, the PDF of S_T is formally expressed as

$$f_g(x) = \int_0^\infty g(t) f_{\text{LN}}(x; \ln S_0 + \tilde{\mu}t, \sigma^2 t) dt.$$

However, the calculation of this integral cannot be performed for general $g(t)$. This study focuses on the case wherein T follows a uniform distribution as a simple case. In this case, the PDF and CCDF of S_T can be exactly calculated although they have complicated forms. We further calculate the moments of S_T and compare the CCDF to the discrete-time process (3).

II. PROBABILITY DENSITY

In this section, we derive the PDF of the observed value S_T of geometric Brownian motion wherein the observation time T follows the uniform distribution on the interval $[0, T_{\max}]$:

$$g_{\text{uni}}(t) = \begin{cases} \frac{1}{T_{\max}} & 0 \leq t \leq T_{\max} \\ 0 & t > T_{\max}. \end{cases}$$

The PDF of S_T in this case is written as $f_{\text{uni}}(x)$ and is expressed as

$$\begin{aligned} f_{\text{uni}}(x) &= \int_0^\infty g_{\text{uni}}(t) f_{\text{LN}}(x; \ln S_0 + \tilde{\mu}t, \sigma^2 t) dt \\ &= \frac{1}{T_{\max}} \int_0^{T_{\max}} f_{\text{LN}}(x; \ln S_0 + \tilde{\mu}t, \sigma^2 t) dt \\ &= \frac{1}{T_{\max}} \int_0^{T_{\max}} \frac{1}{\sqrt{2\pi t} \sigma x} \exp\left(-\frac{(\ln x - \ln S_0 - \tilde{\mu}t)^2}{2\sigma^2 t}\right) dt. \end{aligned}$$

After changing the integral variable to $u = \sqrt{t/T_{\max}}$ and some manipulation, we obtain

$$f_{\text{uni}}(x) = \frac{2}{\sqrt{2\pi T_{\max}} \sigma x} \int_0^1 \exp\left(-\frac{(\ln x - \ln S_0 - T_{\max} \tilde{\mu} u^2)^2}{2T_{\max} \sigma^2 u^2}\right) du.$$

By introducing the scaling transformations $\mu_* = T_{\max} \mu$, $\sigma_* = \sqrt{T_{\max}} \sigma$, and $\tilde{\mu}_* = \mu_* - \sigma_*^2/2 = \tilde{\mu} T_{\max}$, the parameter T_{\max} can be eliminated:

$$\begin{aligned} f_{\text{uni}}(x) &= \frac{2}{\sqrt{2\pi} \sigma_* x} \int_0^1 \exp\left(-\frac{(\ln x - \ln S_0 - \tilde{\mu}_* u^2)^2}{2\sigma_*^2 u^2}\right) du \\ &= \frac{2}{\sqrt{2\pi} \sigma_* x} \left(\frac{x}{S_0}\right)^{\tilde{\mu}_*/\sigma_*^2} \int_0^1 \exp\left(-\frac{\tilde{\mu}_*^2}{2\sigma_*^2} u^2 - \frac{(\ln(x/S_0))^2}{2\sigma_*^2 u^2}\right) du. \end{aligned} \quad (7)$$

Therefore, although f_{uni} originally involves three parameters $\tilde{\mu}$, σ , and T_{\max} , it can be essentially reduced to two ($\tilde{\mu}_*$ and σ_*).

To calculate the integral in Eq. (7), we employed the following formula [11]:

$$\int_z^\infty \exp\left(-a^2 u^2 - \frac{b^2}{u^2}\right) du = \frac{\sqrt{\pi}}{4a} \left[e^{2ab} \operatorname{erfc}\left(az + \frac{b}{z}\right) + e^{-2ab} \operatorname{erfc}\left(az - \frac{b}{z}\right) \right], \quad (8)$$

where $a, b > 0$. Equation (8) can be considered as a generalization (or indefinite integral) of Eq. (5). In fact, Eq. (5) is obtained by considering the $z \rightarrow 0+$ limit and using the limit values $\operatorname{erfc}(+\infty) = 0$ and $\operatorname{erfc}(-\infty) = 2$. Using this relation, we obtain

$$\begin{aligned} \int_0^1 \exp\left(-a^2 u^2 - \frac{b^2}{u^2}\right) du &= \left(\int_0^\infty - \int_1^\infty\right) \exp\left(-a^2 u^2 - \frac{b^2}{u^2}\right) du = \frac{\sqrt{\pi}}{4a} [e^{-2ab}(2 - \operatorname{erfc}(a-b)) - e^{2ab} \operatorname{erfc}(a+b)] \\ &= \frac{\sqrt{\pi}}{4a} [e^{-2ab} \operatorname{erfc}(b-a) - e^{2ab} \operatorname{erfc}(a+b)], \end{aligned} \quad (9)$$

where $\operatorname{erfc}(-z) = 2 - \operatorname{erfc}(z)$ (see Ref. [11]) is used in the final equality.

The integral in Eq. (7) for $\tilde{\mu}_* \neq 0$ can be calculated by inserting $a = |\tilde{\mu}_*|/(\sqrt{2}\sigma_*)$ and $b = |\ln(x/S_0)|/(\sqrt{2}\sigma_*)$ in Eq. (9). The factor $e^{\pm 2ab}$ is reduced to

$$e^{\pm 2ab} = \exp\left(\pm \frac{|\tilde{\mu}_*|}{\sigma_*^2} \left| \ln \frac{x}{S_0} \right| \right) = \left(\frac{x}{S_0}\right)^{\pm \operatorname{sgn}(x-S_0) |\tilde{\mu}_*|/\sigma_*^2},$$

where sgn is the signum function defined by

$$\operatorname{sgn}(x) = \begin{cases} +1 & x > 0, \\ 0 & x = 0, \\ -1 & x < 0. \end{cases}$$

Therefore,

$$f_{\text{uni}}(x) = \frac{1}{2|\tilde{\mu}_*|x} \left[\left(\frac{x}{S_0} \right)^{(1-\text{sgn}(x-S_0))|\tilde{\mu}_*|/\sigma_*^2} \text{erfc} \left(\frac{|\ln(x/S_0)| - |\tilde{\mu}_*|}{\sqrt{2}\sigma_*} \right) - \left(\frac{x}{S_0} \right)^{(1+\text{sgn}(x-S_0))|\tilde{\mu}_*|/\sigma_*^2} \text{erfc} \left(\frac{|\ln(x/S_0)| + |\tilde{\mu}_*|}{\sqrt{2}\sigma_*} \right) \right].$$

When $\tilde{\mu}_* > 0$, the absolute value can be simply removed: $|\tilde{\mu}_*| = \tilde{\mu}_*$. When $\tilde{\mu}_* < 0$, $|\tilde{\mu}_*| = -\tilde{\mu}_*$, and it can be confirmed in a straightforward manner that the expression of $f_{\text{uni}}(x)$ becomes the same as in $\tilde{\mu}_* > 0$. Thus, we finally obtain

$$f_{\text{uni}}(x) = \frac{1}{2\tilde{\mu}_*x} \left[\left(\frac{x}{S_0} \right)^{(1-\text{sgn}(x-S_0))\tilde{\mu}_*/\sigma_*^2} \text{erfc} \left(\frac{|\ln(x/S_0)| - \tilde{\mu}_*}{\sqrt{2}\sigma_*} \right) - \left(\frac{x}{S_0} \right)^{(1+\text{sgn}(x-S_0))\tilde{\mu}_*/\sigma_*^2} \text{erfc} \left(\frac{|\ln(x/S_0)| + \tilde{\mu}_*}{\sqrt{2}\sigma_*} \right) \right]. \quad (10)$$

The $f_{\text{uni}}(x)$ for $\tilde{\mu}_* = 0$ is written as

$$f_{\text{uni}}(x) = \frac{2}{\sqrt{2\pi}\sigma_*x} \int_0^1 \exp \left(-\frac{(\ln(x/S_0))^2}{2\sigma_*^2 u^2} \right) du = \frac{2}{\sqrt{2\pi}\sigma_*x} \int_0^1 \exp \left(-\frac{b^2}{u^2} \right) du,$$

where $b = \frac{|\ln(x/S_0)|}{\sqrt{2}\sigma_*}$, as above. By introducing a new integration variable $v = |b|/u$ and integrating by parts, we obtain

$$\begin{aligned} f_{\text{uni}}(x) &= \frac{2}{\sqrt{2\pi}\sigma_*x} \left[-\frac{|b|}{v} e^{-v^2} \right]_{|b|}^{\infty} - \frac{2}{\sqrt{2\pi}\sigma_*x} \int_{|b|}^{\infty} 2|b|e^{-v^2} dv = \frac{2e^{-b^2}}{\sqrt{2\pi}\sigma_*x} - \frac{\sqrt{2}|b|}{\sigma_*x} \text{erfc}(|b|) \\ &= \frac{2}{\sqrt{2\pi}\sigma_*x} \exp \left(-\frac{(\ln(x/S_0))^2}{2\sigma_*^2} \right) - \frac{|\ln(x/S_0)|}{\sigma_*^2 x} \text{erfc} \left(\frac{|\ln(x/S_0)|}{\sqrt{2}\sigma_*} \right). \end{aligned} \quad (11)$$

The derived PDFs in Eqs. (10) and (11) are complicated compared to the DP distribution (6).

We derive the asymptotic form of $f_{\text{uni}}(x)$ in the $x \rightarrow \infty$ and $x \rightarrow 0+$ limits. Using the asymptotic expansion of the erfc function [11]

$$\text{erfc}(z) \sim \frac{1}{\sqrt{\pi}z} e^{-z^2} \quad (z \rightarrow \infty) \quad (12)$$

and the relation

$$\left(\frac{x}{S_0} \right)^{2\tilde{\mu}_*/\sigma_*} \exp \left(-\frac{(\ln(x/S_0) + \tilde{\mu}_*)^2}{2\sigma_*^2} \right) = \exp \left(-\frac{(\ln(x/S_0) - \tilde{\mu}_*)^2}{2\sigma_*^2} \right),$$

we obtain

$$f_{\text{uni}}(x) \sim \sqrt{\frac{2}{\pi}} \frac{\sigma_*}{x(\ln x)^2} \exp \left(-\frac{(\ln(x/S_0) - \tilde{\mu}_*)^2}{2\sigma_*^2} \right) = \frac{2\sigma_*^2}{(\ln x)^2} f_{\text{LN}}(x; \ln S_0 + \tilde{\mu}_*, \sigma_*^2) \quad (x \rightarrow \infty, 0+).$$

Notably, both limits $x \rightarrow \infty$ and $x \rightarrow 0+$ have the same asymptotic form. Owing to the $(\ln x)^{-2}$ factor, $f_{\text{uni}}(x)$ decays slightly faster than the lognormal PDF.

A. Shape of $f_{\text{uni}}(x)$ graph

The graph of $f_{\text{uni}}(x)$ is shown in Fig. 1. For any $\tilde{\mu}_*$ and σ_* , $f_{\text{uni}}(x)$ is continuous for all $x > 0$ and is not differentiable at $x = S_0$, which corresponds to the discontinuity point of $\text{sgn}(x - S_0)$. Panel (a) shows the graph of fixed $\tilde{\mu}_* = 1$ and varying σ_* , and (b) shows the log-log graph of (a). Panel (c) shows the graph of fixed $\sigma_* = 1$ and varying $\tilde{\mu}_*$, and (d) shows the log-log graph of (c). The decay of $f_{\text{uni}}(x)$ as $x \rightarrow \infty$ is slower for larger σ_* and $\tilde{\mu}_*$. The graph

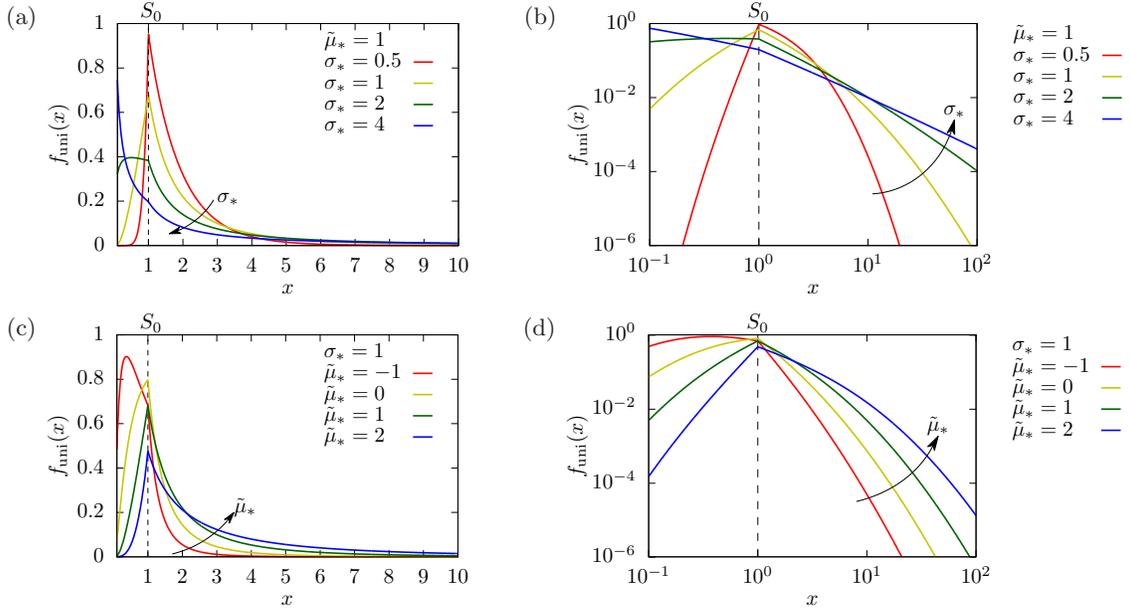


FIG. 1. Graphs of $f_{\text{uni}}(x)$. (a) $\tilde{\mu}_* = 1$ and $\sigma_* = 0.5, 1, 2$, and 4 . (b) Log-log graph of (a). (c) $\sigma_* = 1$ and $\tilde{\mu}_* = -1, 0, 1$, and 2 . (d) Log-log graph of (c). The vertical dashed line indicates $x = S_0 = 1$.

for $\tilde{\mu}_* = 1$ and $\sigma_* = 4$ in Fig. 1(a) appears to increase monotonically as $x \rightarrow 0+$; however, this is not true. As $\lim_{x \rightarrow 0+} f_{\text{uni}}(x) = 0$ for any $\tilde{\mu}_*$ and σ_* , this graph attains a maximum value for a very small (but positive) x .

In Fig. 1, the graph of $f_{\text{uni}}(x)$ has a peak at $x = S_0$ for $(\tilde{\mu}_*, \sigma_*) = (1, 0.5), (1, 1), (0, 1)$, and $(2, 1)$; the peak for the remaining graphs are not at $x = S_0$. We investigate the reason for this qualitative change, by deriving the condition for $(\tilde{\mu}_*, \sigma_*)$ such that $f_{\text{uni}}(x)$ has a peak at $x = S_0$. For sufficiently small $\delta > 0$,

$$f_{\text{uni}}(S_0 \pm \delta) \simeq f_{\text{uni}}(S_0) \pm f'_{\text{uni}}(S_0^\pm)\delta,$$

where

$$f'_{\text{uni}}(S_0^+) = \lim_{h \rightarrow 0+} \frac{f_{\text{uni}}(S_0 + h) - f_{\text{uni}}(S_0)}{h}, \quad f'_{\text{uni}}(S_0^-) = \lim_{h \rightarrow 0+} \frac{f_{\text{uni}}(S_0) - f_{\text{uni}}(S_0 - h)}{h}$$

are the right and left derivatives of $f_{\text{uni}}(x)$ at $x = S_0$, respectively. We have to consider one-sided derivatives because the function $f_{\text{uni}}(x)$ is not differentiable at $x = S_0$. The condition such that $f_{\text{uni}}(x)$ has a peak at $x = S_0$ is

$$f'_{\text{uni}}(S_0^+) < 0, \quad f'_{\text{uni}}(S_0^-) > 0.$$

The differentiation of Eqs. (10) and (11) yields

$$f'_{\text{uni}}(S_0^+) = \begin{cases} \frac{1}{\tilde{\mu}_* S_0^2} \left[\left(1 - \frac{\tilde{\mu}_*}{\sigma_*^2}\right) \operatorname{erfc}\left(\frac{\tilde{\mu}_*}{\sqrt{2}\sigma_*}\right) - 1 \right] & \tilde{\mu}_* \neq 0, \\ -\frac{2}{\sqrt{2\pi}\sigma_* S_0^2} & \tilde{\mu}_* = 0, \end{cases}$$

and

$$f'_{\text{uni}}(S_0^-) = \begin{cases} \frac{1}{\tilde{\mu}_* S_0^2} \left[\frac{2\tilde{\mu}_*}{\sigma_*^2} - 1 + \left(1 - \frac{\tilde{\mu}_*}{\sigma_*^2}\right) \operatorname{erfc}\left(\frac{\tilde{\mu}_*}{\sqrt{2}\sigma_*}\right) \right] & \tilde{\mu}_* \neq 0, \\ \frac{1}{\sigma_*^2 S_0^2} - \frac{2}{\sqrt{2\pi}\sigma_* S_0^2} & \tilde{\mu}_* = 0. \end{cases} \quad (13)$$

First, we prove that $f'_{\text{uni}}(S_0^+) < 0$ always holds. This is trivial for $\tilde{\mu}_* = 0$; thus, we examine the $\tilde{\mu}_* \neq 0$ case:

$$f'_{\text{uni}}(S_0^+) = -\frac{1}{S_0^2 \sigma_*^2} \operatorname{erfc}\left(\frac{\tilde{\mu}_*}{\sqrt{2}\sigma_*}\right) - \frac{1}{\tilde{\mu}_* S_0^2} \left[1 - \operatorname{erfc}\left(\frac{\tilde{\mu}_*}{\sqrt{2}\sigma_*}\right) \right].$$

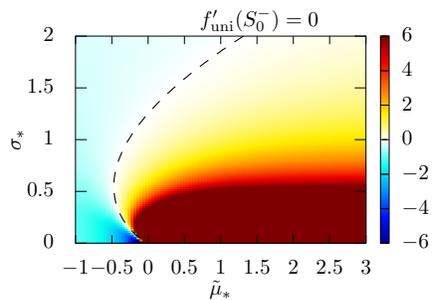


FIG. 2. Diagram of $f'_{\text{uni}}(S_0^-)$ on the $(\tilde{\mu}_*, \sigma_*)$ plane. The dashed curve represents the contour of $f'_{\text{uni}}(S_0^-) = 0$. The graph of $f_{\text{uni}}(x)$ yields a peak at $x = S_0$ if the point $(\tilde{\mu}_*, \sigma_*)$ lies on the right side of this curve.

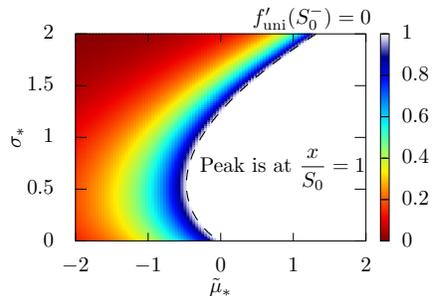


FIG. 3. Numerical result for the peak position depicted in the $(\tilde{\mu}_*, \sigma_*)$ plane. The color bar represents the peak position x divided by S_0 . The dashed curve shows the contour of $f'_{\text{uni}}(S_0^-) = 0$, which is shown in Fig. 2. To the right of this curve, $f_{\text{uni}}(x)$ shows a peak at $x = S_0$.

The first term on the right-hand side is always negative because the erfc function always takes a positive value. The factor $1 - \text{erfc}(\tilde{\mu}_*/(\sqrt{2}\sigma_*))$ in the second term becomes positive for positive $\tilde{\mu}_*$ and becomes negative for negative $\tilde{\mu}_*$. Thus, $1 - \text{erfc}(\tilde{\mu}_*/(\sqrt{2}\sigma_*))$ always has the same sign as $\tilde{\mu}_*$, and the second term is always negative.

In contrast, whether $f'_{\text{uni}}(S_0^-) > 0$ depends on $\tilde{\mu}_*$ and σ_* . Figure 2 shows the $f'_{\text{uni}}(S_0^-)$ values with $S_0 = 1$ for $-1 \leq \tilde{\mu}_* \leq 3$ and $0 < \sigma_* \leq 2$. To ensure clear visualization, we restrict the color bar within the interval $[-6, 6]$, although certain $(\tilde{\mu}_*, \sigma_*)$ yield values outside this range [e.g., $f'_{\text{uni}}(S_0^-) \approx 49$ for $(\tilde{\mu}_*, \sigma_*) = (1, 0.2)$]. The dashed curve in the figure represents the contour of $f'_{\text{uni}}(S_0^-) = 0$. Thus, $f_{\text{uni}}(x)$ exhibits a peak at $x = S_0$ when $(\tilde{\mu}_*, \sigma_*)$ is on the right side of this curve. By using the limit value $\text{erfc}(+\infty) = 0$, it can be proven using Eq. (13) that the curve for $f'_{\text{uni}}(S_0^-) = 0$ asymptotically draws the parabola $\tilde{\mu}_* = \sigma_*^2/2$ for $\sigma_* \gg 1$. When $\tilde{\mu}_* = 0$, $f'_{\text{uni}}(S_0^-) > 0$ can be exactly solved to attain $\sigma_* < \sqrt{\pi/2} \approx 1.25$. However, the solution for the $\tilde{\mu}_* \neq 0$ case will not be obtained exactly, owing to the erfc function.

Further analysis shows that $f_{\text{uni}}(x)$ is unimodal for any $\tilde{\mu}_*$ and σ_* (further details have been presented in Appendix A). Depending on the $(\tilde{\mu}_*, \sigma_*)$ values, the peak position of $f_{\text{uni}}(x)$ is either at $x = S_0$ (corresponding to $f'_{\text{uni}}(S_0^-) \geq 0$) or $x < S_0$ ($f'_{\text{uni}}(S_0^-) < 0$). The peak position x for the latter case is characterized by the solution of the equation $f'_{\text{uni}}(x) = 0$. The explicit form of this equation is presented in Eq. (A2) in Appendix A, which is a transcendental equation involving the erfc function. Rather than providing the exact solution of the peak position, we present a numerical result for the peak position in Fig. 3. The peak tends to be located at small x with decreasing $\tilde{\mu}_*$ and increasing σ_* .

III. COMPLEMENTARY CUMULATIVE DISTRIBUTION

In principle, the CCDF $F_{\text{uni}}(x) = \int_x^\infty f_{\text{uni}}(y)dy$ is derived by integrating the PDF $f_{\text{uni}}(x)$ in Eqs. (10) and (11). However, as the absolute value in the erfc function cannot be handled in a straightforward manner, we first remove the absolute value symbols.

When $x < S_0$ and $\tilde{\mu}_* \neq 0$,

$$f_{\text{uni}}(x) = \frac{1}{2\tilde{\mu}_*x} \left[\text{erfc} \left(\frac{\ln(x/S_0) - \tilde{\mu}_*}{\sqrt{2}\sigma_*} \right) - \left(\frac{x}{S_0} \right)^{2\tilde{\mu}_*/\sigma_*^2} \text{erfc} \left(\frac{\ln(x/S_0) + \tilde{\mu}_*}{\sqrt{2}\sigma_*} \right) + 2 \left(\left(\frac{x}{S_0} \right)^{2\tilde{\mu}_*/\sigma_*^2} - 1 \right) \right],$$

where the relation $\operatorname{erfc}(-z) = 2 - \operatorname{erfc}(z)$ is used. The difference from the case $x > S_0$ is the last term. Therefore, for any $x > 0$,

$$f_{\text{uni}}(x) = \frac{1}{2\tilde{\mu}_*} \left[\operatorname{erfc} \left(\frac{\ln(x/S_0) - \tilde{\mu}_*}{\sqrt{2}\sigma_*} \right) - \left(\frac{x}{S_0} \right)^{2\tilde{\mu}_*/\sigma_*^2} \operatorname{erfc} \left(\frac{\ln(x/S_0) + \tilde{\mu}_*}{\sqrt{2}\sigma_*} \right) + (1 - \operatorname{sgn}(x - S_0)) \left(\left(\frac{x}{S_0} \right)^{2\tilde{\mu}_*/\sigma_*^2} - 1 \right) \right]$$

is valid.

For the derivation of $F_{\text{uni}}(x)$, we calculate the following two integrals as a preliminary step:

$$I_1 = \int_x^\infty \operatorname{erfc} \left(\frac{\ln(y/S_0) - \tilde{\mu}_*}{\sqrt{2}\sigma_*} \right) \frac{dy}{y}, \quad I_2 = \int_x^\infty \left(\frac{y}{S_0} \right)^{2\tilde{\mu}_*/\sigma_*^2} \operatorname{erfc} \left(\frac{\ln(y/S_0) + \tilde{\mu}_*}{\sqrt{2}\sigma_*} \right) \frac{dy}{y}.$$

By setting $z = (\ln(y/S_0) - \tilde{\mu}_*)/(\sqrt{2}\sigma_*)$ and $z_- = (\ln(x/S_0) - \tilde{\mu}_*)/(\sqrt{2}\sigma_*)$,

$$\begin{aligned} I_1 &= \sqrt{2}\sigma_* \int_{z_-}^\infty \operatorname{erfc}(z) dz = \left[\sqrt{2}\sigma_* z \operatorname{erfc}(z) \right]_{z_-}^\infty + 2\sqrt{\frac{2}{\pi}}\sigma_* \int_{z_-}^\infty z e^{-z^2} dz \\ &= - \left(\ln \frac{x}{S_0} - \tilde{\mu}_* \right) \operatorname{erfc} \left(\frac{\ln(x/S_0) - \tilde{\mu}_*}{\sqrt{2}\sigma_*} \right) + \sqrt{\frac{2}{\pi}}\sigma_* \exp \left(-\frac{(\ln(x/S_0) - \tilde{\mu}_*)^2}{2\sigma_*^2} \right). \end{aligned}$$

We use integration by parts in the second equality. Similarly, by setting $z = (\ln(y/S_0) + \tilde{\mu}_*)/(\sqrt{2}\sigma_*)$ and $z_+ = (\ln(x/S_0) + \tilde{\mu}_*)/(\sqrt{2}\sigma_*)$,

$$\begin{aligned} I_2 &= \sqrt{2}\sigma_* e^{-2\tilde{\mu}_*/\sigma_*^2} \int_{z_+}^\infty \exp \left(\frac{2\sqrt{2}\tilde{\mu}_*}{\sigma_*} z \right) \operatorname{erfc}(z) dz \\ &= \left[\frac{\sigma_*^2}{2\tilde{\mu}_*} \exp \left(-\frac{2\tilde{\mu}_*^2}{\sigma_*^2} + \frac{2\sqrt{2}\tilde{\mu}_*}{\sigma_*} z \right) \operatorname{erfc}(z) \right]_{z_+}^\infty + \frac{\sigma_*^2}{\tilde{\mu}_*\sqrt{\pi}} e^{-2\tilde{\mu}_*/\sigma_*^2} \int_{z_+}^\infty \exp \left(-z^2 + \frac{2\sqrt{2}\tilde{\mu}_*}{\sigma_*} z \right) dz \\ &= -\frac{\sigma_*^2}{2\tilde{\mu}_*} \left(\frac{x}{S_0} \right)^{2\tilde{\mu}_*/\sigma_*^2} \operatorname{erfc} \left(\frac{\ln(x/S_0) + \tilde{\mu}_*}{\sqrt{2}\sigma_*} \right) + \frac{\sigma_*^2}{\sqrt{\pi}\tilde{\mu}_*} \int_{z_+ - \sqrt{2}\tilde{\mu}_*/\sigma_*}^\infty e^{-z^2} dz \\ &= -\frac{\sigma_*^2}{2\tilde{\mu}_*} \left(\frac{x}{S_0} \right)^{2\tilde{\mu}_*/\sigma_*^2} \operatorname{erfc} \left(\frac{\ln(x/S_0) + \tilde{\mu}_*}{\sqrt{2}\sigma_*} \right) + \frac{\sigma_*^2}{2\tilde{\mu}_*} \operatorname{erfc} \left(\frac{\ln(x/S_0) - \tilde{\mu}_*}{\sqrt{2}\sigma_*} \right). \end{aligned}$$

The CCDF for $x \geq S_0$ becomes

$$\begin{aligned} F_{\text{uni}}(x) &= \frac{1}{2\tilde{\mu}_*} \int_x^\infty \left[\operatorname{erfc} \left(\frac{\ln(y/S_0) - \tilde{\mu}_*}{\sqrt{2}\sigma_*} \right) - \left(\frac{y}{S_0} \right)^{2\tilde{\mu}_*/\sigma_*^2} \operatorname{erfc} \left(\frac{\ln(y/S_0) + \tilde{\mu}_*}{\sqrt{2}\sigma_*} \right) \right] \frac{dy}{y} \\ &= \frac{I_1 - I_2}{2\tilde{\mu}_*} \\ &= \frac{1}{2\tilde{\mu}_*} \left[\sqrt{\frac{2}{\pi}}\sigma_* \exp \left(-\frac{(\ln(x/S_0) - \tilde{\mu}_*)^2}{2\sigma_*^2} \right) + \frac{\sigma_*^2}{2\tilde{\mu}_*} \left(\frac{x}{S_0} \right)^{2\tilde{\mu}_*/\sigma_*^2} \operatorname{erfc} \left(\frac{\ln(x/S_0) + \tilde{\mu}_*}{\sqrt{2}\sigma_*} \right) \right. \\ &\quad \left. - \left(\ln \frac{x}{S_0} - \tilde{\mu}_* + \frac{\sigma_*^2}{2\tilde{\mu}_*} \right) \operatorname{erfc} \left(\frac{\ln(x/S_0) - \tilde{\mu}_*}{\sqrt{2}\sigma_*} \right) \right], \end{aligned}$$

and for $x < S_0$,

$$F_{\text{uni}}(x) = \frac{I_1 - I_2}{2\tilde{\mu}_*} + \frac{1}{\tilde{\mu}_*} \int_x^{S_0} \left(\left(\frac{y}{S_0} \right)^{2\tilde{\mu}_*/\sigma_*^2} - 1 \right) \frac{dy}{y}.$$

The remaining integral can be calculated as

$$\frac{1}{\tilde{\mu}_*} \int_x^{S_0} \left(\left(\frac{y}{S_0} \right)^{2\tilde{\mu}_*/\sigma_*^2} - 1 \right) \frac{dy}{y} = \frac{1}{\tilde{\mu}_*} \left(\frac{\sigma_*^2}{2\tilde{\mu}_*} - \frac{\sigma_*^2}{2\tilde{\mu}_*} \left(\frac{x}{S_0} \right)^{2\tilde{\mu}_*/\sigma_*^2} + \ln \frac{x}{S_0} \right).$$

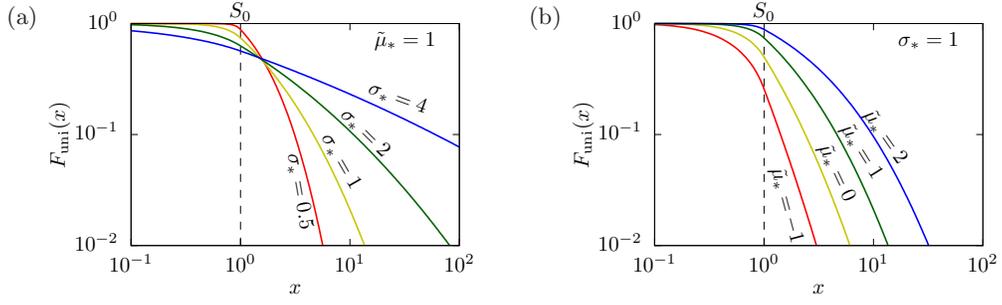


FIG. 4. Log-log graphs of $F_{\text{uni}}(x)$ for $\tilde{\mu}_* = 1$ and $\sigma_* = 0.5, 1, 2$, and 4 (a), and for $\sigma_* = 1$ and $\tilde{\mu}_* = -1, 0, 1$, and 2 (b). The vertical dashed line indicates $x = S_0 = 1$.

Consequently, we obtain

$$\begin{aligned}
 F_{\text{uni}}(x) = & \frac{1}{2\tilde{\mu}_*} \left[\sqrt{\frac{2}{\pi}} \sigma_* \exp\left(-\frac{(\ln(x/S_0) - \tilde{\mu}_*)^2}{2\sigma_*^2}\right) + \frac{\sigma_*^2}{2\tilde{\mu}_*} \left(\frac{x}{S_0}\right)^{2\tilde{\mu}_*/\sigma_*^2} \operatorname{erfc}\left(\frac{\ln(x/S_0) + \tilde{\mu}_*}{\sqrt{2}\sigma_*}\right) \right. \\
 & \left. - \left(\ln \frac{x}{S_0} - \tilde{\mu}_* + \frac{\sigma_*^2}{2\tilde{\mu}_*}\right) \operatorname{erfc}\left(\frac{\ln(x/S_0) - \tilde{\mu}_*}{\sqrt{2}\sigma_*}\right) \right] \\
 & + \frac{1 - \operatorname{sgn}(x - S_0)}{2\tilde{\mu}_*} \left(\frac{\sigma_*^2}{2\tilde{\mu}_*} - \frac{\sigma_*^2}{2\tilde{\mu}_*} \left(\frac{x}{S_0}\right)^{2\tilde{\mu}_*/\sigma_*^2} + \ln \frac{x}{S_0} \right), \tag{14}
 \end{aligned}$$

which is valid for all $x > 0$. This result is extremely complex, but it is exact.

By integrating Eq. (11) or taking the $\tilde{\mu}_* \rightarrow 0$ limit in Eq. (14), $F_{\text{uni}}(x)$ for $\tilde{\mu}_* = 0$ is obtained as

$$F_{\text{uni}}(x) = \frac{1}{2} \operatorname{erfc}\left(\frac{\ln(x/S_0)}{\sqrt{2}\sigma_*}\right) - \frac{1}{2\sqrt{2\pi}\sigma_*} \exp\left(-\frac{(\ln(x/S_0))^2}{2\sigma_*^2}\right) + \frac{1}{2\sigma_*^2} \left(\ln \frac{x}{S_0}\right)^2 \operatorname{erfc}\left(\frac{\ln(x/S_0)}{\sqrt{2}\sigma_*}\right).$$

Figure 4 shows the graph of $F_{\text{uni}}(x)$ on a log-log scale. The graph decays slower for larger σ_* [in (a)] and larger $\tilde{\mu}_*$ [in (b)]. The graphs do not exhibit significant variety in shape as $f_{\text{uni}}(x)$ in Fig. 2, and the overall shapes are similar for different $\tilde{\mu}_*$ and σ_* ,

IV. CALCULATION OF MOMENTS

The k th moment of S_T can be calculated as

$$\begin{aligned}
 E[S_T^k] &= \int_0^\infty x^k f_{\text{uni}}(x) dx \\
 &= \int_0^\infty x^k \frac{1}{T_{\text{max}}} \int_0^{T_{\text{max}}} f_{\text{LN}}(x; \ln S_0 + \tilde{\mu}t, \sigma^2 t) dt dx \\
 &= \frac{1}{T_{\text{max}}} \int_0^{T_{\text{max}}} \int_0^\infty x^k f_{\text{LN}}(x; \ln S_0 + \tilde{\mu}t, \sigma^2 t) dx dt \\
 &= \frac{S_0^k}{T_{\text{max}}} \int_0^{T_{\text{max}}} \exp\left(k\tilde{\mu}t + \frac{k^2}{2}\sigma^2 t\right) dt \\
 &= \frac{2}{k^2\sigma_*^2 + 2k\tilde{\mu}_*} \left[\exp\left(\frac{k^2\sigma_*^2}{2} + k\tilde{\mu}_*\right) - 1 \right] S_0^k.
 \end{aligned}$$

The moment of the lognormal distribution [Eq. (2)] is used in the fourth equality. In contrast to the DP distribution, $E[S_T^k]$ is not divergent for any k .

Specifically, the mean and variance of S_T for $\mu_* \neq 0$ and $\mu_* \neq -\sigma_*^2$ becomes

$$E[S_T] = \frac{e^{\mu_*} - 1}{\mu_*} S_0, \quad V[S_T] = \frac{e^{\sigma_*^2 + \mu_*} - 1}{\sigma_*^2 + \mu_*} S_0^2 - \frac{(e^{\mu_*} - 1)^2}{\mu_*^2} S_0^2.$$

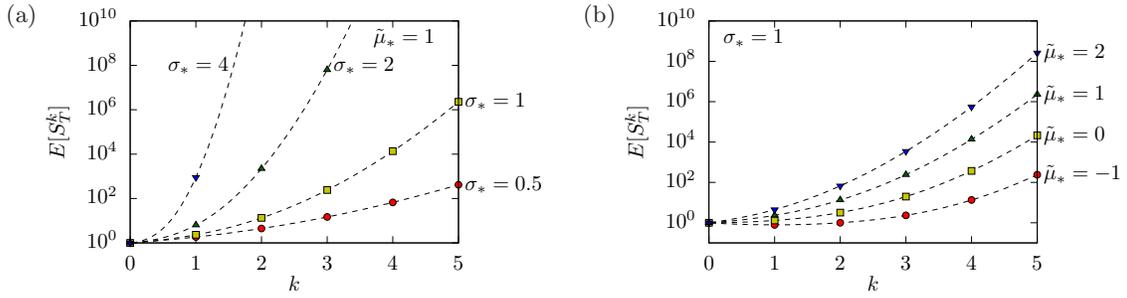


FIG. 5. The k th moment of S_T ($S_0 = 1$) up to $k = 5$ for $\tilde{\mu}_* = 1$ and $\sigma_* = 0.5, 1, 2$, and 4 (a), and for $\sigma_* = 1$ and $\tilde{\mu}_* = -1, 0, 1$, and 2 (b).

We obtain $E[S_T] = S_0$ when $\mu_* = 0$ and

$$V[S_T] = S_0^2 - \frac{(e^{\mu_*} - 1)^2}{\mu_*^2} S_0^2$$

when $\mu_* = -\sigma_*^2$.

Figure 5 shows the semi-log graphs of $E[S_T^k]$ for $S_0 = 1$ with different $\tilde{\mu}_*$ and σ_* . $E[S_T^k]$ becomes large for large $\tilde{\mu}_*$ and σ_* . The graph for $\tilde{\mu}_* = -1$ and $\sigma_* = 1$ in Fig. 5(b) is not monotonically increasing; $E[S_T^1] \approx 0.787 < 1 = E[S_T^0]$. We can easily prove that

$$\frac{2}{k^2\sigma_*^2 + 2k\tilde{\mu}_*} \left[\exp\left(\frac{k^2\sigma_*^2}{2} + k\tilde{\mu}_*\right) - 1 \right] < 1$$

when

$$\frac{k^2\sigma_*^2}{2} + k\tilde{\mu}_* < 0.$$

Therefore, this phenomenon can occur for the $\tilde{\mu}_* < 0$ case.

V. COMPARISON TO DISCRETE-TIME MULTIPLICATIVE PROCESS

We investigate whether the above results are valid for the discrete-time multiplicative stochastic process (3).

If X_0 is a constant value, $\ln X_n = \ln M_1 + \dots + \ln M_n + \ln X_0$ approximately follows the normal distribution with mean μn and variance $\sigma^2 n$, where $\mu = E[\ln M_i]$ and $\sigma^2 = V[\ln M_i]$. Therefore, X_n approximately follows the lognormal distribution with $f_{\text{LN}}(x; \ln X_0 + \mu n, \sigma^2 n)$. This lognormality is similar to S_T for the geometric Brownian motion having $f_{\text{LN}}(x; \ln S_0 + \tilde{\mu}T, \sigma^2 T)$.

If the observation time n is drawn from the discrete uniform distribution on $\{1, 2, \dots, N\}$, the PDF and CCDF of X_n are expected to become approximately similar to $f_{\text{uni}}(x)$ and $F_{\text{uni}}(x)$ derived in Sections II and III, respectively. We compare the numerically generated CCDF and $F_{\text{uni}}(x)$ using two examples for the distribution of M_i .

The first example is the case wherein M_i is uniformly distributed over the interval $[a, b]$ ($a > 0$). The mean μ and variance σ^2 of $\ln M_i$ become

$$\mu = \int_a^b \frac{1}{b-a} \ln y dy = \frac{b \ln b - a \ln a}{b-a} - 1, \quad \sigma^2 = \int_a^b \frac{1}{b-a} (\ln y - \mu)^2 dy = 1 - ab \left(\frac{\ln b - \ln a}{b-a} \right)^2.$$

In the numerical calculation, we set $a = 1/2$ and $b = 3/2$ so that $\mu \approx -0.0452$ and $\sigma^2 \approx 0.0948$. The CCDFs for $N = 2, 5, 10$, and 20 are shown in Fig. 6(a) with points, calculated from 10^4 independent samples each. The solid curves represent $F_{\text{uni}}(x)$ in Eq. (14) by using $\tilde{\mu}_* = \mu N$ and $\sigma_*^2 = \sigma^2 N$.

The second example is where M_i follows the power-law distribution. The mean μ and variance σ^2 of $\ln M_i$ can be computed to be

$$\mu = \int_{m_0}^{\infty} \frac{\nu-1}{m_0} \left(\frac{y}{m_0} \right)^{-\nu} \ln y dy = \frac{1}{\nu-1} + \ln m_0, \quad \sigma^2 = \int_{m_0}^{\infty} \frac{\nu-1}{m_0} \left(\frac{y}{m_0} \right)^{-\nu} (\ln y - \mu)^2 dy = \frac{1}{(\nu-1)^2},$$

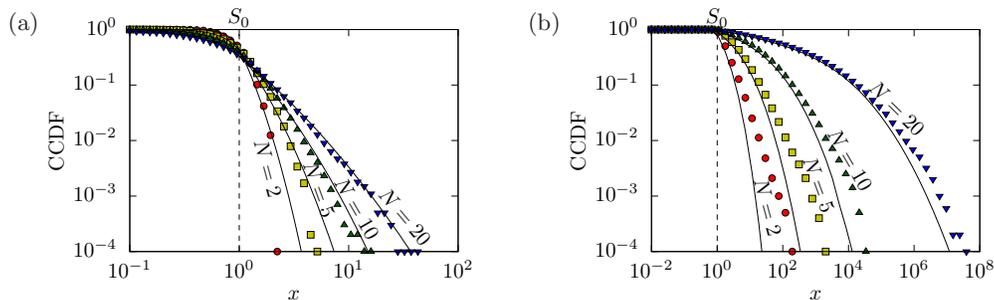


FIG. 6. Comparison between numerically calculated CCDF of X_n for $N = 2, 5, 10,$ and 20 (points) and the corresponding $F_{\text{uni}}(x)$ (curves). The random variable M_i is drawn from the uniform distribution on the interval $[1/2, 3/2]$ in (a) and the power-law distribution with exponent $\nu = 3$ and lower bound $m_0 = 1$ in (b).

respectively, where the power-law PDF with exponent $\nu (> 1)$ and lower bound $m_0 (> 0)$ is expressed as

$$\frac{\nu - 1}{m_0} \left(\frac{y}{m_0} \right)^{-\nu}.$$

In the numerical calculation, we used $\nu = 3$ and $m_0 = 1$, so that $\mu = 1/2$ and $\sigma^2 = 1/4$. Figure 6(b) shows the numerical result for $N = 2, 5, 10,$ and 20 with points and $F_{\text{uni}}(x)$ with curves.

A common property in these two examples is that the CCDF of X_n gradually approaches $F_{\text{uni}}(x)$ with increase in N . The deviation from $F_{\text{uni}}(x)$ in small N is attributed to the fact that the lognormal approximation based on the central limit theorem fails. In other words, the distribution of X_n for small N is dependent on the individuality of the distribution of M_i . For example, the numerical CCDF of X_n for the uniform M_i decays faster than $F_{\text{uni}}(x)$, whereas that for the power-law M_i decays slower than $F_{\text{uni}}(x)$. This observation is consistent with the property that the lognormal distribution has a tail heavier than the uniform distribution and the power-law distribution has an even further heavier tail.

VI. DISCUSSION

This study examined the geometric Brownian motion S_T under uniformly distributed observation time T . This model is a deformation of the DP distribution in which the observation time follows an exponential distribution. The PDF $f_{\text{uni}}(x)$ and CCDF $F_{\text{uni}}(x)$ are exactly calculated, although they have nontrivial complicated forms compared to those of the DP distribution. Consequently, the power-law form of the DP distribution is realized owing to the balance of the geometric Brownian motion and the exponential distribution of T .

By changing the observation time distribution, the DP distribution can be generalized. However, it remains unclear whether the exact expressions of the PDF and CCDF can be obtained in such a generalized case. In particular, the difficulty will arise in the computation of the (indefinite) integral, which involves $\exp(-a^2u^2 - b^2/u^2)$ factor. Therefore, observation time distributions that provide exact PDF and CCDF of S_T will be highly limited.

In this study, we assume that the initial value of the geometric Brownian motion, S_0 , is a constant, which is common to the DP distribution [26]. A possible extension of this study involves changing S_0 to a random variable. When S_0 is distributed lognormally and the observation time T is exponentially distributed, the distribution of S_T is exactly calculated and is referred to as the double Pareto-lognormal distribution [25, 29]. In the context of the double Pareto-lognormal distribution, an interesting challenge associated with this study is the derivation of the distribution of S_T with uniform T when the constant S_0 in this study is replaced with a lognormal random variable.

In Section V, the applicability of $F_{\text{uni}}(x)$ to discrete-time stochastic process (3) is examined. $F_{\text{uni}}(x)$ is expected to provide a reasonable estimate when the maximum number N of steps is large, whereas the distribution of X_n is strongly dependent on the property of the random variable M_i for small N . A further systematic and quantitative study is required to establish the connection between continuous- and discrete-time processes.

From a practical viewpoint, the exploration of empirical datasets that exhibit the proposed distribution is an important future research direction. Because the double Pareto distribution has been observed in various fields related to human activities and social phenomena [27, 28], we believe that the proposed distribution, which can be regarded as a modification of the DP distribution, is useful in the analysis of empirical data. In future studies, the practical importance of this study will be tested by its application to data analysis.

VII. CONCLUSION

The double Pareto (DP) distribution, a double-sided power-law distribution, can be obtained by geometric Brownian motion with a constant initial value and exponentially distributed observation time. This study investigates a deformation of the DP distribution by replacing the exponential distribution for the observation time with a continuous uniform distribution.

The exact forms of the probability density function (PDF) and complementary cumulative distribution function (CCDF) are derived using the error function [see Eqs. (10) and (14)]. Furthermore, the detailed shape of the PDF, e.g., the asymptotic form, unimodality, and peak position, is analyzed. The main finding of this study is the establishment of a parametric distribution that is an extension of the lognormal and power-law distributions.

Appendix A: Unimodality of $f_{\text{uni}}(x)$

Here we show that the PDF $f_{\text{uni}}(x)$ given in Eqs. (10) and (11) is unimodal. That is, $f_{\text{uni}}(x)$ has exactly one peak, at $x = S_0$ or $0 < x < S_0$.

Before the main proof of the unimodality, we prove that the function

$$Q(w) = \ln[e^{w^2} \operatorname{erfc}(w)]$$

is strictly convex. The second derivative of $Q(w)$ becomes

$$Q''(w) = \frac{2e^{-w^2}}{\operatorname{erfc}(w)^2} \left[e^{w^2} \operatorname{erfc}(w)^2 + \frac{2}{\sqrt{\pi}} w \operatorname{erfc}(w) - \frac{2}{\pi} e^{-w^2} \right] =: \frac{2e^{-w^2}}{\operatorname{erfc}(w)^2} q(w).$$

Notably, the key to a simple proof is to separate e^{-w^2} in the numerator. We analyze the function $q(w)$. The derivative of $q(w)$ becomes

$$q'(w) = 2e^{w^2} \operatorname{erfc}(w) \left[w \operatorname{erfc}(w) - \frac{e^{-w^2}}{\sqrt{\pi}} \right] = -2e^{w^2} \operatorname{erfc}(w) \int_w^\infty \operatorname{erfc}(u) du,$$

where

$$\int_w^\infty \operatorname{erfc}(u) du = -w \operatorname{erfc}(w) + \frac{e^{-w^2}}{\sqrt{\pi}}$$

can be derived by integration by parts. The erfc function always takes a positive value, and its integral always becomes positive. Therefore, $q'(w) < 0$, i.e., $q(w)$ is a decreasing function. Using the asymptotic expansion (12) of the erfc function, we obtain

$$\lim_{w \rightarrow \infty} q(w) = 0.$$

Consequently, $q(w)$ takes a positive value for any w . Thus, $Q''(w) > 0$, i.e., the strict convexity of $Q(w)$, is proven. For reference, Fig. 7 shows graphs of $Q(w)$, $Q'(w)$, and $Q''(w)$.

If $f_{\text{uni}}(x)$ has a peak at $x \neq S_0$, this peak position is characterized by the equation $f'_{\text{uni}}(x) = 0$. First, we prove the unimodality for the $\tilde{\mu}_* = 0$ case. By differentiating Eq. (11), the equation $f'_{\text{uni}}(x) = 0$ becomes

$$\left(1 - \ln \frac{x}{S_0} \right) \operatorname{erfc} \left(\pm \frac{\ln(x/S_0)}{\sqrt{2}\sigma_*} \right) = \mp \sqrt{\frac{2}{\pi}} \sigma_* \exp \left(-\frac{(\ln(x/S_0))^2}{2\sigma_*^2} \right),$$

where the upper and lower signs refer to the cases of $x > S_0$ and $x < S_0$, respectively. Taking the logarithm of this equation for $x < S_0$ and introducing $w = -\ln(x/S_0)/(\sqrt{2}\sigma_*)$, we obtain

$$Q'(w) = -\frac{\sqrt{2}}{\sigma_*}, \tag{A1}$$

where

$$Q'(w) = -\frac{2e^{-w^2}}{\sqrt{\pi} \operatorname{erfc}(w)} + 2w.$$

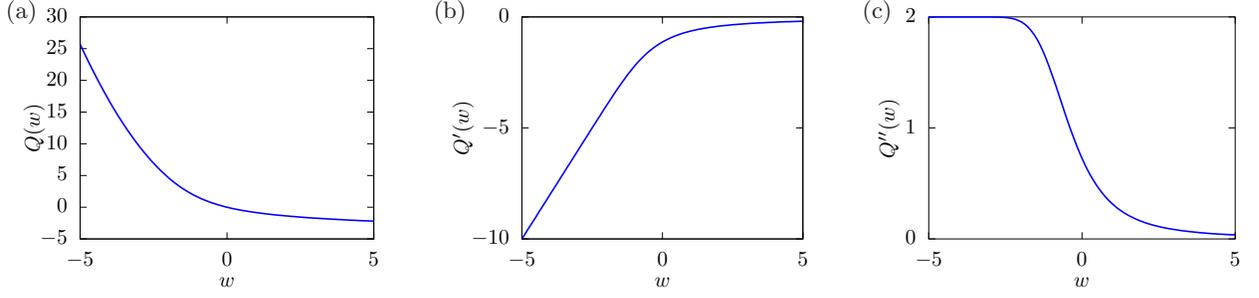


FIG. 7. Graphs of $Q(w)$, $Q'(w)$, and $Q''(w)$ in $-5 \leq w \leq 5$. (a) The function $Q(w)$ is strictly convex. (b) The derivative $Q'(w)$ is increasing and negative-valued. (c) The second derivative $Q''(w)$ is positive-valued.

Owing to the strict convexity of $Q(w)$, the derivative $Q'(w)$ is an increasing function. Moreover, $Q'(w)$ satisfies

$$\lim_{w \rightarrow -\infty} Q'(w) = -\infty, \quad \lim_{w \rightarrow \infty} Q'(w) = 0,$$

where the asymptotic expansion (12) is required to compute $Q'(\infty)$. Therefore, Eq. (A1) has a unique solution w for any $\sigma_* > 0$. If this solution is $w > 0$, the peak position of $f_{\text{uni}}(x)$ is $x = S_0 \exp(-\sqrt{2}\sigma_* w) < S_0$, and otherwise, $f_{\text{uni}}(x)$ is monotonically increasing in $0 < x < S_0$ and the peak is at $x = S_0$. The solution w is positive if and only if

$$-\frac{\sqrt{2}}{\sigma_*} < Q'(0) = -\frac{2}{\sqrt{\pi}},$$

which is equivalent to

$$\sigma_* > \sqrt{\frac{\pi}{2}}.$$

This threshold value $\sqrt{\pi/2}$ can be obtained as the solution of $f'_{\text{uni}}(S_0^-) = 0$ stated in Section II A.

The equation for $x > S_0$ becomes

$$Q'(w) = \frac{\sqrt{2}}{\sigma_*},$$

with $w = \ln(x/S_0)/(\sqrt{2}\sigma_*)$. This equation does not have solutions because $Q'(w)$ is always negative, as stated above; $f_{\text{uni}}(x)$ does not have a peak in $x > S_0$. Thus, the unimodality of $f_{\text{uni}}(x)$ for $\tilde{\mu}_* = 0$ is proven.

Next, we show the unimodality of $f_{\text{uni}}(x)$ for $\tilde{\mu}_* \neq 0$. The equation $f'_{\text{uni}}(x) = 0$ becomes

$$\operatorname{erfc}\left(\pm \frac{\ln(x/S_0) - \tilde{\mu}_*}{\sqrt{2}\sigma_*}\right) = \left(1 - \frac{2\tilde{\mu}_*}{\sigma_*^2}\right) \left(\frac{x}{S_0}\right)^{2\tilde{\mu}_*/\sigma_*^2} \operatorname{erfc}\left(\pm \frac{\ln(x/S_0) + \tilde{\mu}_*}{\sqrt{2}\sigma_*}\right), \quad (\text{A2})$$

where \pm indicates “+” for $x > S_0$ and “-” for $x < S_0$, as in the $\tilde{\mu}_* = 0$ case above. Since the erfc function is positive-valued, the existence of a solution x requires $1 - 2\tilde{\mu}_*/\sigma_*^2 > 0$. Otherwise, if $1 - 2\tilde{\mu}_*/\sigma_*^2 \leq 0$, Eq. (A2) does not have solutions, which means that the peak of $f_{\text{uni}}(x)$ is $x = S_0$. Taking the logarithm of Eq. (A2) and introducing $w = \pm \ln(x/S_0)/(\sqrt{2}\sigma_*)$, we obtain

$$Q\left(w + \frac{\tilde{\mu}_*}{\sqrt{2}\sigma_*}\right) - Q\left(w - \frac{\tilde{\mu}_*}{\sqrt{2}\sigma_*}\right) = \begin{cases} -\ln\left(1 - \frac{2\tilde{\mu}_*}{\sigma_*^2}\right) & x > S_0, \\ \ln\left(1 - \frac{2\tilde{\mu}_*}{\sigma_*^2}\right) & x < S_0. \end{cases} \quad (\text{A3})$$

The condition $1 - 2\tilde{\mu}_*/\sigma_*^2 > 0$ ensures that the term $\ln(1 - 2\tilde{\mu}_*/\sigma_*^2)$ takes a real value.

Let us focus on the $\tilde{\mu}_* > 0$ case. According to the theory of convex functions [30], for any strictly convex function ψ and constant $a > 0$, the function $\psi(w+a) - \psi(w-a)$ is an increasing function of w . Hence, the left-hand side of Eq. (A3) is an increasing function of w , and satisfies

$$\lim_{w \rightarrow -\infty} \left[Q\left(w + \frac{\tilde{\mu}_*}{\sqrt{2}\sigma_*}\right) - Q\left(w - \frac{\tilde{\mu}_*}{\sqrt{2}\sigma_*}\right) \right] = -\infty, \quad \lim_{w \rightarrow \infty} \left[Q\left(w + \frac{\tilde{\mu}_*}{\sqrt{2}\sigma_*}\right) - Q\left(w - \frac{\tilde{\mu}_*}{\sqrt{2}\sigma_*}\right) \right] = 0.$$

That is, the left-hand side of Eq. (A3) is always negative and can take any negative value by tuning w . Noting that $\ln(1 - 2\tilde{\mu}_*/\sigma_*^2) < 0$ for $\tilde{\mu}_* > 0$, we conclude that Eq. (A3) for $x > S_0$ does not have solution and for $x < S_0$ has a unique solution. As stated in the $\tilde{\mu}_* = 0$ case, the solution $w < 0$ corresponds to $x > S_0$, and the peak of $f_{\text{uni}}(x)$ is at $x = S_0$.

For the $\tilde{\mu}_* < 0$ case, the left-hand side of Eq. (A3) becomes positive-valued decreasing function and $\ln(1 - 2\tilde{\mu}_*/\sigma_*^2) > 0$. Therefore, Eq. (A3) for $x > S_0$ does not have solution and for $x < S_0$ has a unique solution, which indicates that the $f_{\text{uni}}(x)$ is unimodal.

ACKNOWLEDGMENTS

The authors are grateful to referees for providing details regarding the double Pareto-lognormal distribution. One of the authors (K.Y.) was supported by a Grant-in-Aid for Scientific Research (C) 19K03656 and 23K03264 from Japan Society for the Promotion of Science.

-
- [1] van Kampen N G 1992 *Stochastic Processes in Physics and Chemistry* (Elsevier)
 - [2] Redner S 2001 *A Guide to First-Passage Processes* (Cambridge University Press)
 - [3] Nair J, Wierman A and Zwart B 2022 *The Fundamentals of Heavy Tails* (Cambridge University Press)
 - [4] Nishimori H and Ortiz G 2011 *Elements of Phase Transitions and Critical Phenomena* (Oxford University Press)
 - [5] ben-Avraham D and Havlin S 2000 *Diffusion and Reaction in Fractals and Disordered Systems* (Cambridge University Press)
 - [6] Zwanzig R 2001 *Nonequilibrium Statistical Mechanics* (Oxford University Press)
 - [7] Newman M E J 2005 Power laws, Pareto distributions and Zipf's law *Contemp. Phys.* **46** 323–351
 - [8] Kobayashi N, Kuninaka H, Wakita J and Matsushita M 2011 Statistical features of complex systems—toward establishing sociological physics— *J. Phys. Soc. Jpn.* **80** 072001
 - [9] Limpert E, Stahel W A and Abbt M 2001 Log-normal distributions across the sciences: keys and clues *BioScience* **51** 341–352
 - [10] Crow E L and Shimizu K (ed.) 1988 *Lognormal distributions* (Dekker)
 - [11] Olver F W, Lozier D W, Boisvert R F and Clark C W 2010 *NIST Handbook of Mathematical Functions* (Cambridge University Press)
 - [12] Uttley P, McHardy I M and Vaughan S 2005 Non-linear X-ray variability in X-ray binaries and active galaxies *Mon. Not. R. Astron. Soc.* **359** 345–362
 - [13] Yamamoto K and Wakita J 2016 Analysis of a stochastic model for bacterial growth and the lognormality of the cell-size distribution *J. Phys. Soc. Jpn.* **85** 074004
 - [14] Koyama K, Yamamoto K and Ushio M 2017 A lognormal distribution of the lengths of terminal twigs on self-similar branches of elm trees *Proc. R. Soc. B* **284** 20162395
 - [15] Kolmogorov A N 1941 On the log-normal distribution of particles sizes during breakup process *Dokl. Akad. Nauk SSSR* **31** 99–101
 - [16] Takayasu H, Sato A and Takayasu M 1997 Stable infinite variance fluctuations in randomly amplified Langevin systems *Phys. Rev. Lett.* **79** 966–969
 - [17] Manrubia S C and Zanette D H 1999 Stochastic multiplicative processes with reset events *Phys. Rev. E* **59** 4945–4948
 - [18] Yamamoto K and Yamazaki Y 2012 Power-law behavior in a cascade process with stopping events *Phys. Rev. E* **85** 011145
 - [19] Yamamoto K 2014 Stochastic model of Zipf's law and the universality of the power-law exponent *Phys. Rev. E* **89** 042115
 - [20] Yamamoto K 2015 A simple view of the heavy-tailed sales distributions and application to the box-office grosses of U.S. movies *Europhys. Lett.* **108** 68004
 - [21] Levy M and Solomon S 1996 Power laws are logarithmic Boltzmann laws *Int. J. Mod. Phys. C* **7** 595–601
 - [22] Yamamoto K and Yamazaki Y 2022 Analysis and application of multiplicative stochastic process with a sample-dependent lower bound *J. Phys. Soc. Jpn.* **91** 064803
 - [23] Øksendal B 2013 *Stochastic Differential Equations* (Springer)
 - [24] Paul W and Baschnagel J 2013 *Stochastic Processes: From Physics to Finance* (Springer)
 - [25] Reed W J and Jorgensen M 2004 The double Pareto-lognormal distribution—a new parametric model for size distributions *Commun. Statist.* **33** 1733–1753
 - [26] Mitzenmacher M 2004 Dynamic models for file sizes and double Pareto distributions *Internet Math.* **1** 305–333
 - [27] Reed W J 2001 The Pareto, Zipf and other power laws *Econ. Lett.* **74** 15
 - [28] Wang C, Guan X, Qin T and Yang T 2016 Modeling heterogeneous and correlated human dynamics of online activities with double Pareto distribution *Information Sciences* **330** 186–198
 - [29] Grbac N and Grbac T G 2023 Letter to the editor: on the paper “The double Pareto-lognormal distribution—a new parametric model for size distributions” and its correction *Commun. Statistics* (DOI: 10.1080/03610926.2023.2174788)
 - [30] Roberts A W and Varberg D E 1973 *Convex Functions* (Academic Press)