Log-amplitude statistics for Beck-Cohen superstatistics

Ken Kiyono[1] and Hidetoshi Konno[2]

[1]*Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama-cho, Toyonaka, Osaka 560-8531, Japan*

[2]*Department of Risk Engineering, Faculty of Information and Systems, University of Tsukuba, Tsukuba, Ibaraki 305-8573, Japan*

(Received 6 April 2013; published 6 May 2013)

As a possible generalization of Beck-Cohen superstatistical processes, we study non-Gaussian processes with temporal heterogeneity of local variance. To characterize the variance heterogeneity, we define log-amplitude cumulants and log-amplitude autocovariance and derive closed-form expressions of the log-amplitude cumulants for χ^2, inverse χ^2, and log-normal superstatistical distributions. Furthermore, we show that χ^2 and inverse χ^2 superstatistics with degree 2 are closely related to an extreme value distribution, called the Gumbel distribution. In these cases, the corresponding superstatistical distributions result in the q-Gaussian distribution with $q = 5/3$ and the bilateral exponential distribution, respectively. Thus, our finding provides a hypothesis that the asymptotic appearance of these two special distributions may be explained by a link with the asymptotic limit distributions involving extreme values. In addition, as an application of our approach, we demonstrated that non-Gaussian fluctuations observed in a stock index futures market can be well approximated by the χ^2 superstatistical distribution with degree 2.

DOI: 10.1103/PhysRevE.87.052104 PACS number(s): 05.40.−a, 02.50.−r, 87.15.Ya, 05.45.Tp

I. INTRODUCTION

Non-Gaussian fluctuations observed in complex systems have attracted a wide interest in statistical physics [1,2]. The importance of understanding such non-Gaussian phenomena is underscored by studies demonstrating that the non-Gaussian properties were associated with crisis or catastrophic events, such as higher mortality [3,4], stock market crashes [5,6], and major earthquakes [7,8]. To gain helpful and insightful information about these systems, a widely applicable method for characterizing non-Gaussian time series is required.

Complex nonequilibrium systems are often effectively described by a superposition of different statistical models [9]. Based on this paradigm, Beck and Cohen proposed superstatistics [10,11] in which superposition of different dynamics on different time scales is assumed. The stationary distributions of such superstatistical models typically exhibit non-Gaussian probability density functions (PDFs) with fat tails and have been successful in describing non-Gaussian fluctuations observed in a wide variety of real-world signals [2,11]. In addition, from the viewpoint of a maximum entropy principle, the theoretical framework of superstatistical processes has been developed [12–14].

As a simple dynamical realization of superstatistics, Beck introduced a nonequilibrium process described by stochastic differential equations with fluctuating parameters [15]. To explain this idea, let us consider a Langevin equation for a variable u:

$$\frac{du}{dt} = \gamma F(u) + \hat{\sigma} L(t), \tag{1}$$

where $L(t)$ is Gaussian white noise, γ is a positive friction constant, $\hat{\sigma}$ describes the strength of the noise, and $F(u) = -(\partial/\partial u)V(u)$ is a drift force. If γ and $\hat{\sigma}$ are constant, the equilibrium distribution of u is described by

$$p(u|\beta) = \frac{1}{Z(\beta)} e^{-\beta V(u)}, \tag{2}$$

where an intensive parameter is defined as $\beta = 2\gamma/\hat{\sigma}^2$ as in Brownian motion of a particle with unit mass and $Z(\beta)$ is a normalization constant. In the framework of superstatistics, the intensive parameter β is not constant but stochastically fluctuates so that β has the probability density function $f(\beta)$. The parameter fluctuations are assumed to be on a long time scale so that the system can temporarily reach local equilibrium. In this case, the marginal probability $p(u)$ is described by

$$p(u) = \int_0^\infty p(u|\beta)\, f(\beta)\, d\beta. \tag{3}$$

In this paper, we will consider only the case of linear drift force $F(u) = -u$. In this case, the local equilibrium distribution $p(u|\beta)$ is restricted to a Gaussian:

$$p(u|\beta) = \sqrt{\frac{\beta}{2\pi}} \exp\left(-\frac{\beta}{2}u^2\right). \tag{4}$$

For practical applications of superstatistical distributions [Eq. (3)], Beck *et al.* reported that many experimental data fall into three different classes [11]: χ^2 superstatistics, inverse χ^2 superstatistics, and log-normal superstatistics (as shown in Fig. 1). In the χ^2 superstatistics (also called gamma superstatistics), the intensity parameter β is χ^2 distributed with degree n:

$$f(\beta) = \frac{1}{\Gamma\left(\frac{n}{2}\right)} \left(\frac{n}{2\beta_0}\right)^{\frac{n}{2}} \beta^{\frac{n}{2}-1} e^{-\frac{n\beta}{2\beta_0}}, \tag{5}$$

where n takes a positive real number, β_0 is the mean of β, and $\Gamma(x)$ is the Gamma function. Equation (5) has the same functional form as the gamma distribution [2], because n is not restricted to the integer value, and β_0 corresponds to a scale parameter of the gamma distribution. In this case, the marginal

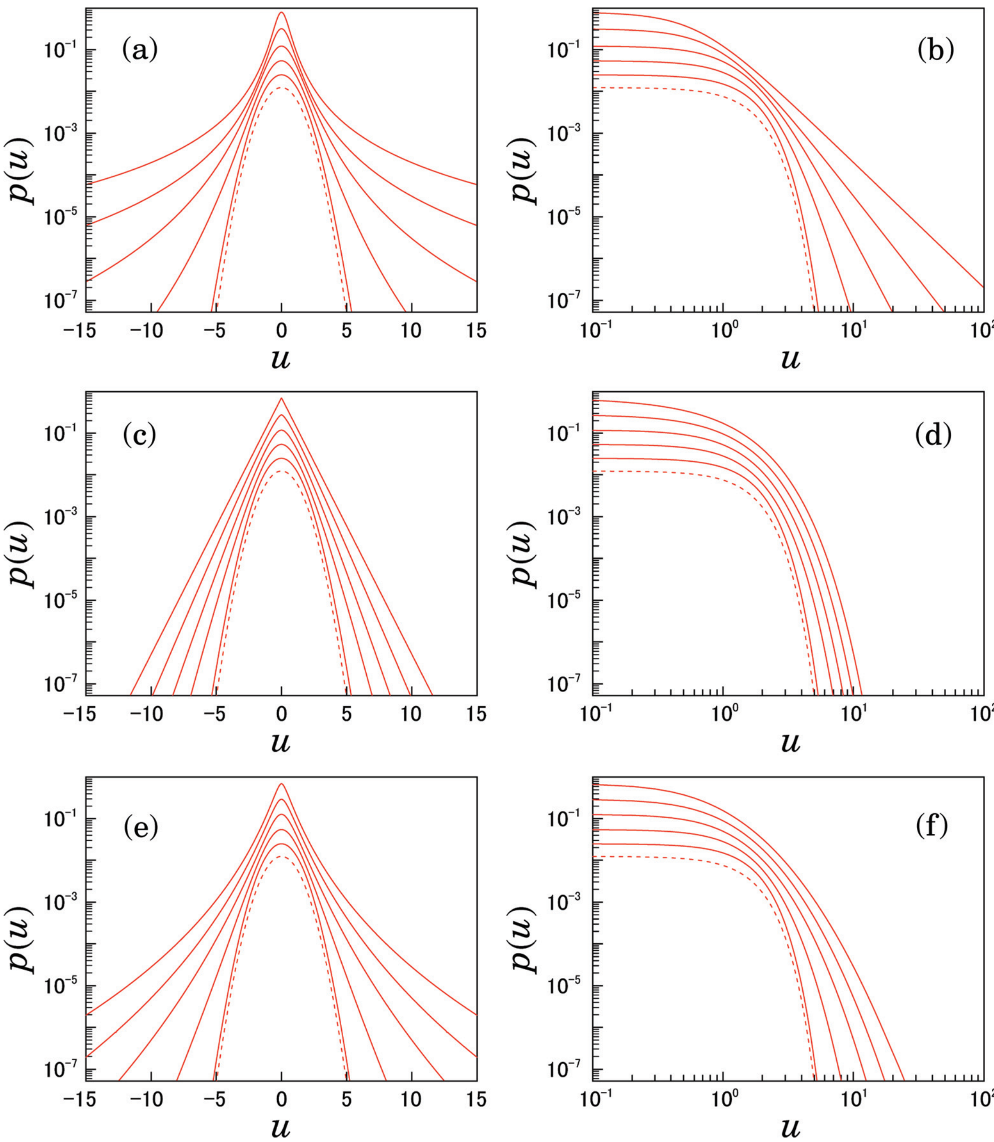

FIG. 1. (Color online) Examples of superstatistical distributions in (a, b) χ^2 superstatistics, (c, d) inverse χ^2 superstatistics, and (e, f) log-normal superstatistics. For (a–d) χ^2 and inverse χ^2 superstatistics, $n = 2$, 3, 5, 10, and 100 from top to bottom. For log-normal superstatistics, $s = 1.2$, 1.0, 0.8, 0.5, and 0.1 from top to bottom. For comparison, the dashed lines indicate a Gaussian distribution.

probability $p(u)$ [Eq. (3)] results in

$$p(u) = \frac{\sqrt{\beta_0}\,\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi\, n}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{\beta_0}{n}u^2\right)^{-\frac{n+1}{2}}, \tag{6}$$

which is the same form as the q-Gaussian distribution with $q = (n+3)/(n+1)$, also called the Student t distribution. This superstatistical distribution [Eq. (6)] appears in the framework of Tsallis statistics [16]. As a characteristic feature of this class, the superstatistical distributions exhibit power-law tails for large $|u|$:

$$p(u) \sim |u|^{-(n+1)}. \tag{7}$$

On the other hand, in the inverse χ^2 superstatistics (also called inverse gamma superstatistics), β^{-1} rather than β is χ^2 distributed with degree n. That is, $f(\beta)$ is given by the inverse χ^2 distribution [17]:

$$f(\beta) = \frac{1}{\Gamma\left(\frac{n}{2}\right)}\left(\frac{n\beta_0}{2}\right)^{\frac{n}{2}}\beta^{-\frac{n}{2}-1}e^{-\frac{n\beta_0}{2\beta}}, \tag{8}$$

where n takes real positive values, β_0 is a scale parameter, and this equation has the same functional form as the inverse gamma distribution [2]. In this case, the marginal probability $p(u)$ [Eq. (3)] results in

$$p(u) = \frac{(n\beta_0)^{\frac{n+1}{4}}}{\sqrt{\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(\frac{|u|}{2}\right)^{\frac{n-1}{2}} K_{\frac{n-1}{2}}(\sqrt{n\beta_0}\,|u|), \tag{9}$$

where $K_\alpha(u)$ is the modified Bessel function of the second kind [18]. In the special case of $n = 2$, $p(u)$ results in the bilateral exponential distribution, also known as the Laplace distribution:

$$p(u) = \sqrt{\frac{\beta_0}{2}}\, e^{-\sqrt{2\beta_0}|u|}. \tag{10}$$

In the log-normal superstatistics, β is log-normally distributed:

$$f(\beta) = \frac{1}{\sqrt{2\pi}s\beta} \exp\left\{-\frac{\ln^2(\beta/m)}{2s^2}\right\}, \tag{11}$$

where m and s^2 are suitable mean and variance parameters, respectively. In this case, the marginal probability $p(u)$ [Eq. (3)] is given by

$$p(u) = \frac{1}{2\pi s}\int_0^\infty \beta^{-1/2} \exp\left\{-\frac{\beta}{2}u^2 - \frac{\ln^2(\beta/m)}{2s^2}\right\} d\beta, \tag{12}$$

which does not have a closed form. This distribution [Eq. (12)] has the same form as the multiplicative log-normal distribution [19,20].

In this paper, as a possible generalization of the above superstatistical processes, we study non-Gaussian processes with variance heterogeneity. Our approach is closely related with Castaing's formulation of a PDF observed in turbulencelike fluctuations [19–21]. In the study of the velocity difference between two points in fully developed turbulent flows, Castaing *et al.* [19] introduced the following equation:

$$p(u) = \int_0^\infty \frac{1}{\sigma} P_L\left(\frac{u}{\sigma}\right) G(\ln\sigma)\, d(\ln\sigma), \tag{13}$$

where P_L describes velocity fluctuations at integral scale L and G describes the nature of the fluctuating energy dissipation. Under Kolmogorov's refined similarity hypothesis [22], G is assumed to be an infinitely divisible distribution, such as a Gaussian distribution. In this framework, if we consider the PDF observed at a fixed scale, Castaing's equation [Eq. (13)] can be seen as another form of superstatistics. That is, Eq. (13) can be recast as

$$p(u) = \int_0^\infty p(u|\sigma)\, g(\sigma)\, d\sigma, \tag{14}$$

where $p(u|\sigma) = P_L(u/\sigma)/\sigma$ and $g(\sigma) = G(\ln\sigma)/\sigma$. In this case, if $P_L(u)$ in Eq. (13) is assumed to be the standard Gaussian distribution, $p(u|\sigma)$ is described by

$$p(u|\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{u^2}{2\sigma^2}\right), \tag{15}$$

and the fluctuation of the standard deviation σ is described by $g(\sigma)$. In the framework of superstatistics, if the noise strength $\hat{\sigma}$, instead of β, fluctuates so that $\sigma = \hat{\sigma}/\sqrt{2\gamma}$ has the probability density function $g(\sigma)$, the marginal probability $p(u)$ is described by Eq. (14).

In the next section, to characterize superstatistical distributions, we will introduce log-amplitude cumulants, defined as cumulants of G in Eq. (13). As we will see in Sec. III, the advantage of our approach is that the log-amplitude cumulants can provide closed-form expressions for all of the log-amplitude cumulants of the above-mentioned superstatistical distributions. Because the log-amplitude cumulants can be directly estimated from the observed time series, they are applicable to parameter estimation and model selection for superstatistical distributions. Furthermore, we will show that χ^2 and inverse χ^2 superstatistics with $n = 2$ have an intriguing relation with an extreme value distribution known as the Gumbel distribution. In these cases, the marginal probabilities result in the Student t distribution with degree 2, which is the same as the q-Gaussian distribution with $q = 5/3$ and the bilateral exponential distribution, respectively. Thus, our finding suggests that the extreme value distribution could be involved in the appearance of two special non-Gaussian distributions.

In addition, it is also important to note that superstatistics assumes the time-scale separation between fast relaxation of the local dynamics and slow driving of the intensity parameter. Hence, to study the superstatistical processes, a method to estimate the time scale characteristic of the intensity parameter fluctuations is required. To this end, we will introduce the log-amplitude autocovariance and characterize superstatistical processes in Sec. IV. As an application of our approach, we will study non-Gaussian fluctuations observed in a stock index futures market.

II. DEFINITION OF LOG-AMPLITUDE CUMULANTS

In this section, we define log-amplitude cumulants to characterize non-Gaussian distributions. By the change of the variable $y = \ln\sigma$ in Eq. (13), the marginal probability $p(u)$ is expressed by

$$p(u) = \int_{-\infty}^\infty p(u|y)\, G(y)\, dy, \tag{16}$$

where $p(u|y) = P_L(u\, e^{-y})\, e^{-y}$, and y is referred to as the log amplitude [21]. In this framework, we define the kth log-amplitude cumulant C_k as the kth cumulant of Y, where Y is a random variable following the distribution $G(y)$. The log-amplitude cumulants C_k are given by the cumulant-generating function:

$$\Phi_G(q) = \ln\langle e^{qY}\rangle = \ln\left\{\int_{-\infty}^\infty e^{qy}\, G(y)\, dy\right\}, \tag{17}$$

where $\langle\cdot\rangle$ denotes the expectation value. Using $\Phi_G(q)$, C_k can be extracted as

$$C_k = \left.\frac{d^k\Phi(q)}{dq^k}\right|_{q=0}. \tag{18}$$

For instance, from a formal power series of $\Phi_G(q)$,

$$\begin{aligned}\ln\langle e^{qY}\rangle &= -\sum_{i=1}^\infty \frac{(1-\langle e^{qY}\rangle)^i}{i} \\ &= -\sum_{i=1}^\infty \frac{1}{i}\left(-\sum_{j=1}^\infty \langle Y^j\rangle \frac{q^j}{j!}\right)^i \\ &= \langle Y\rangle q + (\langle Y^2\rangle - \langle Y\rangle^2)\frac{q^2}{2!} + \cdots,\end{aligned} \tag{19}$$

the first three log-amplitude cumulants can be expressed by

$$C_1 = \langle Y\rangle, \tag{20}$$

$$C_2 = \langle Y^2\rangle - \langle Y\rangle^2, \tag{21}$$

$$C_3 = \langle Y^3\rangle - 3\langle Y^2\rangle\langle Y\rangle + 2\langle Y\rangle^3. \tag{22}$$

Note that C_2 and C_3 are equal to, respectively, the second and third central moments of Y.

If we assume $G(y) = \delta(y - \ln \sigma_0)$, $p(u)$ is described by a Gaussian distribution with variance σ_0^2. In this case, the C_k are given by

$$C_1 = \ln \sigma_0, \tag{23}$$

$$C_k = 0 \quad \text{for } k \geqslant 2. \tag{24}$$

Therefore, nonzero values of C_k for $k \geqslant 2$ can characterize systematic deviations from a Gaussian shape.

It is important to point out that the log-amplitude cumulants C_k can be directly estimated from the observed time series $\{u_i\}$. To explain this, we assume a multiplicative stochastic process $\{U_1, U_2, \ldots, U_i, \ldots\}$ as described by

$$U_i = W_i \exp Y_i, \tag{25}$$

where W_i and Y_i are random variables independent of each other and obey $p(u|y = 1)$ and $G(y)$, respectively, in Eq. (16). In this process, the PDF of U_i is given by Eq. (16). If we assume that W_i is a standard Gaussian random variable with zero mean and unit variance, it is possible to derive the relation between logarithmic absolute moments of U_i and log-amplitude cumulants C_k [21]. For instance, the first three log-amplitude cumulants are given by

$$C_1 = \langle \ln |U_i| \rangle + \frac{\ln 2 + \gamma}{2}, \tag{26}$$

$$C_2 = \langle (\ln |U_i| - \langle \ln |U_i| \rangle)^2 \rangle - \frac{\pi^2}{8}, \tag{27}$$

$$C_3 = \langle (\ln |U_i| - \langle \ln |U_i| \rangle)^3 \rangle + \frac{7}{4}\zeta(3), \tag{28}$$

where $\gamma \approx 0.57721566$ is the Euler-Mascheroni constant and $\zeta(n)$ is the Riemann zeta function [$\zeta(3) \approx 1.2020569$]. Thus, using Eqs. (26)–(28), we can estimate C_k from an observed time series $\{u_i\}$, where u_i is a realization of U_i.

III. LOG-AMPLITUDE CUMULANTS OF SUPERSTATISTICAL DISTRIBUTIONS

In this section, we study the relation between superstatistical distributions [Eq. (3)] and the Castaing-type description [Eq. (16)] and provide closed-form expressions of log-amplitude cumulants for superstatistical distributions [Eqs. (6), (9), and (12)].

From the comparison between Eqs. (3) and (16), the $f(\beta)$ can be transformed into the $G(y)$ as

$$G(y) = 2e^{-2y} f(e^{-2y}). \tag{29}$$

Using this relation, we can obtain the Castaing-type description of a superstatistical distribution.

In the case of χ^2 superstatistics [Eq. (5)], the corresponding $G(y)$ is given by

$$G(y) = \frac{2}{\Gamma\left(\frac{n}{2}\right)} \exp\left\{-\frac{y - \frac{1}{2}\ln\frac{n}{2\beta_0}}{1/n} - \exp\left(-\frac{y - \frac{1}{2}\ln\frac{n}{2\beta_0}}{1/2}\right)\right\}, \tag{30}$$

and its cumulant-generating function is given by

$$\Phi_G(q) = \frac{1}{\Gamma\left(\frac{n}{2}\right)} \left(\frac{n}{2\beta_0}\right)^{q/2} \Gamma\left(\frac{n-q}{2}\right). \tag{31}$$

From this, we can obtain all of the log-amplitude cumulants C_k as

$$C_1 = \frac{1}{2}\left\{\ln\left(\frac{n}{2\beta_0}\right) - \psi^{(0)}\left(\frac{n}{2}\right)\right\}, \tag{32}$$

$$C_k = \frac{1}{(-2)^k} \psi^{(k-1)}\left(\frac{n}{2}\right) \quad \text{for } k \geqslant 2, \tag{33}$$

where $\psi^{(m)}(x)$ is the polygamma function of order m, defined as the $(m+1)$th derivative of the logarithm of the gamma function.

Intriguingly, when $n = 2$, Eq. (30) is coincident with the Gumbel distribution for the maximum extreme [23]. In extreme value statistics, the Gumbel distribution is known as one of the extreme value distributions [24] which are the limiting distributions for the maximum or the minimum among a large number of independent identically distributed random variables [23]. As Gaussian and stable distributions are natural limit distributions when considering linear combinations such as sums of independent variables [25], extreme value distributions are natural limit distributions when considering min and max operations of independent variables. They naturally emerge in contexts related to reliability and risk assessments where one needs to consider extreme events such as floods, hurricanes, and earthquakes. Using extreme value distributions, it is possible to assess the occurrence probability of such extreme events. In the χ^2 superstatistics when $n = 2$, the marginal probability $p(u)$ results in the Student t distribution with degree 2, which exhibits power-law tails, asymptotically as $|u|^{-3}$. This distribution is also known as the q-Gaussian distribution with $q = 5/3$. Hence, the appearance of this distribution may indicate the existence of the max operation in the system.

In the case of inverse χ^2 superstatistics [Eq. (8)], the corresponding $G(y)$ is given by

$$G(y) = \frac{2}{\Gamma\left(\frac{n}{2}\right)} \exp\left\{\frac{y + \frac{1}{2}\ln\frac{n}{2\beta_0}}{1/n} - \exp\left(\frac{y + \frac{1}{2}\ln\frac{n}{2\beta_0}}{1/2}\right)\right\}, \tag{34}$$

and its cumulant-generating function is given by

$$\Phi_G(q) = \frac{1}{\Gamma\left(\frac{n}{2}\right)} \left(\frac{2}{n\beta_0}\right)^{q/2} \Gamma\left(\frac{n+q}{2}\right). \tag{35}$$

From this, we can obtain the log-amplitude cumulants C_k as

$$C_1 = \frac{1}{2}\left\{\ln\left(\frac{2}{n\beta_0}\right) + \psi^{(0)}\left(\frac{n}{2}\right)\right\}, \tag{36}$$

$$C_k = \frac{1}{2^k} \psi^{(k-1)}\left(\frac{n}{2}\right) \quad \text{for } k \geqslant 2. \tag{37}$$

When $n = 2$, Eq. (34) is coincident with the Gumbel distribution for the minimum extreme, which is again an extreme value distribution. In this case, the marginal probability $p(u)$ results in the bilateral exponential distribution [Eq. (10)].

In the case of log-normal superstatistics [Eq. (11)], the corresponding $G(y)$ is given by

$$G(y) = \frac{1}{\sqrt{2\pi}\,(s/2)} \exp\left\{-\frac{\left(y+\frac{1}{2}\ln m\right)^2}{2(s/2)^2}\right\}, \tag{38}$$

and its cumulant-generating function is given by

$$\Phi_G(q) = m^{-\frac{q}{2}}\, e^{\frac{q^2 s^2}{8}}. \tag{39}$$

From this, we can obtain the log-amplitude cumulants C_k as

$$C_1 = -\frac{\ln m}{2}, \tag{40}$$

$$C_2 = \frac{s^2}{4}, \tag{41}$$

$$C_k = 0 \quad \text{for } k \geqslant 3. \tag{42}$$

From the viewpoint of log-amplitude cumulants, the log-normal superstatistical distribution plays a central role in non-Gaussian distribution families, analogous to a Gaussian distribution in which all of the cumulants beyond the second cumulant are zero. When we consider the fluctuation of the standard deviation σ, a log-normally distributed σ can be derived from the maximum entropy principle [26]. Because of $\sigma > 0$, we can infer that at least the mean and variance of $\ln\sigma$ are fixed throughout time. Under these constraints, the log-normal distribution can be obtained as the maximum entropy distribution of σ. Therefore, the normally distributed log amplitude [Eq. (38)] may emerge naturally.

IV. PARAMETER ESTIMATION FOR SUPERSTATISTICAL DISTRIBUTIONS

In the previous section, we derived closed-form expressions of log-amplitude cumulants C_k for three superstatistical distributions. As shown in Figs. 2(a) and 2(b), there exists the one-to-one correspondence between C_2 and model parameters. Furthermore, as shown in Fig. 2(c), the difference among three superstatistical models can be characterized by C_3. In the (C_2,C_3) plane shown in Fig. 2(c), we can clearly see the symmetric structure between χ^2 and inverse χ^2 superstatistics, where the line of symmetry corresponds to log-normal superstatistics. Hence, the log-amplitude cumulants are applicable to parameter estimation and model selection for superstatistical distributions. In this section, we propose a parameter estimation method using C_k, and we study numerical and real-world examples involved with superstatistical distributions.

To explain our approach, let us assume that a stationary time series $\{u_i\}_{i=1}^N$ follows a superstatistical distribution. To estimate the logarithmic absolute moments, $\left\langle(\ln|U_i| - \langle\ln|U_i|\rangle)^k\right\rangle$, in Eqs. (26)–(28), we employ the following estimators:

$$M_k = \frac{1}{N}\sum_{i=1}^{N}(\ln|u_i| - M_1)^k \quad (k = 2,3,\ldots), \tag{43}$$

where M_1 is the estimator of $\langle\ln|U_i|\rangle$:

$$M_1 = \frac{1}{N}\sum_{i=1}^{N}\ln|u_i|. \tag{44}$$

Using M_k, the parameter n in χ^2 [Eq. (5)] and inverse χ^2 [Eq. (8)] superstatistics can be estimated from

$$n = 2\Psi^{(1)}\left(4M_2 - \frac{\pi^2}{2}\right), \tag{45}$$

where $\Psi^{(m)}(x)$ is the inverse function of the polygamma function $\psi^{(m)}(x)$. In log-normal superstatistics, the parameter s^2 [Eq. (11)] can be estimated from

$$s^2 = 4M_2 - \frac{\pi^2}{2}. \tag{46}$$

Except for the case of $n \leqslant 2$ in χ^2 superstatistics, the time series $\{u_i\}$ can be standardized to zero mean and unit variance. If $\{u_i\}$ is already standardized, the value of a scale parameter, such as β_0 and m, is uniquely determined by the other parameter value. That is, for the χ^2 superstatistical distribution with $n > 2$, the β_0 is given by

$$\beta_0 = \frac{n}{n-2}, \tag{47}$$

for the inverse χ^2 superstatistical distribution,

$$\beta_0 = 1, \tag{48}$$

and for the log-normal superstatistical distribution,

$$m = e^{\frac{s^2}{2}}. \tag{49}$$

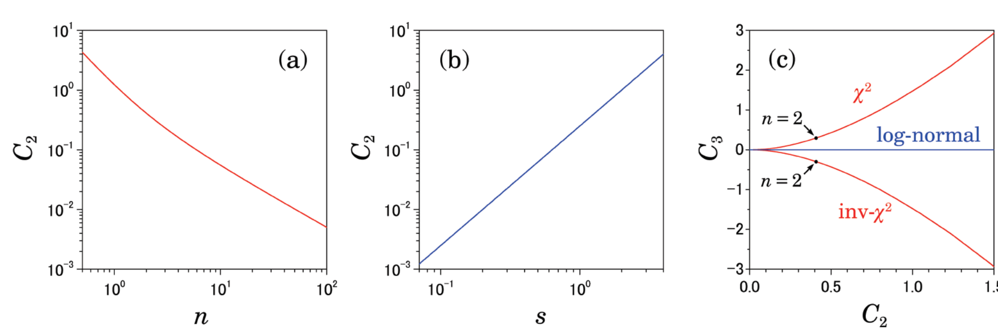

FIG. 2. (Color online) Relation between the first three log-amplitude cumulants and parameters in superstatistical distributions. (a) The second log-amplitude cumulant C_2 vs the shape parameter n of χ^2 and inverse χ^2 superstatistical distributions. (b) C_2 vs the shape parameter s of log-normal superstatistical distribution. (c) The third log-amplitude cumulant C_3 vs C_2 of χ^2, inverse χ^2 (inv-χ^2), and log-normal superstatistical distributions.

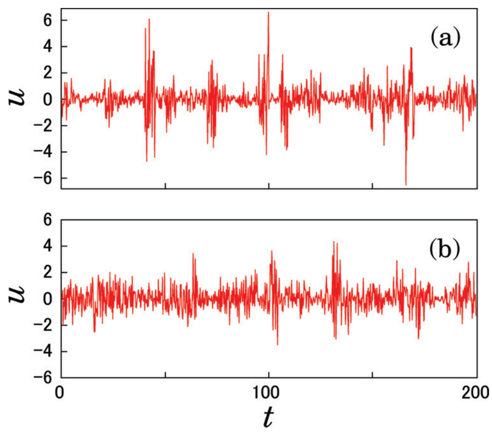

FIG. 3. (Color online) Numerically generated time series of superstatistical processes [Eqs. (50) and (51)], where $\gamma = 10$ and $T = 5$. (a) The χ^2 superstatistical process, where $f(\beta)$ is Eq. (5) with $n = 2$ and $\beta_0 = 7$. (b) The inverse χ^2 superstatistical process, where $f(\beta)$ is Eq. (8) with $n = 2$ and $\beta_0 = 1$.

By substitution of those relations into the first log-amplitude cumulant C_1, the value of C_1 can be estimated from the standardized time series. Hence, based on the estimated C_1 and C_2, it is possible to determine a candidate distribution for the non-Gaussian time series. Although agreement of higher-order log-amplitude cumulants is also important for the model selection, an accurate estimation of such higher-order statistics requires much larger amounts of data points. Hence, lower-order statistics, such as C_1 and C_2, can provide more accurate estimation.

For the case of χ^2 superstatistical distributions with $n \leqslant 2$, the variance is not finite. Therefore, proper standardization of the observed time series is impossible, which means that the time series no longer satisfy Eq. (47). However, all of the log-amplitude cumulants have finite values. Thus, the estimated values of C_2 and C_3 can provide useful information for the model selection. In addition, C_1 can be used for estimation of the scale parameter β_0.

A. Numerical tests

To test our approach, we carry out numerical experiments. To generate time series following superstatistical distributions, we consider a superstatistical process as mentioned in the Introduction. Here we perform numerical integrations of the stochastic differential equation:

$$du(t) = -\gamma u(t)\,dt + \sqrt{\frac{2\gamma}{\beta(t)}}\,dW(t), \tag{50}$$

where γ is a constant and $dW(t)$ is the infinitesimal increment of a standard Wiener process $W(t)$. In this case, $\beta(t)$ is assumed to be a step function with a time-step length T:

$$\beta(t) = \sum_{j=0}^{\infty} B_j\,\chi_{I_j}(t), \tag{51}$$

where B_j is a random variable following $f(\beta_j)$, $I_j = [jT,(j+1)T)$ is the interval of time, and $\chi_{I_j}(t)$ is the indicator function:

$$\chi_{I_j}(t) = \begin{cases} 1 & \text{if } t \in I_j \\ 0 & \text{if } t \notin I_j \end{cases}. \tag{52}$$

In superstatistical processes, the time scale T must be larger than the time scale of the relaxation time $1/\gamma$.

In our numerical study, we chose $\gamma = 10$ and assumed $f(\beta_j)$ to be a χ^2 distribution [Eq. (5)] or inverse χ^2 distribution [Eq. (8)] with $n = 2$. For the numerical integration, we used the Euler-Maruyama method with time step $\Delta = 0.001$. Samples of the numerically generated time series are shown in Fig. 3. As shown in Fig. 4, the estimated PDFs (open circles) from the time series are in good agreement with the theoretical predictions (dashed lines).

In general, accurate estimation of higher-order cumulants from the observed time series requires large amounts of data. To evaluate the accuracy of the estimated C_k, we generated 1000 samples with N data points at sampling intervals at T and estimated C_1, C_2, and C_3 from these samples. As shown in Fig. 5, the sample averages were in a good agreement with the theoretical predictions (dashed lines).

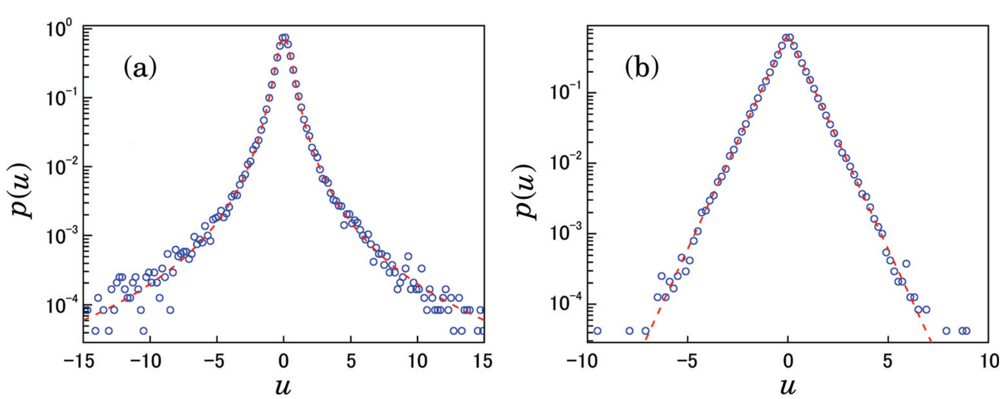

FIG. 4. (Color online) The probability density function (PDF) of superstatistical processes where $\gamma = 10$ and $T = 5$. (a) The χ^2 superstatistical process, where $f(\beta)$ is Eq. (5) with $n = 2$ and $\beta_0 = 7$. (b) The inverse χ^2 superstatistical process, where $f(\beta)$ is Eq. (8) with $n = 2$ and $\beta_0 = 1$. The PDFs (open circles) were estimated from the numerically generated time series (Fig. 3). The dashed lines indicate the theoretical distributions.

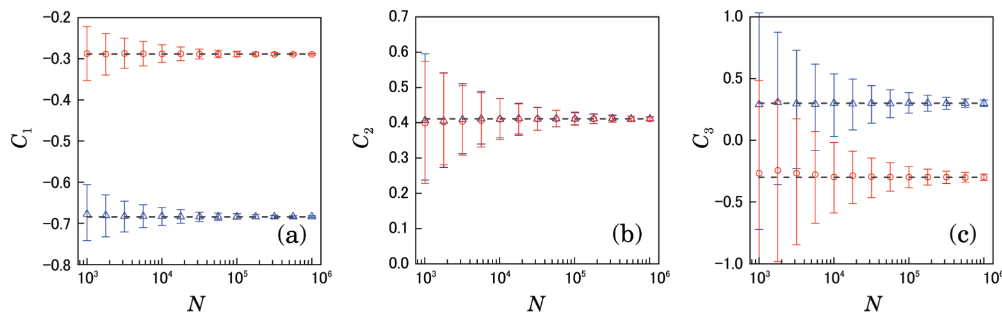

FIG. 5. (Color online) Estimated values of log-amplitude cumulants C_k from the numerically generated time series (Fig. 3) with number of data points N. The mean values of 1000 samples of χ^2 and inverse χ^2 superstatistical processes with $n = 2$ are indicated by triangles (blue) and circles (red), respectively. The dashed lines correspond to the theoretical values. The error bars denote the 5th and 95th percentiles of the estimated values.

In the study of superstatistical processes, it is also important to evaluate the time scale T, because the time-scale separation between fast relaxation of the local dynamics and slow driving of the intensity parameter β is essential for the superstatistical process. To estimate the time scale T in our model, we here define the log-amplitude autocovariance of a stochastic process $\{U(t)\}$ as

$$R_{\ln|u|}(s) = \langle \ln|U(t)|\,\ln|U(t+s)|\rangle - \langle \ln|U(t)|\rangle^2. \tag{53}$$

By taking the limit $1/\gamma \to 0$ in Eq. (50), we obtain

$$R_{\ln|u|}(s) = \begin{cases} C_2 + \frac{\pi^2}{8} & \text{if } s = 0 \\ C_2\,|T - s| & \text{if } 0 < s < T \\ 0 & \text{if } T \leqslant s \end{cases}, \tag{54}$$

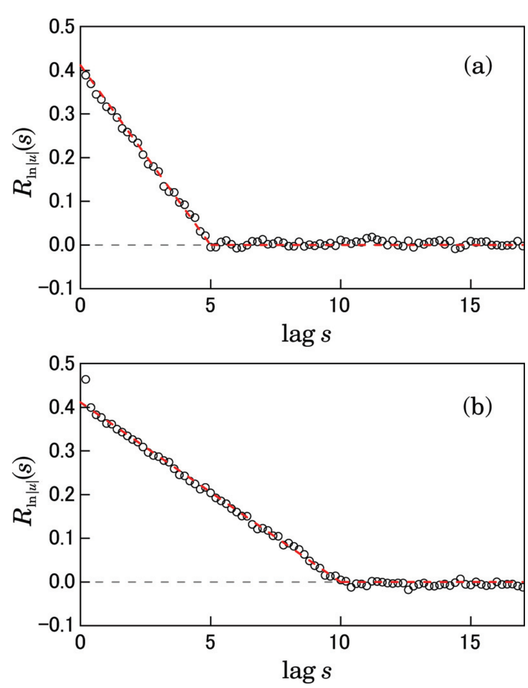

FIG. 6. (Color online) The log-amplitude autocovariance of χ^2 superstatistical processes in which $n = 2$, $\beta_0 = 7$, and $\gamma = 10$. (a) $T = 5$. (b) $T = 10$. The estimates (open circles) from a sample time series are in good agreement with the theoretical prediction (dashed lines) of Eq. (54).

where C_2 is the second log-amplitude cumulant of $\{U(t)\}$. To derive Eq. (54), a Gaussian white noise process under a constant β_j is assumed. As shown in Fig. 6, the estimated $R_{\ln|u|}(s)$ are in good agreement with the theoretical prediction of Eq. (54). Hence, the time scale T can be evaluated as the time lag where $R_{\ln|u|}(s)$ reaches zero. On the other hand, if an observed time series $\{u_i\}$ is a realization of a stochastic process described by independent identically distributed random variables, we obtain $R_{\ln|u|}(s) = 0$ for $s \neq 0$.

B. Analysis of real-world time series

As an application of our approach, we study non-Gaussian properties of real-world time series observed in a stock index futures market. The data are the historical data of Nikkei 225 futures for the 3-yr period from 1 August 2005 to 31 July 2008, with a sampling frequency of 1-min intervals, as shown in Fig. 7. The Nikkei 225 futures market is a form of futures contract where the underlying commodity is the Nikkei 225 stock average. In this market, a futures contract is an agreement to buy or sell a standardized value of the stock index, on a future date at a specified price. For Nikkei 225 futures, the contract unit is defined as the value of the Nikkei 225 $\times$ 1000 yen. In addition, the minimum fluctuation has been set to 10 yen, causing each tick to cause a fluctuation in price of 10 000 yen.

In our analysis, we study the price fluctuations $\{r_{\Delta t}(t)\}$ defined as the log return:

$$r_{\Delta t}(t) = \ln \frac{y(t+\Delta t)}{y(t)}, \tag{55}$$

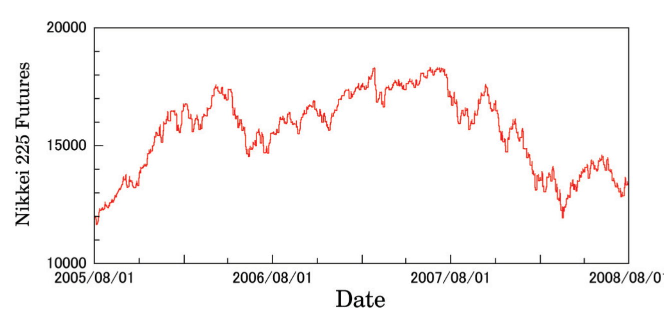

FIG. 7. (Color online) Historical data of Nikkei 225 futures for the 3-yr period, from 1 August 2005 to 31 July 2008.

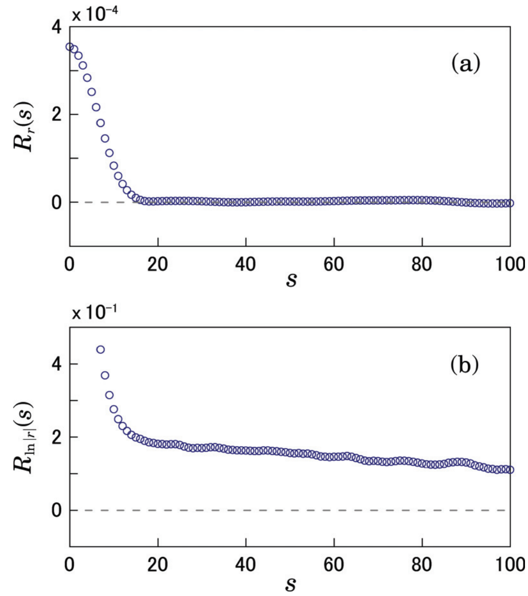

FIG. 8. (Color online) (a) Autocovariance $R_r(s)$ [Eq. (56)] and (b) log-amplitude autocovariance $R_{\ln|r|}(s)$ [Eq. (53)] estimated from $\{r_{\Delta t}(t)\}$ of Nikkei 225 futures.

where $y(t)$ denotes the price of the Nikkei 225 futures at time t. The total number of data points is about 2.5×10^5.

As already mentioned, the time-scale separation between the local equilibrium relaxation and slow driving of the intensity parameter is assumed in superstatistical processes. To evaluate the local equilibrium relaxation time scale, we employ the autocovariance of $\{U(t)\}$ defined as

$$R_u(s) = \langle U(t)\, U(t+s) \rangle\,, \qquad (56)$$

where $\langle U(t) \rangle = 0$ is assumed [11]. In contrast, to evaluate the driving time scale of log-amplitude fluctuations, we employ the log-amplitude autocovariance $R_{\ln|u|}(s)$ [Eq. (53)].

The estimated $R_u(s)$ and $R_{\ln|u|}(s)$ are shown in Fig. 8. In the case of Nikkei futures $\{r_{10\mathrm{min}}(t)\}$, the autocovariance $R_r(s)$ decayed to near zero at $s \approx 10$ min [Fig. 8(a)], where the positive correlation was observed only in the overlapping interval of the log returns within $\Delta t = 10$ min. This implies that the local equilibrium relaxation is much faster than exponential decay and that its relaxation time scale is shorter than 10 min. Compared to this time scale, the log-amplitude autocovariance $R_{\ln|r|}(s)$ exhibited long range persistent correlation, shown in Fig. 8(b). The strong correlations observed in the log-amplitude fluctuation indicate heterogeneous and clustered behavior of the variance of the time series $\{r_{10\mathrm{min}}(t)\}$. Hence, the presumed main cause for the emergence of non-Gaussian distributions is the variance heterogeneity.

Based on the decay time observed in $R_u(s)$ [Fig. 8(a)], we analyzed the log-return time series on scales equal to or larger than 10 min. Estimated log-amplitude cumulants are shown in Table I. If we estimate the parameter n for χ^2 superstatistical distributions, the estimated values were close to 2.0, as shown in Table I. Moreover, as shown in Fig. 9, the observed PDFs were well approximated by the χ^2 superstatistical distribution with $n = 2$.

TABLE I. The first three log-amplitude cumulants C_k and shape parameter n in χ^2 superstatistical distributions estimated from log returns $\{r_{\Delta t}(t)\}$ of Nikkei 225 futures, where $\{r_{\Delta t}(t)\}$ was not standardized and n was calculated from C_2.

Δt	C_1	C_2	C_3	n
10 min	−4.52	0.510	−0.486	1.73
15 min	−4.23	0.428	0.076	1.95
20 min	−4.03	0.387	0.371	2.09
25 min	−3.88	0.363	0.534	2.18
30 min	−3.76	0.365	0.557	2.17

As we discussed before, the χ^2 and inverse χ^2 superstatistical distributions with $n = 2$ are special cases where the log-amplitude fluctuations are subjected to extreme value distributions. Hence, the appearance of such special PDFs may be explained by a mechanism involving the asymptotic limit distribution of extreme values. However, our finding is not sufficient to confirm this point. Thus, further studies are required to understand the fundamental mechanism.

V. SUMMARY AND DISCUSSION

We studied superstatistical processes including log-amplitude fluctuations [Eq. (16)]. Using this framework, we showed that χ^2 and inverse χ^2 superstatistics with $n = 2$ are closely related to extreme value distributions. In these cases, the marginal distributions result in the q-Gaussian distribution with $q = 5/3$ (also known as the Student t distribution with degree 2) and the bilateral exponential distribution, respectively. This finding may help us understand the asymptotic appearance of these superstatistical distributions.

In addition, to characterize superstatistical processes, we proposed log-amplitude cumulants and log-amplitude autocovariance. The advantage of the log-amplitude statistics is that the log-amplitude cumulants can provide closed-form expressions for χ^2, inverse χ^2, and log-normal superstatistical distributions. As we demonstrated in numerical and real-world examples, the log-amplitude cumulants can be

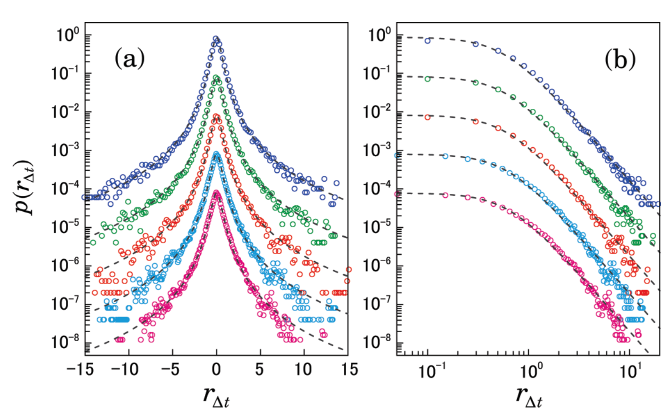

FIG. 9. (Color online) Probability density functions estimated from $\{r_{\Delta t}(t)\}$ of Nikkei 225 futures, where $\Delta t =$ 10, 15, 20, 25, and 30 min from top to bottom. In (a, b), the χ^2 superstatistical distributions with $n = 2$ are represented by the dashed lines.

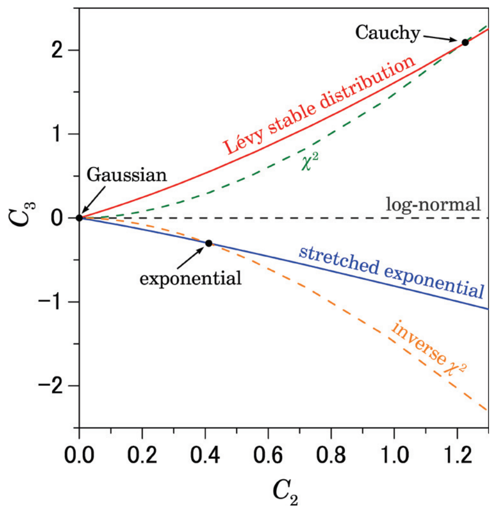

FIG. 10. (Color online) The (C_2, C_3) plane for Lévy stable distributions [Eq. (57)], stretched exponential distributions [Eq. (58)], and superstatistical distributions (dashed lines). The Cauchy distribution (Cauchy) lies at the intersection of the Lévy stable distribution and χ^2 superstatistical distribution families. The exponential distribution (exponential) lies at the intersection of the stretched exponential distribution and inverse χ^2 superstatistical distribution families. The origin $(C_2,C_3) = (0,0)$ corresponds to a Gaussian distribution (Gaussian).

directly estimated from the observed time series. Beck-Cohen superstatistics [2,10,11] has been very successful in describing non-Gaussian fluctuations. Including such examples, our method can be used to characterize non-Gaussian fluctuations observed in a wide variety of real-world signals.

While in this paper we focused on superstatistical models, it is important to note that the log-amplitude cumulants are applicable to a large class of symmetric non-Gaussian distributions [21]. To discuss some general features of the log-amplitude statistics, we consider the symmetric Lévy stable distribution:

$$P_{\text{Lévy}}(x) = \frac{1}{\pi}\int_0^\infty \exp\left(-\xi^\alpha q^\alpha\right)\cos(qx)\,dq, \tag{57}$$

where $0 < \alpha \leqslant 2$ and $\xi > 0$, and the two-tailed stretched exponential distribution:

$$P_{\exp}(x) = \frac{\alpha\xi^{1/\alpha}}{2\Gamma(1/\alpha)}\exp\left(-\xi\,|x|^\alpha\right), \tag{58}$$

where $0 < \alpha \leqslant 2$ and $\xi > 0$. In both cases α is the shape parameter and ξ is the scale parameter. When $\alpha = 2$, Eqs. (57) and (58) coincide with a Gaussian distribution. Note that here we consider independent and identically distributed random variables, different from superstatistical models with variance heterogeneity.

In the case of the Lévy stable distribution [Eq. (57)] with $0 < \alpha < 2$, the variance and higher moments are infinite. In contrast, all of the log-amplitude cumulants are finite and given by closed-form expressions. For instance, the second and third log-amplitude cumulants of $P_{\text{Lévy}}(x)$ are, respectively, given by

$$C_2 = \left(\frac{1}{6\alpha^2} - \frac{1}{24}\right)\pi^2, \tag{59}$$

$$C_3 = \left(\frac{2}{\alpha^3} - \frac{1}{4}\right)\zeta(3). \tag{60}$$

In the case of the stretched exponential distribution [Eq. (58)], all of the log-amplitude cumulants are also given by closed-form expressions, such as

$$C_2 = \frac{1}{\alpha^2}\,\psi^{(1)}\left(\frac{1}{\alpha}\right) - \frac{\pi^2}{8}, \tag{61}$$

$$C_3 = \frac{1}{\alpha^3}\,\psi^{(2)}\left(\frac{1}{\alpha}\right) + \frac{7}{4}\,\zeta(3). \tag{62}$$

The relations between C_2 and C_3 for the above distributions [Eqs. (57) and (58)], together with superstatistical distributions, are shown in Fig. 10. Because all of the log-amplitude cumulants beyond the first one are independent of the scale parameter ξ, the non-Gaussian properties depending on the shape parameter α are mainly characterized by C_2 and C_3. Namely, in a broad sense the deviation from a Gaussian shape can be quantified by C_2, and the thickness of tails compared to log-normal superstatistical distributions can be quantified by C_3. In the case of inverse χ^2 superstatistical distributions [Eq. (9)] the asymptotic behavior for large $|u|$ is evaluated as

$$p(u) \sim |u|^{n-2}\exp(-|u|). \tag{63}$$

When $n < 2$, this decay is faster than that of the stretched exponential distribution [Eq. (58)], $P_{\exp}(x) \sim \exp(-|x|^\alpha)$, with the same value of C_2. Therefore, as shown in Fig. 10, the smaller values of C_3 in the inverse χ^2 superstatistics imply a faster decay in the tails. As we have shown, it is possible to derive log-amplitude cumulants for a variety of symmetric non-Gaussian distributions. Hence, the log-amplitude statistics are applicable to parameter estimation and model selection for a large class of non-Gaussian distributions.

In our analysis of stock index futures, we demonstrated that the non-Gaussian distributions of the log-return time series were well approximated by the χ^2 superstatistical distribution with $n = 2$ [Fig. 9]. Of particular interest to us is the appearance of this distribution suggesting a link with the extreme value distribution for maxima. It is known that a stop-loss order is one of the most crucial orders in trading. As a reference for placing a stop loss, traders may use the biggest change in the high-low range over a fixed period or the lowest low (or highest high) of the preceding few days. Such trading strategies may result in the appearance of the extreme value distribution. However, our analysis is not sufficient to confirm this point. A detailed discussion of this problem is beyond the scope of this paper and will be discussed in future work.

ACKNOWLEDGMENTS

The author would like to thank Professors Taishin Nomura and Yasuyuki Suzuki for fruitful comments. This work was supported by Japan Society for the Promotion of Science KAKENHI Grant No. 23700544.

[1] R. Friedrich, J. Peinke, M. Sahimi, and M. R. Rahimi Tabar, Phys. Rep. **506**, 87 (2011).

[2] E. Van der Straeten and C. Beck, Phys. Rev. E **80**, 036108 (2009).

[3] K. Kiyono, J. Hayano, E. Watanabe, Z. R. Struzik, and Y. Yamamoto, Heart Rhythm **5**, 261 (2008).

[4] J. Hayano, K. Kiyono, Z. R. Struzik, Y. Yamamoto, E. Watanabe, P. K. Stein, L. L. W. J. A. Blumenthal, and R. M. Carney, Front. Physiol. **2**, 65 (2011).

[5] Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C.-K. Peng, and H. E. Stanley, Phys. Rev. E **60**, 1390 (1999).

[6] K. Kiyono, Z. R. Struzik, and Y. Yamamoto, Phys. Rev. Lett. **96**, 068701 (2006).

[7] P. Manshour, S. Saberi, M. Sahimi, J. Peinke, A. F. Pacheco, and M. R. Rahimi Tabar, Phys. Rev. Lett. **102**, 014101 (2009).

[8] B. A. Toledo, A. C.-L. Chian, E. L. Rempel, R. A. Miranda, P. R. Munoz, and J. A. Valdivia, Phys. Rev. E **87**, 022821 (2013).

[9] G. Wilk and Z. Włodarczyk, Phys. Rev. Lett. **84**, 2770 (2000).

[10] C. Beck and E. G. D. Cohen, Physica A **322**, 267 (2003).

[11] C. Beck, E. G. D. Cohen, and H. L. Swinney, Phys. Rev. E **72**, 056133 (2005).

[12] C. Tsallis and A. M. C. Souza, Phys. Rev. E **67**, 026106 (2003).

[13] F. Sattin, Eur. Phys. J. B **49**, 219 (2006).

[14] E. Van der Straeten and C. Beck, Phys. Rev. E **78**, 051101 (2008).

[15] C. Beck, Phys. Rev. Lett. **87**, 180601 (2001).

[16] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).

[17] F. Sattin and L. Salasnich, Phys. Rev. E **65**, 035106(R) (2002).

[18] Equation (9) can be derived using the following formula [I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series and Products* (Academic, New York, 1994), p. 368]:

$$\int_0^\infty x^{\nu-1} e^{-\frac{\beta}{x}-\gamma x}\, dx = 2\left(\frac{\beta}{\gamma}\right)^{\nu/2} K_\nu(2\sqrt{\beta\gamma}),$$

where β and γ are positive real numbers.

[19] B. Castaing, Y. Gagne, and E. J. Hopfinger, Physica D **46**, 177 (1990).

[20] K. Kiyono, Z. R. Struzik, and Y. Yamamoto, Phys. Rev. E **76**, 041113 (2007).

[21] K. Kiyono, Phys. Rev. E **79**, 031129 (2009).

[22] A. N. Kolmogorov, J. Fluid Mech. **13**, 82 (1962).

[23] E. J. Gumbel, *Statistics of Extremes* (Columbia University Press, New York, 1958).

[24] There are three types of extreme value distributions: the Gumbel, Fréchet, and Weibull distributions. The extreme value distributions arise as limiting distributions for maximums or minimums of a sample of independent identically distributed random variables, as the sample size increases. For instance, if random variables obey a Gaussian distribution, the Gumbel distribution appears as the corresponding limiting distribution. In addition, exponential, log-normal, and gamma distributions belong to the domain of attraction of the Gumbel distribution.

[25] W. Feller, *An Introduction to Probability Theory and its Applications*, 2nd ed. (Wiley, New York, 1971), Vol. 2.

[26] E. T. Jaynes, Phys. Rev. **106**, 620 (1957).