

A new life of Pearson's skewness

Yevgeniy Kovchegov *

Abstract

In this work we show how coupling and stochastic dominance methods can be successfully applied to a classical problem of rigorizing Pearson's skewness. Here, we use Fréchet means to define generalized notions of positive and negative skewness that we call *truly positive* and *truly negative*. Then, we apply stochastic dominance approach in establishing criteria for determining whether a continuous random variable is truly positively skewed. Intuitively, this means that scaled right tail of the probability density function exhibits strict stochastic dominance over equivalently scaled left tail. Finally, we use the stochastic dominance criteria and establish some basic examples of true positive skewness, thus demonstrating how the approach works in general.

1 Introduction

Consider a positively skewed (right-skewed) unimodal distribution with finite second moment. It is expected that the mode, the median, and the mean line up in an increasing order, i.e.,

$$\text{mode} < \text{median} < \text{mean}.$$

Symmetrically, for a negatively skewed (left-skewed) unimodal distribution with finite second moment,

$$\text{mean} < \text{median} < \text{mode}.$$

Here, when saying that a distribution is positively skewed we usually mean *Pearson's moment coefficient of skewness* (the standardized third central moment) is positive. However, there are notable exceptions to the above orderings, also known as the *mean-median-mode inequalities*. See [1, 17]. It all depends on how we measure skewness. Indeed, there are other measurements of skewness besides the moment coefficient such as Pearson's *first skewness coefficient* also known as *mode skewness* defined as

$$\frac{\text{mean} - \text{mode}}{\text{standard deviation}}$$

and Pearson's *second skewness coefficient* also known as *median skewness* defined as

$$3 \times \frac{\text{mean} - \text{median}}{\text{standard deviation}}.$$

*Department of Mathematics, Oregon State University, Corvallis, OR 97331, USA
Email: kovchegy@math.oregonstate.edu

The sign of the latter two measurements of skewness is consistent with the above discussed ordering of the mode, the median, and the mean.

Thus, whether the signs of the mode skewness, the median skewness, and the moment coefficient of skewness align depends on how we define skewness. First, we need an approach that unifies the different measurements of skewness in determining the sign of skewness (this is done in Def. 3 and 4 below). Second, when defining positive skewness, the objective is to characterize the distributions for which the left tail is “spreading short” and the right tail is “spreading longer”. If looking at this problem from the stochastic dominance perspective, we want to dissect the distribution into the left and the right parts so that when we reflect the left part, align both at zero, and multiply each by a certain monotone function (positive constant times a power of x) so that each part is converted into a distribution over positive half-line. If the distribution obtained from the right part exhibits stochastic dominance over the distribution obtained from the left part, then this should imply that the left tail is “spreading short” and the right tail is “spreading longer”. Hence, such distribution is positively skewed (reverse left and right for the negatively skewed distributions). Thus it is natural to consider finding criteria for the skewness to be positive or negative that is based on stochastic dominance of one tail over the other. Definition 3 of *true positive skewness* and its variation Def. 4 for unimodal distributions yield such stochastic dominance criteria which we establish in Sect. 2, and apply in Sect. 3.

We will generalize positive and negative skewness by consider the following class of centroids known as Fréchet p -means [4].

Definition 1. For $p \in [1, \infty)$ and a random variable X with the finite p -th moment, the quantity

$$\nu_p = \operatorname{argmin}_{a \in \mathbb{R}} E[|X - a|^p] \quad (1)$$

is called **Fréchet p -mean**, or simply the **p -mean**.

The theoretical p -mean ν_p is uniquely defined for all $p \geq 1$ as $E[|X - a|^p]$ is a strictly convex function of a . Moreover, the p -mean ν_p is a unique solution of

$$E[(X - \nu_p)_+^{p-1}] = E[(\nu_p - X)_+^{p-1}]. \quad (2)$$

Notice that (2) can be used as an extended definition of the p -mean that only requires finiteness of the $(p - 1)$ -st moment.

Definition 2. For $p \in [1, \infty)$ and a continuous random variable X with the finite $(p - 1)$ -st moment, the unique solution ν_p of (2) is the **p -mean** of X .

For the rest of the paper the **p -mean** ν_p of X is as defined in Def. 2, and we let

$$\mathcal{D} = \{p \geq 1 : E[|X|^{p-1}] < \infty\}$$

denote the domain of ν_p . Next, we observe that ν_1 and ν_2 are respectively the median (in continuous case) and the mean of X . If the distribution of X is unimodal, we let ν_0 denote the mode. In the unimodal case, we will use the domain $\mathcal{D}_0 = \mathcal{D} \cup \{0\}$.

For $p \in (0, 1)$, there are examples of no uniqueness of ν_p in (1). For instance, if X is a Bernoulli random variable with parameter $1/2$, there will be two values of ν_p for each $p \in (0, 1)$. In Sect. 4, we will observe that often Definition 2 can be extended to include $p \in (0, 1)$.

Next, we observe that the p -means ν_0 , ν_1 , ν_2 , and ν_4 are definitive for the notion of *positive skewness* defined via the Pearson's first and second skewness coefficients, as well as the Pearson's moment coefficient of skewness. Indeed, Karl Pearson's first skewness coefficient (mode skewness) for a unimodal random variable is expressed as $\frac{\nu_2 - \nu_0}{\sigma}$, where σ denotes the standard deviation, while Pearson's second skewness coefficient (median skewness) is given by $\frac{3(\nu_2 - \nu_1)}{\sigma}$. Finally, the renown Pearson's moment coefficient of skewness

$$\gamma = E \left[\left(\frac{X - \nu_2}{\sigma} \right)^3 \right]$$

is also related to the p -means via the following proposition.

Proposition 1. *Consider a random variable X with a finite third moment. The Pearson's moment coefficient of skewness γ is positive if and only if $\nu_4 > \nu_2$.*

Proof. Notice that the following cubic equation in a ,

$$E[(X - a)^3] = 0, \quad (3)$$

can be rewritten as

$$(a - \nu_2)^3 + 3\sigma^2(a - \nu_2) - \sigma^3\gamma = 0. \quad (4)$$

The 4-mean ν_4 is the only real root of (3), and thus of (4). Therefore, ν_4 is now being obtained from (4) as follows:

$$\gamma = \left(\frac{\nu_4 - \nu_2}{\sigma} \right)^3 + 3 \left(\frac{\nu_4 - \nu_2}{\sigma} \right). \quad (5)$$

□

Throughout the rest of the paper, we consider a continuous random variable X with density function $f(x)$. Furthermore, we suppose $f(x)$ has support $\text{supp}(f) = (L, R)$, where L may take value at $-\infty$ and R may take value at ∞ .

By Proposition 1, the second skewness coefficient and the moment coefficient of skewness are both positive if and only if

$$\nu_1 < \nu_2 < \nu_4.$$

In the case of unimodal continuous distribution, both, the moment coefficient of skewness is positive and the mean-median-mode inequality holds if and only if

$$\nu_0 < \nu_1 < \nu_2 < \nu_4.$$

See Fig. 1. This observation suggests the notion of *true positive skewness* defined below.

Definition 3. *We say that a random variable X , or its distribution, is truly positively skewed if and only if ν_p is an increasing function of p in the domain \mathcal{D} , provided the interior of \mathcal{D} is nonempty. Analogously, X is truly negatively skewed if and only if ν_p is a decreasing function of p in \mathcal{D} .*

The above defined *true positive skewness* insures $\nu_1 < \nu_2 < \nu_4$.

Definition 4. We say that a unimodal distribution is truly mode positively skewed if and only if ν_p is an increasing function of p in the domain \mathcal{D}_0 . Analogously, it is truly mode negatively skewed if and only if ν_p is a decreasing function of p in \mathcal{D}_0 .

Observe that *true mode positive skewness* guarantees $\nu_0 < \nu_1 < \nu_2 < \nu_4$. Notice also that a distribution can be *truly positively skewed* or *truly mode positively skewed* even in the absence of finite second moment.

Besides obtaining the sign of skewness, Definitions 3 and 4 do not immediately provide a way of measuring the magnitude of skewness. Defining the corresponding measures of skewness and extending the approach to multidimensional distributions is touched upon in the discussion section (Sect. 5). Importantly, this theoretical work concentrates on the problem of rigorously defining the sign of skewness. We do not intend to venture into statistical analysis or statistical applications of here defined concepts. In general, the question of centroids and their role in rigorously defining skewness that we considered in this paper has a long and interesting history; see for example [7, 16, 19, 20, 12, 5, 6, 3, 2]. Yet, this problem can still generate new challenges for theoretical probabilists and statisticians alike.

In this paper we will show how true positive skewness (and analogously, true negative skewness) can be validated using the methods of *coupling* and *stochastic dominance*. These methods are widely used in statistical mechanics and interacting particle systems, the theory of mixing times, and beyond. See [8, 9, 10, 11, 13] and references therein. Specifically we will need the following well known result.

Lemma 1 ([11, 13]). Suppose X and Y are real valued random variables with cumulative distribution functions denoted by F_X and F_Y respectively and satisfying $F_X(x) \geq F_Y(x)$, i.e., Y exhibits stochastic dominance over X , then, for any increasing function $h : \mathbb{R} \rightarrow \mathbb{R}$ we have $E[h(Y)] \geq E[h(X)]$. Moreover, if $F_X \not\equiv F_Y$, i.e., Y exhibits strict stochastic dominance over X , and if $h(x)$ is strictly increasing, then $E[h(Y)] > E[h(X)]$.

Lemma 1 is usually proved via a coupling argument. If X and Y are continuous random variables, the inequality in Lemma 1 follows from the integration by parts. In some instances, instead of stating that one random variable exhibits (strict) stochastic dominance over another, it is more convenient to say that one distribution or p.d.f. exhibits (strict) stochastic dominance over the other.

Before presenting the general approach in Sect. 2, we show how stochastic dominance method can be used to establish true positive skewness of exponential random variables.

Example 1 (Exponential distribution). Consider an exponential random variable X with parameter $\lambda > 0$. Without loss of generality let $\lambda = 1$. Equation (2) implies

$$\int_0^{\nu_p} (\nu_p - x)^{p-1} e^{-x} dx = \int_{\nu_p}^{\infty} (x - \nu_p)^{p-1} e^{-x} dx.$$

This simplifies to

$$\int_0^{\nu_p} x^{p-1} e^x dx = \Gamma(p). \quad (6)$$

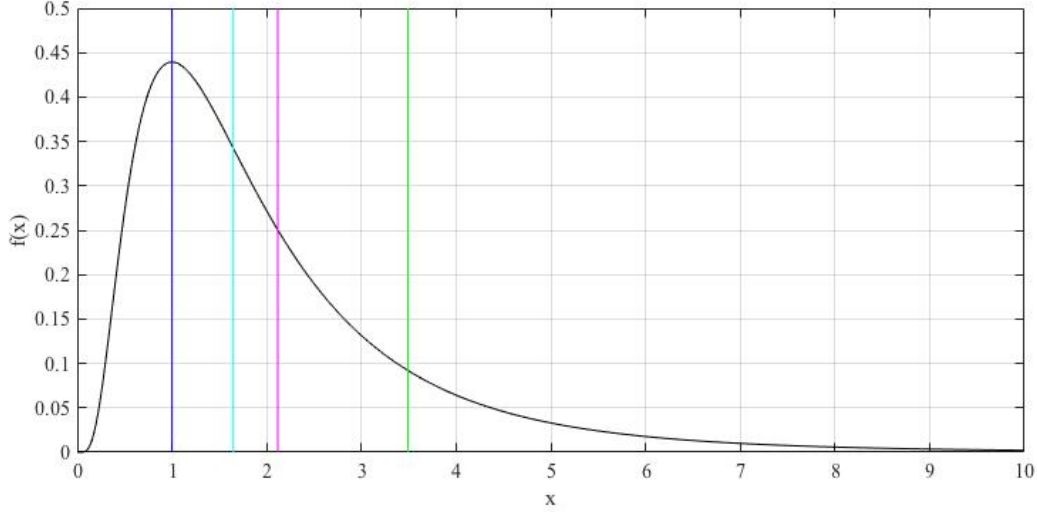


Figure 1: For the log-normal density function with parameters $\mu = \sigma^2 = 1/2$ and p.d.f. $f(x) = \frac{1}{\sqrt{\pi x}} \exp \{-(\log x - 1/2)^2\}$, we have the following centroids indicated on the plot: the mode $\nu_0 = \exp \{\mu - \sigma^2\} = 1$, the median $\nu_1 = \exp \{\mu\} = e^{1/2}$, the mean $\nu_2 = \exp \left\{ \mu + \frac{\sigma^2}{2} \right\} = e^{3/4}$, and the 4-mean $\nu_4 = \exp \left\{ \mu + \frac{3\sigma^2}{2} \right\} = e^{5/4}$.

Differentiating $\frac{d}{dp}$ in (6) yields

$$\nu_p^{p-1} e^{\nu_p} \frac{d\nu_p}{dp} + \int_0^{\nu_p} x^{p-1} e^x \log x \, dx = \int_0^\infty x^{p-1} e^{-x} \log x \, dx \quad (7)$$

Next, we observe that the gamma distribution with the p.d.f. $\frac{1}{\Gamma(p)} x^{p-1} e^{-x} \mathbf{1}_{(0,\infty)}(x)$ stochastically dominates the distribution with the p.d.f. $\frac{1}{\Gamma(p)} x^{p-1} e^x \mathbf{1}_{(0,\nu_p)}(x)$. Thus, Lemma 1 implies

$$\int_0^{\nu_p} x^{p-1} e^x \log x \, dx < \int_0^\infty x^{p-1} e^{-x} \log x \, dx \quad (8)$$

as $\log x$ is an increasing function. Consequently, equations (7) and (8) imply $\frac{d\nu_p}{dp} > 0$ for all $p > 0$. Hence, exponential random variables are proved to be *truly positively skewed* (Def. 3) and, as the mode $\nu_0 = 0$, *truly mode positively skewed* (Def. 4).

2 True positive skewness via stochastic dominance

For a given $p \in \mathcal{D}$, the theoretical p -mean ν_p defined in (2) solves

$$H_p := \int_0^{\nu_p - L} x^{p-1} f(\nu_p - x) \, dx = \int_0^{R - \nu_p} x^{p-1} f(\nu_p + x) \, dx. \quad (9)$$

Next, we state and prove a criterion for ν_p to be increasing, and therefore for the p.d.f. $f(x)$ to be truly positively skewed.

Theorem 1. Consider a continuous random variable with p.d.f. $f(x)$ supported on $\text{supp}(f) = (L, R)$, and a real number p in the interior of \mathcal{D} . If p.d.f. $\frac{1}{H_p} x^{p-1} f(\nu_p + x) \mathbf{1}_{(0, R-\nu_p)}(x)$ exhibits strict stochastic dominance over p.d.f. $\frac{1}{H_p} x^{p-1} f(\nu_p - x) \mathbf{1}_{(0, \nu_p-L)}(x)$, then function ν_p is increasing at p . Consequently, if the above stochastic dominance holds for all p in the interior of \mathcal{D} , the distribution is truly positively skewed (Def. 3).

Proof. For p in the interior of \mathcal{D} , we use Leibniz integral rule to obtain

$$\begin{aligned} 0 &= \frac{d}{dp} \left(\int_L^{\nu_p} (\nu_p - x)^{p-1} f(x) dx - \int_{\nu_p}^R (x - \nu_p)^{p-1} f(x) dx \right) \\ &= \left((p-1) \int_L^{\nu_p} (\nu_p - x)^{p-2} f(x) dx + \int_L^{\nu_p} (\nu_p - x)^{p-1} f(x) dx \right) \frac{d\nu_p}{dp} \\ &\quad + \left((p-1) \int_{\nu_p}^R (x - \nu_p)^{p-2} f(x) dx - \int_{\nu_p}^R (x - \nu_p)^{p-1} f(x) dx \right) \frac{d\nu_p}{dp} \\ &\quad + \int_0^{\nu_p-L} x^{p-1} \log x f(\nu_p - x) dx - \int_0^{R-\nu_p} x^{p-1} \log x f(\nu_p + x) dx, \end{aligned}$$

which simplifies to

$$\frac{d\nu_p}{dp} = \frac{\int_0^{R-\nu_p} x^{p-1} \log x f(\nu_p + x) dx - \int_0^{\nu_p-L} x^{p-1} \log x f(\nu_p - x) dx}{(p-1) \left[\int_0^{\nu_p-L} x^{p-2} f(\nu_p - x) dx + \int_0^{R-\nu_p} x^{p-2} f(\nu_p + x) dx \right]}. \quad (10)$$

Under the strict stochastic dominance assumption of the theorem, Lemma 1 implies

$$\int_0^{R-\nu_p} x^{p-1} \log x f(\nu_p + x) dx > \int_0^{\nu_p-L} x^{p-1} \log x f(\nu_p - x) dx. \quad (11)$$

Together, equations (10) and (11) imply $\frac{d\nu_p}{dp} > 0$ for all p in the interior of \mathcal{D} . \square

Remark 1. Returning to the discussion in the introduction of this paper (Sec. 1), Thm. 1 states that if we dissect $f(x)$ at ν_p into the left and the right parts so that when we reflect the left part, and align both at zero, and multiply each by $\frac{1}{H_p} x^{p-1}$, then each part is converted into a distribution over positive half-line, i.e., densities $\frac{1}{H_p} x^{p-1} f(\nu_p - x) \mathbf{1}_{(0, \nu_p-L)}(x)$ and $\frac{1}{H_p} x^{p-1} f(\nu_p + x) \mathbf{1}_{(0, R-\nu_p)}(x)$. If p.d.f. $\frac{1}{H_p} x^{p-1} f(\nu_p + x) \mathbf{1}_{(0, R-\nu_p)}(x)$ obtained from the right part exhibits stochastic dominance over p.d.f. $\frac{1}{H_p} x^{p-1} f(\nu_p - x) \mathbf{1}_{(0, \nu_p-L)}(x)$ obtained from the left part, then this should imply that the left tail is “spreading short” and the right tail is “spreading longer”. Hence, such distribution is truly positively skewed.

Importantly, this stochastic dominance argument remains valid and true positive skewness can be established even in the case of a distribution with infinite first moment and $\mathcal{D} \subseteq [1, 2)$. See Pareto distribution example in Sect. 3.4. This is one of the advantages of using Def. 3 and 4 in determining the sign of skewness.

Next, we present a criterion for the p.d.f. $\frac{1}{H_p} x^{p-1} f(\nu_p + x) \mathbf{1}_{(0, R-\nu_p)}(x)$ to exhibit strict stochastic dominance over the p.d.f. $\frac{1}{H_p} x^{p-1} f(\nu_p - x) \mathbf{1}_{(0, \nu_p-L)}(x)$.

Lemma 2. *Consider a continuous random variable supported over (L, R) with p.d.f. $f(x)$, and a real number $p \in \mathcal{D}$. Suppose there exists $c > 0$ such that $f(\nu_p - c) = f(\nu_p + c)$, and $f(\nu_p - x) > f(\nu_p + x)$ for $x \in (0, c)$, while $f(\nu_p - x) < f(\nu_p + x)$ for $x > c$. Suppose also that $\nu_p - L \leq R - \nu_p$. Then, a random variable with p.d.f. $\frac{1}{H_p} x^{p-1} f(\nu_p + x) \mathbf{1}_{(0, R-\nu_p)}(x)$ exhibits strict stochastic dominance over a random variable with p.d.f. $\frac{1}{H_p} x^{p-1} f(\nu_p - x) \mathbf{1}_{(0, \nu_p-L)}(x)$.*

Proof. Observe that c is the maximum and the only local extrema of

$$\int_0^x \frac{1}{H_p} y^{p-1} f(\nu_p - y) dy - \int_0^x \frac{1}{H_p} y^{p-1} f(\nu_p + y) dy$$

for $x \in [0, \nu_p - L]$, while

$$\int_0^{\nu_p-L} \frac{1}{H_p} y^{p-1} f(\nu_p - y) dy - \int_0^{\nu_p-L} \frac{1}{H_p} y^{p-1} f(\nu_p + y) dy \geq 0,$$

where $\int_0^{\nu_p-L} \frac{1}{H_p} y^{p-1} f(\nu_p - y) dy = 1$. Hence,

$$\int_0^x \frac{1}{H_p} y^{p-1} f(\nu_p - y) dy - \int_0^x \frac{1}{H_p} y^{p-1} f(\nu_p + y) dy > 0$$

for all $x \in (0, R - \nu_p)$. □

The following simple criterion follows immediately from the definition of stochastic dominance.

Proposition 2. *Consider a continuous random variable supported over (L, R) with p.d.f. $f(x)$. Suppose $f(x)$ is a decreasing function for $x \in (L, R)$. Then, for all $p \in \mathcal{D}$, density function $\frac{1}{H_p} x^{p-1} f(\nu_p + x) \mathbf{1}_{(0, R-\nu_p)}(x)$ exhibits strict stochastic dominance over density function $\frac{1}{H_p} x^{p-1} f(\nu_p - x) \mathbf{1}_{(0, \nu_p-L)}(x)$.*

3 Examples of true positive skewness

In this section, we will use stochastic dominance method and the toolbox developed in Sect. 2 for establishing true positive skewness for the gamma distribution, the unimodal beta distribution with the mode in the first half interval, the log-normal distribution, and Pareto distribution.

3.1 Gamma distribution

Consider a gamma random variable with p.d.f.

$$f(x) = \frac{1}{\Gamma(\alpha)} \lambda^\alpha x^{\alpha-1} e^{-\lambda x}$$

over $(L, R) = (0, \infty)$, with parameters $\alpha > 0$ and $\lambda > 0$. We will consider cases $\alpha > 1$ and $0 < \alpha < 1$ separately.

Case i: $\alpha > 1$. Here, the mode equals $\nu_0 = \frac{\alpha-1}{\lambda}$. Differentiating

$$\log f(\nu_p - x) - \log f(\nu_p + x) = 2\lambda x + (\alpha - 1) \log(\nu_p - x) - (\alpha - 1) \log(\nu_p + x)$$

with respect to x and setting it equal to zero yields

$$x^2 = \nu_p^2 - \frac{(\alpha - 1)\nu_p}{\lambda} = \nu_p(\nu_p - \nu_0). \quad (12)$$

Suppose $0 < \nu_p \leq \nu_0$ for some p , then (12) has no positive real solution. Thus, since $f(0) = 0 < f(2\nu_p)$, we have $f(\nu_p - x) < f(\nu_p + x)$ for all $x > 0$. Therefore, assumption $0 < \nu_p \leq \nu_0$ contradicts (9).

Therefore, $\nu_p > \nu_0$ for all $p \in \mathcal{D}$. Function $\frac{f(\nu_p - x)}{f(\nu_p + x)}$ equals 1 at $x = 0$ and equals 0 at $x = \nu_p$ with the only extremum at $c^* = \sqrt{\nu_p(\nu_p - \nu_0)}$. Since by (9), function $\frac{f(\nu_p - x)}{f(\nu_p + x)}$ cannot be ≤ 1 for all $x > 0$, the extremum c^* is the location of the maximum of $\frac{f(\nu_p - x)}{f(\nu_p + x)}$. Function $\frac{f(\nu_p - x)}{f(\nu_p + x)}$ increases on $[0, c^*)$, and decreases on (c^*, ν_p) . Thus, there is a point $c \in (c^*, \nu_p)$ such that $\frac{f(\nu_p - c)}{f(\nu_p + c)} = 1$ and the conditions in Lemma 2 are satisfied.

We conclude that $\frac{1}{H_p} x^{p-1} f(\nu_p + x) \mathbf{1}_{(0, R - \nu_p)}(x)$ exhibits strict stochastic dominance over $\frac{1}{H_p} x^{p-1} f(\nu_p - x) \mathbf{1}_{(0, \nu_p - L)}(x)$. Hence, $f(x)$ is *truly mode positively skewed* (Def. 4) by Thm. 1.

Case ii: $0 < \alpha < 1$. Here, $f(x)$ is *truly positively skewed* (Def. 3) by Prop. 2 and Thm. 1.

3.2 Beta distribution

Consider a beta random variable with p.d.f.

$$f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}$$

over $(L, R) = (0, 1)$, with parameters $\beta > \alpha > 1$. Here, $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ denotes the *beta function*. The mode equals $\nu_0 = \frac{\alpha-1}{\alpha+\beta-2} < \frac{1}{2}$. Differentiating

$$\begin{aligned} \log f(\nu_p - x) - \log f(\nu_p + x) &= (\alpha - 1) \log(\nu_p - x) - (\alpha - 1) \log(\nu_p + x) \\ &\quad + (\beta - 1) \log(1 - \nu_p + x) - (\beta - 1) \log(1 - \nu_p - x) \end{aligned}$$

with respect to x and setting it equal to zero, we obtain

$$(\alpha - 1) \frac{\nu_p}{\nu_p^2 - x^2} = (\beta - 1) \frac{1 - \nu_p}{(1 - \nu_p)^2 - x^2},$$

yielding

$$x^2 = -(1 - \nu_p)\nu_p \frac{\alpha - 1 - (\alpha + \beta - 2)\nu_p}{\beta - 1 - (\alpha + \beta - 2)\nu_p} = (1 - \nu_p)\nu_p \frac{\nu_p - \nu_0}{(1 - \nu_0) - \nu_p}. \quad (13)$$

Recall that $\nu_0 < \frac{1}{2}$. Suppose $0 < \nu_p \leq \nu_0$ for $p > 0$, then (13) has no positive real solution. Thus, since $f(\nu_p - \nu_p) = f(0) = 0 < f(2\nu_p) = f(\nu_p + \nu_p)$, we have $f(\nu_p - x) < f(\nu_p + x)$ for all $x \in (0, 1 - \nu_p)$. Therefore, assumption $0 < \nu_p \leq \nu_0$ contradicts (9). Hence, $\nu_p > \nu_0$.

Next, suppose $\nu_p \geq \frac{1}{2}$ for $p > 0$, then $1 - \nu_p \leq \nu_p$ and

$$\left(\frac{1 - \nu_p + x}{1 - \nu_p - x} \right)^{\beta-1} > \left(\frac{\nu_p + x}{\nu_p - x} \right)^{\alpha-1} \quad \forall x \in (0, 1 - \nu_p).$$

Thus, $f(\nu_p - x) > f(\nu_p + x)$ for all $x \in (0, \nu_p)$. Hence, assumption $\nu_p \geq \frac{1}{2}$ also contradicts equation (9).

Consequently, $\nu_0 < \nu_p < \frac{1}{2}$ for all $p \in \mathcal{D}$, and $c^* = \sqrt{(1 - \nu_p)\nu_p \frac{\nu_p - \nu_0}{(1 - \nu_0) - \nu_p}}$ is the only extremum of $\frac{f(\nu_p - x)}{f(\nu_p + x)}$ in $(0, \nu_p)$. Now, $\frac{f(\nu_p - 0)}{f(\nu_p + 0)} = 1$ and $\frac{f(\nu_p - \nu_p)}{f(\nu_p + \nu_p)} = \frac{f(0)}{f(2\nu_p)} = 0$. Since by (9), $\frac{f(\nu_p - x)}{f(\nu_p + x)}$ cannot be ≤ 1 for all $x > 0$, the extremum c^* is the point of the maximum of $\frac{f(\nu_p - x)}{f(\nu_p + x)}$. Therefore, there is a point $c \in (c^*, \nu_p)$ such that $\frac{f(\nu_p - c)}{f(\nu_p + c)} = 1$ and the conditions in Lemma 2 are satisfied.

We conclude that $\frac{1}{H_p} x^{p-1} f(\nu_p + x) \mathbf{1}_{(0, R - \nu_p)}(x)$ exhibits strict stochastic dominance over $\frac{1}{H_p} x^{p-1} f(\nu_p - x) \mathbf{1}_{(0, \nu_p - L)}(x)$. Thus, $f(x)$ is *truly mode positively skewed* (Def. 4) by Thm. 1.

3.3 Log-normal distribution

Consider a log-normal random variable with p.d.f.

$$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(\log x - \mu)^2}{2\sigma^2} \right\}$$

over $(L, R) = (0, \infty)$, with parameters μ and σ^2 . Here, the mode equals $\nu_0 = \exp \{\mu - \sigma^2\}$, the median is $\nu_1 = \exp \{\mu\}$, and the mean $\nu_2 = \exp \left\{ \mu + \frac{\sigma^2}{2} \right\}$. Next, we find ν_p and conclude that ν_p increases as a power of p . See Fig. 1.

Theorem 2. *For a log-normal random variable with parameters μ and σ^2 ,*

$$\nu_p = \exp \left\{ \mu + \frac{p-1}{2} \sigma^2 \right\}$$

solves (9) for all $p \in (0, \infty)$.

Proof. Without loss of generality, let $\mu = 0$. We need to show that $\nu_p = \exp \left\{ \frac{p-1}{2} \sigma^2 \right\}$ with $f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(\log x)^2}{2\sigma^2} \right\}$ satisfy (9) for all $p \in (0, \infty)$.

First, letting $z = \log \nu_p - \log(\nu_p - x)$ in the left hand side of (9), we have

$$\begin{aligned} \int_0^{\nu_p} x^{p-1} f(\nu_p - x) dx &= \frac{1}{\sqrt{2\pi\sigma^2}} \nu_p^{p-1} \int_0^\infty (1 - e^{-z})^{p-1} \exp \left\{ -\frac{(z - \log \nu_p)^2}{2\sigma^2} \right\} dz \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \nu_p^{p-1} \int_0^\infty (e^z - 1)^{p-1} \exp \left\{ -\frac{z^2 + (\log \nu_p)^2}{2\sigma^2} + \left(\frac{\log \nu_p}{\sigma^2} + 1 - p \right) z \right\} dz. \end{aligned} \quad (14)$$

Next, we let $z = \log(\nu_p + x) - \log \nu_p$ in the right hand side of (9), obtaining

$$\begin{aligned} \int_0^\infty x^{p-1} f(\nu_p + x) dx &= \frac{1}{\sqrt{2\pi\sigma^2}} \nu_p^{p-1} \int_0^\infty (e^z - 1)^{p-1} \exp \left\{ -\frac{(z + \log \nu_p)^2}{2\sigma^2} \right\} dz \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \nu_p^{p-1} \int_0^\infty (e^z - 1)^{p-1} \exp \left\{ -\frac{z^2 + (\log \nu_p)^2}{2\sigma^2} - \frac{\log \nu_p}{\sigma^2} z \right\} dz. \end{aligned} \quad (15)$$

Hence, the integrals in (14) and (15) are equal if and only if $\nu_p = \exp \left\{ \frac{p-1}{2} \sigma^2 \right\}$. \square

Corollary 1. *A log-normal random variable is truly mode positively skewed (Def. 4).*

Observe that $\nu_2 = \exp \left\{ \mu + \frac{\sigma^2}{2} \right\}$ and $\nu_4 = \exp \left\{ \mu + \frac{3\sigma^2}{2} \right\}$ together with the standard deviation $\sqrt{\exp \{ \sigma^2 \} - 1} \exp \left\{ \mu + \frac{\sigma^2}{2} \right\}$ and the Pearson's moment coefficient of skewness $\gamma = (\exp \{ \sigma^2 \} + 2) \sqrt{\exp \{ \sigma^2 \} - 1}$ satisfy the equality in (5).

3.4 Pareto distribution

For a given parameter $\alpha > 0$, consider a random variable with p.d.f.

$$f(x) = \frac{\alpha}{x^{\alpha+1}}$$

over $(L, R) = (1, \infty)$. Here, the mode equals $\nu_0 = 1$, and $\mathcal{D} = [1, \alpha + 1)$. Equation (9) implies

$$\int_0^{\nu_p-1} \frac{x^{p-1}}{(\nu_p - x)^{\alpha+1}} dx = \int_0^\infty \frac{x^{p-1}}{(\nu_p + x)^{\alpha+1}} dx. \quad (16)$$

Substituting $z = x/\nu_p$ into (16), we have

$$Z_p := \int_0^{1-1/\nu_p} \frac{z^{p-1}}{(1-z)^{\alpha+1}} dz = \int_0^\infty \frac{z^{p-1}}{(1+z)^{\alpha+1}} dz, \quad (17)$$

where $Z_p = \alpha^{-1} \nu_p^{1+\alpha-p} H_p$. Thus, $\nu_p > 1 = \nu_0$. Now, since $\frac{z^{p-1}}{Z_p(1+z)^{\alpha+1}} \mathbf{1}_{(0,\infty)}(z)$ exhibits strict stochastic dominance over $\frac{z^{p-1}}{Z_p(1-z)^{\alpha+1}} \mathbf{1}_{(0,1-1/\nu_p)}(z)$, Lemma 1 implies

$$\int_0^\infty \log z \frac{z^{p-1}}{(1+z)^{\alpha+1}} dz > \int_0^{1-1/\nu_p} \log z \frac{z^{p-1}}{(1-z)^{\alpha+1}} dz. \quad (18)$$

Next, we differentiate both integrals in (17) with respect to p , obtaining

$$\nu_p^{\alpha-p} (\nu_p - 1)^{p-1} \frac{d\nu_p}{dp} + \int_0^{1-1/\nu_p} \log z \frac{z^{p-1}}{(1-z)^{\alpha+1}} dz = \int_0^\infty \log z \frac{z^{p-1}}{(1+z)^{\alpha+1}} dz.$$

Therefore, by (18),

$$\frac{d\nu_p}{dp} = \nu_p^{p-\alpha}(\nu_p - 1)^{1-p} \left(\int_0^\infty \log z \frac{z^{p-1}}{(1+z)^{\alpha+1}} dz - \int_0^{1-1/\nu_p} \log z \frac{z^{p-1}}{(1-z)^{\alpha+1}} dz \right) > 0.$$

Hence, $f(x)$ is *truly mode positively skewed*. Notice that for $\alpha \in (0, 1)$, the quantities ν_1 , ν_2 , and γ do not exist. Yet, using Def. 3 and stochastic dominance, we established the positive sign of skewness. Naturally, this was expected of a distribution with only right tail and no left tail. This sends us back to the discussion in Remark 1.

4 Extending to $p \in (0, 1)$

In this section we will show that Definition 2 of p -mean can often be extended to include all $p \in (0, 1)$. For instance, in the case when X is an exponential random variable, equation (6) defines ν_p *uniquely* for all real $p > 0$.

Similarly, in the case of a gamma random variable with parameters $\alpha, \lambda > 0$, equation (9) implies

$$\int_0^{\nu_p} x^{p-1} (\nu_p - x)^{\alpha-1} e^{-\lambda(\nu_p-x)} dx = \int_0^\infty x^{p-1} (\nu_p + x)^{\alpha-1} e^{-\lambda(\nu_p+x)} dx. \quad (19)$$

Substituting $z = x/\nu_p$ into (19), we obtain

$$Z_p := \int_0^1 z^{p-1} (1-z)^{\alpha-1} e^{\lambda\nu_p z} dz = \int_0^\infty z^{p-1} (1+z)^{\alpha-1} e^{-\lambda\nu_p z} dz, \quad (20)$$

where $Z_p = \nu_p^{1-\alpha-p} e^{\lambda\nu_p} H_p$. Thus, for a gamma distribution, equation (20) also defines ν_p *uniquely* for all $p > 0$.

For the log-normal random variables, equations (14) and (15) imply the *uniqueness* of ν_p for all $p > 0$. Moreover, Thm. 2 finds the close form expression for the unique ν_p that solves (2) for all $p > 0$. Finally, for the Pareto random variables, (17) also defines ν_p *uniquely* for all real $p > 0$.

Even for a Bernoulli random variable with parameter $1/2$ and $p \in (0, 1)$, ν_p defined as in Def. 2 is unique, while ν_p defined as in Def. 1 is not.

We will try to give an argument for extending Definition 2 to $p \in (0, 1)$ in a more general way. Suppose $f(x)$ is differentiable in (L, R) , then, by (9), we have

$$\begin{aligned} 0 &= \frac{d}{dp} \left(\int_0^{\nu_p-L} x^{p-1} f(\nu_p - x) dx - \int_0^{R-\nu_p} x^{p-1} f(\nu_p + x) dx \right) \\ &= \frac{d\nu_p}{dp} \left((\nu_p - L)^{p-1} f(L) + (R - \nu_p)^{p-1} f(R) + \int_0^{\nu_p-L} x^{p-1} f'(\nu_p - x) dx - \int_0^{R-\nu_p} x^{p-1} f'(\nu_p + x) dx \right) \\ &\quad + \int_0^{\nu_p-L} x^{p-1} \log x f(\nu_p - x) dx - \int_0^{R-\nu_p} x^{p-1} \log x f(\nu_p + x) dx, \end{aligned}$$

and therefore,

$$\frac{d\nu_p}{dp} = \frac{\int_0^{R-\nu_p} x^{p-1} \log x f(\nu_p + x) dx - \int_0^{\nu_p-L} x^{p-1} \log x f(\nu_p - x) dx}{(\nu_p - L)^{p-1} f(L) + (R - \nu_p)^{p-1} f(R) + \int_0^{\nu_p-L} x^{p-1} f'(\nu_p - x) dx - \int_0^{R-\nu_p} x^{p-1} f'(\nu_p + x) dx}. \quad (21)$$

Quantities $(\nu_p - L)^{p-1} f(L)$ and $(R - \nu_p)^{p-1} f(R)$ are interpreted as the corresponding left and right limits. This includes the case when $L = -\infty$ and the case when $R = \infty$.

Now, by the Picard-Lindelöf existence and uniqueness theorem [18], if the right hand side in (21) is a Lipschitz function in ν_p , then the solution ν_p of (21) extends uniquely to $p \in (0, 1)$.

So, if we can extend the Definition 2 of p -mean to all $p \in (0, 1)$, then we can extend the definition of true positive skewness (Def. 3). Here, we say that a random variable (or its distribution) is *truly positively skewed over the full domain* if ν_p is increasing for p in the domain $(0, 1) \cup \mathcal{D}$. In the unimodal case, we say that a distribution is *truly mode positively skewed over the full domain* if ν_p is increasing for p in the domain $[0, 1) \cup \mathcal{D}$.

In order to establish true positive skewness over the full domain, we need to show that $\frac{d\nu_p}{dp} > 0$ for all $p \in (0, 1) \cup \mathcal{D}$. Often, we can show that the numerator and the denominator in (21) are both positive. Notice that the criteria in Lemma 2 and Prop. 2 work for $p \in (0, 1)$ as well, establishing stochastic dominance of density $\frac{1}{H_p} x^{p-1} f(\nu_p + x) \mathbf{1}_{(0, R-\nu_p)}(x)$ over density $\frac{1}{H_p} x^{p-1} f(\nu_p - x) \mathbf{1}_{(0, \nu_p-L)}(x)$. Hence, Lem. 2 and Prop. 2 can be used to prove positivity of the numerator in the RHS of (21) via equation (11). Finally, we notice that the denominator in the RHS of (21) is positive if $\log f(x)$ is concave over (L, R) .

Lemma 3. *Consider a continuous random variable supported over (L, R) with p.d.f. $f(x)$ differentiable on $\text{supp}(f) = (L, R)$. Suppose Definition 2 of p -mean ν_p can be extended to $p \in (0, 1)$, and suppose $\log f(x)$ is concave. If a random variable with density function $\frac{1}{H_p} x^{p-1} f(\nu_p + x) \mathbf{1}_{(0, R-\nu_p)}(x)$ exhibits strict stochastic dominance over a random variable with density function $\frac{1}{H_p} x^{p-1} f(\nu_p - x) \mathbf{1}_{(0, \nu_p-L)}(x)$, then function ν_p is increasing at $p \in (0, 1)$.*

Proof. Suppose $\log f(x)$ is concave, then $\frac{f'(x)}{f(x)} = \frac{d}{dx} \log f(x)$ is nonincreasing, and

$$\begin{aligned} & \int_0^{\nu_p-L} x^{p-1} f'(\nu_p - x) dx - \int_0^{R-\nu_p} x^{p-1} f'(\nu_p + x) dx \\ &= \int_0^{\nu_p-L} x^{p-1} \frac{f'(\nu_p - x)}{f(\nu_p - x)} f(\nu_p - x) dx - \int_0^{R-\nu_p} x^{p-1} \frac{f'(\nu_p + x)}{f(\nu_p + x)} f(\nu_p + x) dx \\ &\geq \frac{f'(\nu_p)}{f(\nu_p)} \left(\int_0^{\nu_p-L} x^{p-1} f(\nu_p - x) dx - \int_0^{R-\nu_p} x^{p-1} f(\nu_p + x) dx \right) = 0 \end{aligned} \quad (22)$$

by (9). If $\log f(x)$ is not constant in (L, R) , then the inequality in (22) is strict. If $\log f(x)$ is constant in (L, R) , then L needs to be finite, and therefore

$$(\nu_p - L)^{p-1} f(L) > 0.$$

In either case,

$$(\nu_p - L)^{p-1} f(L) + (R - \nu_p)^{p-1} f(R) + \int_0^{\nu_p - L} x^{p-1} f'(\nu_p - x) dx - \int_0^{R - \nu_p} x^{p-1} f'(\nu_p + x) dx > 0. \quad (23)$$

□

Additionally, for $p = 1$, integration yields

$$f(L) + f(R) + \int_0^{\nu_1 - L} f'(\nu_1 - x) dx - \int_0^{R - \nu_1} f'(\nu_1 + x) dy = 2f(\nu_1) > 0,$$

whence, from equations (11) and (21), we have

$$\left. \frac{d\nu_p}{dp} \right|_{p=1} = \frac{1}{2f(\nu_1)} \left(\int_0^{R - \nu_1} \log x f(\nu_1 + x) dx - \int_0^{\nu_1 - L} \log x f(\nu_1 - x) dx \right) > 0. \quad (24)$$

Next, we go through some examples of true positive skewness over the full domain.

Example 2 (Exponential distribution). For an exponential random variable, equations (7) and (8) imply *true mode positive skewness over the full domain*.

Example 3 (Gamma distribution). Consider a gamma distribution. If $\alpha > 1$, then, the calculations in Sect. 3.2 yield Lemma 2 is satisfied for all $p \in (0, 1) \cup \mathcal{D}$. Therefore, density function $\frac{1}{H_p} x^{p-1} f(\nu_p + x) \mathbf{1}_{(0, R - \nu_p)}(x)$ exhibits strict stochastic dominance over density function $\frac{1}{H_p} x^{p-1} f(\nu_p - x) \mathbf{1}_{(0, \nu_p - L)}(x)$. Moreover,

$$\frac{d}{dx} \log f(x) = \frac{f'(x)}{f(x)} = \frac{\alpha - 1}{x} - \lambda$$

is a decreasing function over $(0, \infty)$ as $\alpha > 1$. Hence, $\log f(x)$ is concave, and Lem. 3 implies *true mode positive skewness over the full domain*.

If $0 < \alpha < 1$, equation (20) implies that $\frac{1}{Z_p} z^{p-1} (1 + z)^{\alpha-1} e^{-\lambda \nu_p z} \mathbf{1}_{(0, \infty)}(z)$ exhibits strict stochastic dominance over $\frac{1}{Z_p} z^{p-1} (1 - z)^{\alpha-1} e^{\lambda \nu_p z} \mathbf{1}_{(0, \nu_p)}(z)$, and therefore, Lemma 1 implies

$$\int_0^\infty (\log z) z^{p-1} (1 + z)^{\alpha-1} e^{-\lambda \nu_p z} dz > \int_0^1 (\log z) z^{p-1} (1 - z)^{\alpha-1} e^{\lambda \nu_p z} dz. \quad (25)$$

We differentiate both integrals in (20) with respect to p , obtaining

$$\begin{aligned} & \int_0^1 (\log z) z^{p-1} (1 - z)^{\alpha-1} e^{\lambda \nu_p z} dz + \frac{d\nu_p}{dp} \int_0^1 z^p (1 - z)^{\alpha-1} e^{\lambda \nu_p z} dz \\ &= \int_0^\infty (\log z) z^{p-1} (1 + z)^{\alpha-1} e^{-\lambda \nu_p z} dz - \frac{d\nu_p}{dp} \int_0^\infty z^p (1 + z)^{\alpha-1} e^{-\lambda \nu_p z} dz. \end{aligned}$$

Thus, for all $p > 0$, we have

$$\frac{d\nu_p}{dp} = \frac{\int_0^\infty (\log z) z^{p-1} (1+z)^{\alpha-1} e^{-\lambda\nu_p z} dz - \int_0^1 (\log z) z^{p-1} (1-z)^{\alpha-1} e^{\lambda\nu_p z} dz}{\int_{-1}^\infty |z|^p (1+z)^{\alpha-1} e^{-\lambda\nu_p z} dz} > 0$$

by (25). Hence, $f(x)$ is *truly positively skewed over the full domain*.

Example 4 (Beta distribution). For a beta distribution with parameters $\beta > \alpha > 1$, Sect. 3.2 yields Lemma 2 is satisfied. Hence, $\frac{1}{H_p} x^{p-1} f(\nu_p + x) \mathbf{1}_{(0, R-\nu_p)}(x)$ exhibits strict stochastic dominance over $\frac{1}{H_p} x^{p-1} f(\nu_p - x) \mathbf{1}_{(0, \nu_p-L)}(x)$. Next,

$$\frac{d}{dx} \log f(x) = \frac{f'(x)}{f(x)} = \frac{\alpha - 1}{x} - \frac{\beta - 1}{1 - x}$$

is a decreasing function over $(0, 1)$ as both, $\alpha > 1$ and $\beta > 1$. Hence, $\log f(x)$ is concave, and Lem. 3 implies *true mode positive skewness over the full domain*.

Example 5 (Log-normal distribution). Consider a log-normal distribution. Observe that Thm. 2 works for all real $p > 0$. Hence, a log-normal distribution is *truly mode positively skewed over the full domain*.

Example 6 (Pareto distribution). For a Pareto distribution, calculations in Sect. 3.4 stay valid for all $p \in (0, \alpha + 1)$. Thus, it is *truly mode positively skewed over the full domain*.

Notice that even in the case of a unimodal continuous random variable with probability density function (p.d.f.) in C^∞ , the mode ν_0 does not have to be equal $\lim_{p \downarrow 0} \nu_p$. Indeed, the case of log-normal random variables is an example of discontinuity of ν_p at 0 as

$$\nu_0 = \exp \left\{ \mu - \sigma^2 \right\} < \exp \left\{ \mu - \frac{\sigma^2}{2} \right\} = \lim_{p \downarrow 0} \nu_p. \quad (26)$$

5 Discussion

While the unification of the three classical characterizations of skewness introduced here is important in its own right, the main insight of true positive/negative skewness was articulated in Remark 1 following Theorem 1. Specifically, for $p \in \mathcal{D}$, let us break the probability density function $f(x)$ into the left and right parts, $f(\nu_p - x) \mathbf{1}_{(0, \nu_p-L)}(x)$ and $f(\nu_p + x) \mathbf{1}_{(0, R-\nu_p)}(x)$. We multiply the two parts by $\frac{1}{H_p} x^{p-1}$ to make the two probability density functions, $\frac{1}{H_p} x^{p-1} f(\nu_p - x) \mathbf{1}_{(0, \nu_p-L)}(x)$ and $\frac{1}{H_p} x^{p-1} f(\nu_p + x) \mathbf{1}_{(0, R-\nu_p)}(x)$. It is natural to expect that if $f(x)$ is positively skewed, the right tail p.d.f. would exhibit stochastic dominance over the left tail p.d.f. for all $p \in \mathcal{D}$. Theorem 1 asserts that this stochastic dominance holds if the distribution is truly positively skewed. The proposed principle of the dominating left tail p.d.f. over the right tail p.d.f. works regardless of the size of the domain \mathcal{D} , and the main advantage of this new approach is that true positive (or negative) skewness can be established even in the case of infinite first moment when all three of Pearson's skewness measures (and even their numerators) are undefined. This was demonstrated in Sect. 3.4

for Pareto distribution with $\alpha \in (0, 1)$, while the most compelling example was established in the work [14] of students advised by the author of this current manuscript, where it was shown that Lévy distribution is truly positively skewed. This result will be published in a separate paper.

Now, we will proceed by mentioning some precursors of the approach to skewness presented in this current paper. There is a relation between the results in [20] and this current paper. In [20], van Zwet proves that the mean-median-mode inequality

$$\text{mode} \leq \text{median} \leq \text{mean}$$

holds whenever the cumulative distribution function $F(x)$ satisfies

$$F(\nu_1 - x) + F(\nu_1 + x) \geq 1 \quad \forall x > 0. \quad (27)$$

Observe that for a continuous unimodal random variable with p.d.f. $f(x)$, condition (27) is equivalent to

$$\frac{1}{H_1} \int_0^x f(\nu_1 + y) dy \geq \frac{1}{H_1} \int_0^x f(\nu_1 - y) dy,$$

where $H_1 = \frac{1}{2}$ for the median ($p = 1$). In other words, the mean-median-mode inequality holds whenever p.d.f. $\frac{1}{H_1} f(\nu_1 + x) \mathbf{1}_{(0, R - \nu_1)}(x)$ exhibits stochastic dominance over p.d.f. $\frac{1}{H_1} f(\nu_1 - x) \mathbf{1}_{(0, \nu_1 - L)}(x)$. That is, the stochastic dominance condition required for the true positive skewness criteria established in Sect. 2 holds for $p = 1$. This connection of (27) to stochastic dominance was also noticed by Dharmadhikari and Joag-dev [5, 6]. Observe that in the unimodal case, the above result of van Zwet [20] (with strict inequalities) together with Thm. 1 imply that if $\frac{1}{H_p} x^{p-1} f(\nu_p + x) \mathbf{1}_{(0, R - \nu_p)}(x)$ exhibits strict stochastic dominance over $\frac{1}{H_p} x^{p-1} f(\nu_p - x) \mathbf{1}_{(0, \nu_p - L)}(x)$ for all $p \in \mathcal{D}$, then the distribution is *truly mode positively skewed*.

As we have seen, the notion of true skewness (i.e., monotonicity of ν_p) is based on the comparison of the left and right halves of a continuous distribution dissected at the centroids ν_p for $p \in \mathcal{D}$. As such, true mode positive skewness is a more restrictive property than positivity of all three Pearson's skewness metrics. For instance, in [1] there is an example of a unimodal continuous random variable with positive Pearson's moment coefficient of skewness γ and negative median skewness. Naturally, in this example $\nu_2 < \nu_1$ while Prop. 1 yields $\nu_2 < \nu_4$. Thus, the distribution is not truly positively or negatively skewed.

Finally, Oja [15] used convexity of order k notion to extend the convex transformation approach to skewness and kurtosis developed in van Zwet [19].

Next, we outline some potentially advantageous directions spinning out of this work. First, in this paper, Definition 3 only speaks of the sign of skewness (i.e., positive or negative). Yet, we would like to consider measures of skewness based on the trajectory of ν_p . One may consider the following approach generating a family of skewness measures. For a probability density function φ over $[0, \infty)$, let

$$\ell_\varphi = \frac{1}{\left. \frac{d\nu_p}{dp} \right|_{p=1}} \frac{\int_{\mathcal{D}} \varphi(p-1) d\nu_p}{\int_{\mathcal{D}} \varphi(p-1) dp}.$$

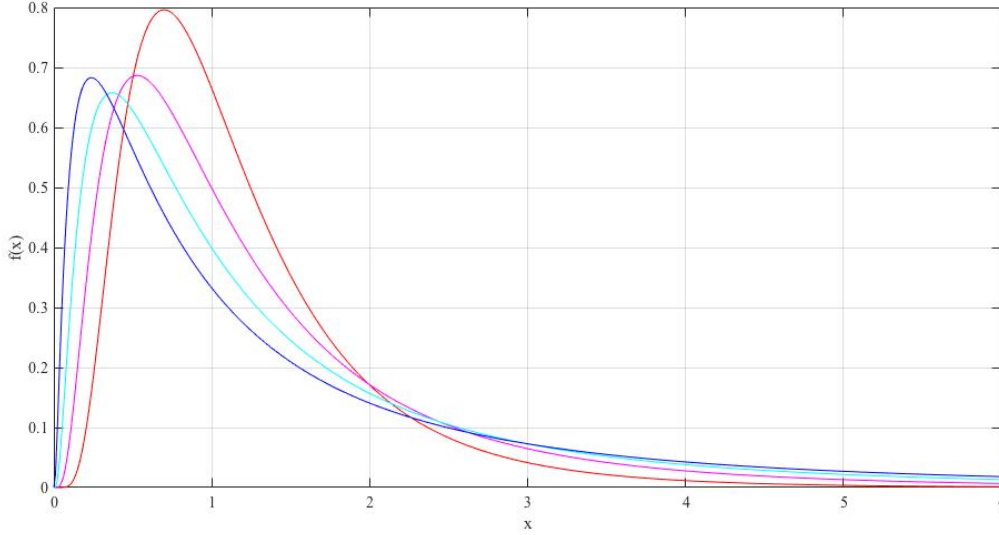


Figure 2: The log-normal density functions with different parameter values: $(\mu = 0, \sigma = 0.6)$ in red, $(\mu = 0, \sigma = 0.8)$ in magenta, $(\mu = 0, \sigma = 1)$ in light blue, $(\mu = 0, \sigma = 1.2)$ in dark blue. Evidently, as parameter μ does not affect the skewness, the distribution skews more to the right with greater values of σ .

Then, for truly positively/negatively skewed continuous distributions, positive quantities ℓ_φ , if finite, may be used to measure the magnitude of skewness.

In the case of log-normal distribution, Pearson's moment coefficient of skewness equals $\gamma = (\exp\{\sigma^2\} + 2) \sqrt{\exp\{\sigma^2\} - 1}$. Thus, γ is an increasing function of σ^2 , reflecting evident increase in skewness as one increases the value of σ^2 . At the same time, contrary to the observed evolution of log-normal density (see Fig. 2), Pearson's first and second skewness coefficients would decrease down to zero as $\sigma^2 \rightarrow \infty$. See [3] for a relevant discussion. Now, Thm. 2 yields

$$\ell_\varphi = \frac{1}{\left. \frac{d\nu_p}{dp} \right|_{p=1}} \frac{\int_{\mathcal{D}} \varphi(p-1) \frac{d\nu_p}{dp} dp}{\int_{\mathcal{D}} \varphi(p-1) dp} = \int_0^\infty \varphi(x) e^{x\sigma^2/2} dx.$$

Observe that in this case, if $\lim_{x \rightarrow \infty} \frac{-\ln \varphi(x)}{x} = \infty$, then ℓ_φ is a well-defined quantity that increases with increasing σ^2 .

Second, in a multidimensional case, for a random vector $X \in \mathbb{R}^d$, consider Fréchet p -means $\nu_p \in \mathbb{R}^d$ as defined in Def. 1. We believe that the trajectory of ν_p can be interpreted as a tailbone of the distribution, or the *trajectory of skewness*. Moreover, if the asymptotic limit

$$\zeta = \lim_{p \rightarrow \infty} \left(\frac{d\nu_p}{dp} \middle/ \left\| \frac{d\nu_p}{dp} \right\| \right),$$

exists, then ζ can be interpreted as the asymptotic direction of skewness in \mathbb{R}^d . Estimating this tailbone trajectory and its limit ζ can be done numerically.

Acknowledgements

The author would like to thank Anatoly Yambartsev, Ilya Zaliapin, Jordan Stoyanov, and the anonymous referee for their helpful comments and suggestions. This research was supported by FAPESP award 2018/07826-5 and by NSF award DMS-1412557.

References

- [1] K. M. Abadir, *The Mean-Median-Mode Inequality: Counterexamples*. Econometric Theory **21**, 477–482 (2005)
- [2] B. Abdous and R. Theodorescu, *Mean, Median, Mode IV* Statistica Neerlandica **52**(3), 356–359 (1998)
- [3] B. C. Arnold and R. A. Groeneveld, *Measuring skewness with respect to the mode* The American Statistician, 49(1), 34–38 (1995)
- [4] F. Barbaresco, *Information geometry of covariance matrix: Cartan-Siegel homogeneous bounded domains, Mostow/Berger fibration and Frechet median* In *Matrix information geometry* Springer, Berlin, Heidelberg, 199–255 (2013)
- [5] S. W. Dharmadhikari and K. Joag-dev, *Mean, Median, Mode III* Statistica Neerlandica **37**(4), 165–168 (1983)
- [6] S. W. Dharmadhikari and K. Joag-dev, *Unimodality, Convexity, and Applications* Academic Press, New York (1988)
- [7] C. F. Gauss, *Theoria Combinationis Observationum Erroribus Minimis Obnoxiae, Pars Prior* Commentationes Societatis Regiae Scientiarum Gottingensis Recentiores. 5. (1823)
- [8] F. den Hollander, *Probability theory: The coupling method* Leiden University, Lectures Notes-Mathematical Institute (2012)
- [9] Y. Kovchegov and P. T. Otto, *Path Coupling and Aggregate Path Coupling* Springer-Briefs in Probability and Mathematical Statistics, Springer (2018)
- [10] T. M. Liggett *Interacting Particle Systems* Springer (1985)
- [11] T. Lindvall, *Lectures on the coupling method* Dover Publications (2002). Originally published in 1992 by John Wiley & Sons, Inc., New York.
- [12] H. L. MacGillivray, *Skewness and asymmetry: measures and orderings* The Annals of Statistics **14** no.3, 994–1011 (1986).
- [13] A. Müller and D. Stoyan, *Comparison methods for stochastic models and risks* Vol. 389. Wiley (2002)
- [14] A. Negrón, C. Pertel, and C. Wang, *Extensions of True Positive Skewness for Unimodal Distributions* Oregon State University Research Experiences for Undergraduates (REU) proceedings (2021) https://sites.science.oregonstate.edu/math_reu/proceedings/REU_Proceedings/Proceedings2021/TeamKovchegov.pdf

- [15] H. Oja, *On Location, Scale, Skewness and Kurtosis of Univariate Distributions* Scandinavian Journal of Statistics **8**(3), 154–168 (1981)
- [16] K. Pearson, *Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material* Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **186**, 343–414 (1895)
- [17] J. M. Stoyanov, *Counterexamples in probability* (3rd ed.) Dover Publications (2014).
- [18] G. Teschl, *Ordinary differential equations and dynamical systems* Graduate Studies in Mathematics **140**, American Mathematical Society, Providence, RI (2012).
- [19] W. R. van Zwet, *Convex transformations: a new approach to skewness and kurtosis* Statistica Neerlandica **18**, 433–441 (1964)
- [20] W. R. van Zwet, *Mean, median, mode II* Statistica Neerlandica **33**(1), 1–5 (1979)