EXTENSIONS OF TRUE SKEWNESS FOR UNIMODAL DISTRIBUTIONS

YEVGENIY KOVCHEGOV, ALEX NEGRÓN, CLARICE PERTEL, AND CHRISTOPHER WANG

ABSTRACT. A 2022 paper [8] introduced the notion of true positive and negative skewness for continuous random variables via Fréchet p-means. In this work, we find novel criteria for true skewness and establish true skewness for the Weibull, Lévy, and skew-normal distributions. Furthermore, some relevant properties of the p-means of random variables are established.

CONTENTS

Date: July 9, 2023.
Key words and phrases. skewness, true skewness, Frechet mean, stochastic dominance, skew-normal distribution, Lévy distribution, Weibull distribution.

1. Introduction

A commonly accepted measure for the skewness of a probability distribution is given by *Pearson's moment coefficient of skewness*, or the standarized third central moment

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

where X has mean μ and variance σ^2. Usually, we say that a distribution is positively skewed if $\text{Skew}[X] > 0$ and negatively skewed if $\text{Skew}[X] < 0$. It is also expected that positively skewed distributions satisfy the *mean-median-mode inequalities*

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

while negatively skewed distributions satisfy the reverse inequalities

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

However, a distribution with positive moment coefficient of skewness does not always satisfy the mean-median-mode inequalities. For instance, the Weibull distribution with shape parameter $3.44 < \beta < 3.60$ has positive moment coefficient of skewness but satisfies the reversed mean-median-mode inequalities, corresponding to negative skew [6]. (For other counterexamples, see [2].) This discrepancy is important when comparing Pearson's moment coefficient with other measures of skewness, such as *Pearson's first skewness coefficient*

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

and *Pearson's second skewness coefficient*

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

The direction of skewness can therefore be inconsistent between different skewness measures. In 2022, Kovchegov [8] introduced the notion of *true positive and negative skewness* to unify Pearson's coefficients in determining the sign of skewness. It relies on the idea that for a positively skewed distributions, the left part of the distribution should stochastically dominate the right part, resulting in a left tail that "spreads shorter" and a right tail that "spreads longer." We use a class of centroids known as p-means. See [8].

Definition 1.1 (Kovchegov [8], 2022)**.** For $p \in [1, \infty)$ and random variable X with finite $(p-1)$-st moment, the ***p*-mean**, of X is the unique solution ν_p of the equation

$$E[(X-\nu_p)_+^{p-1}] = E[(\nu_p - X)_+^{p-1}]. \tag{1.1}$$

If moreover X has finite p-th moment, the notion of ***p*-mean** in (1.1) is equivalent to the **Fréchet *p*-mean** defined as

$$\nu_p = \underset{a\in\mathbb{R}}{\arg\min}\, E|X-a|^p. \tag{1.2}$$

For $p > 1$, the uniqueness of the ***p*-mean** ν_p follows from the fact that

$$E[(a-X)_+^{p-1}] - E[(X-a)_+^{p-1}]$$

is a strictly increasing continuous function of a. Occasionally, we write $\nu(p)$ to emphasize that ν_p is a function of p.

Notice that identically distributed random variables have the same p-means and that ν_1 and ν_2 correspond to the median and mean of the distribution, respectively. For a random variable X let

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

be the domain of ν_p. Notice that p is a real number; it does not have to be an integer. If X has a unique mode, then we denote it by ν_0. In this case, we include 0 in the set $\mathcal{D}_X$. We omit the subscript X when the random variable is unambiguous.

Remark 1.2. For distributions which are "nice enough," the domain $\mathcal{D}_X$ can be extended in a well-defined way to include values of p in the interval $(0, 1)$. The conditions for which this is permissible is discussed in [8]. In this work, we limit our consideration to $p \geq 1$.

Definition 1.3 (Kovchegov [8], 2022)**.** We say a random variable X (resp. its distribution and density) with a uniquely defined median ν_1 is **truly positively skewed** if ν_p is a strictly increasing function of p in $\mathcal{D}$, provided $\mathcal{D}$ has non-empty interior. Analogously, X is **truly negatively skewed** if ν_p is a strictly decreasing function of p in $\mathcal{D}$.

Remark 1.4. It is possible that, for a unimodal distribution, ν_p is strictly increasing only on $\mathcal{D} \setminus \{0\}$ and that there exists $p \in \mathcal{D} \setminus \{0\}$ such that $\nu_p < \nu_0$. Kovchegov [8] differentiates between this case, which is referred to in that work as true positive skewness, and the case where $\nu_p > \nu_0$ holds for all $p \in \mathcal{D}$, which is referred to as the stronger *true mode positive skewness*. Because we consider only unimodal distributions, for simplicity we take "true positive skewness" to mean true mode positive skewness in the sense of Kovchegov [8], unless explicitly mentioned otherwise.

Consider X that has unique mode and median. Naturally, Pearson's first skewness coefficient is positive if and only if $\nu_2 > \nu_0$ and Pearson's second skewness coefficient is positive if and only if $\nu_2 > \nu_1$. Additionally, it was noticed in [8] that Pearson's moment coefficient of skewness γ is positive if and only if $\nu_4 > \nu_2$. Now, true mode positive skewness guarantees $\nu_0 < \nu_1 < \nu_2 < \nu_4$ provided finiteness of the corresponding moments. Thus, Pearson's first and second skewness coefficients and Pearson's moment coefficient of skewness coincide in sign. Therefore, the direction of skewness is unified across several different criteria under true positive skewness. Moreover, Theorem 1.7 cited in Sect. 1.1 below gives direct geometric justification for the notion of true positive skewness as well as to the original Pearson's moment and second coefficients of skewness. See Remark 1.8.

Remark 1.5. An additional advantage of the notion of true positive skewness is that it allows us to characterize the skewness of distributions that have infinite integer moments. Indeed, each of Pearson's skewness coefficients requires at least a finite mean, which excludes a large class of heavy-tailed distributions from the classical study of skewness.

1.1. **Existing criteria for true skewness.** In general, demonstrating the true skewness of an arbitrary distribution requires a detailed analysis of the solutions of a class of integral equations (1.1). Such analysis is simplified by introducing arguments from stochastic ordering.

Definition 1.6 ([10])**.** For random variables X and Y, we say that X **stochastically dominates** Y (resp. the distribution of X stochastically dominates the distribution of Y) if the cumulative distribution function F_X of X is majorized by the cumulative distribution function F_Y of Y, i.e. if $F_X(x) \le F_Y(x)$ holds for every $x \in \mathbb{R}$. We say that the stochastic dominance is *strict* if there exists a point x_0 for which $F_X(x_0) < F_Y(x_0)$.

If X is a continuous random variable with a density function f, then we may rewrite (1.1) as an equality of integrals, the quantities of which we define as a normalizing term H_p, i.e.,

$$H_p := \int_0^{\nu_p - L} y^{p-1} f(\nu_p - y)\, dy = \int_0^{R-\nu_p} y^{p-1} f(\nu_p + y)\, dy, \qquad p \in \mathcal{D} \setminus \{0\}, \tag{1.3}$$

where X has support on the possibly infinite interval (L, R). The following result establishes the relationship between true positive skewness and stochastic dominance of the left and right parts of a distribution.

Theorem 1.7 (Kovchegov [8], 2022)**.** *Let X be a continuous random variable supported on (L, R) with density function f. For fixed $p \in \mathcal{D}$, if the distribution with density function $y \mapsto \frac{1}{H_p} y^{p-1} f(\nu_p + y)\mathbf{1}_{(0,R-\nu_p)}(y)$ exhibits strict stochastic dominance over the distribution with density function $y \mapsto \frac{1}{H_p} y^{p-1} f(\nu_p - y)\mathbf{1}_{(0,\nu_p-L)}(y)$, then $\nu(\cdot)$ is increasing at p.*

Remark 1.8. In the above Theorem 1.7, the stochastic dominance of the density function $\frac{1}{H_p} y^{p-1} f(\nu_p + y)\mathbf{1}_{(0,R-\nu_p)}(y)$ representing the right part of f over the density function $\frac{1}{H_p} y^{p-1} f(\nu_p - y)\mathbf{1}_{(0,\nu_p-L)}(y)$ representing the left part of f for all $p \in \mathcal{D}$ captures the geometric nature of skewness. Thus, Theorem 1.7 explains why the above stochastic dominance implies positive Pearson's moment coefficient of skewness ($\nu_4 > \nu_2$) and positive Pearson's second skewness coefficient ($\nu_2 > \nu_1$). Importantly, Theorem 1.7 provides geometric justification for the notion of true positive skewness as well as to the original Pearson's coefficients of skewness.

Remark 1.9. The crux of the proof of Theorem 1.7 is that ν_p is increasing if and only if

$$\int_0^{R-\nu_p} y^{p-1} \log y \; f(\nu_p + y)\, dy - \int_0^{\nu_p - L} y^{p-1} \log y \; f(\nu_p - y)\, dy > 0. \tag{1.4}$$

Therefore, weaker versions of stochastic ordering between the left and right parts actually suffice for Theorem 1.7, in particular the concave ordering (see, e.g., [10, 13]).

We will frequently use the following criterion for true positive skewness.

Lemma 1.10 (Kovchegov [8], 2022)**.** *Let X be a continuous random variable supported on the possibly infinite interval (L, R) with density function f. For fixed $p \in \mathcal{D}$, suppose there exists $c > 0$ such that*

(a) $f(\nu_p - c) = f(\nu_p + c)$;
(b) $f(\nu_p - x) > f(\nu_p + x)$ for $x \in (0, c)$; and
(c) $f(\nu_p - x) < f(\nu_p + x)$ for $x > c$.

If $\nu_p - L \le R - \nu_p$, or if $L = -\infty$ and $R = \infty$, then $\nu(\cdot)$ is increasing at p.

The following is a special case.

Proposition 1.11 (Kovchegov [8], 2022)**.** *If f is strictly decreasing on its support, then X is truly positively skewed.*

The requirement of strict monotonicity in the proposition can be relaxed, which will be necessary when we consider uniform mixtures in a later section.

Proposition 1.12. *If f is non-increasing on its support (L, R), and there exist any two points $y_1, y_2 \in (L, R)$, $y_1 < y_2$, such that $f(y_1) > f(y_2)$, then X is truly positively skewed.*

The strict inequalities in Lemma 1.10 can be relaxed in a similar manner.

2. Results

2.1. **Properties of p-means.** In this section, we establish several simple but fundamental properties of p-means and their behavior. Here and throughout this work, let ν_p^X denote the p-mean of a random variable X whenever defined. We use ν_p when the random variable in question is unambiguous.

Proposition 2.1. *Let X be a random variable supported on the possibly infinite interval (L, R). Then $\nu_p \in (L, R)$ for all $p \in \mathcal{D}$.*

The following fact was used implicitly in [8], but we prove it for completeness.

Proposition 2.2. *The map $p \mapsto \nu_p$ is continuously differentiable on $\mathcal{D} \cap (1, \infty)$.*

When investigating specific distribution families, we may assume that the scale and location parameters are 1 and 0 respectively unless otherwise noted. This is justified by the following, which implies that true positive skewness is preserved under positive affine transformations.

Proposition 2.3. *For any $c, s \in \mathbb{R}$ and $p \in \mathcal{D}_X$, $\nu_p^{cX+s} = c\nu_p^X + s$.*

Next, we consider the asymptotic behavior of ν_p as $p \to \infty$. This requires X to have finite moments of all orders, which clearly holds if X has bounded support. We consider only continuous random variables, but analogous results hold in the discrete case.

Proposition 2.4. *Let X be a continuous random variable supported on the finite interval (L, R). Then $\nu_p \to (L + R)/2$ as $p \to \infty$.*

Suppose instead X has infinite support that is bounded below. We have an analogous result if the support is instead bounded from above.

Proposition 2.5. *Let X be a continuous random variable supported on (L, ∞) for finite L. If X has finite moments of all orders and $P(X > x) > 0$ holds for every $x > L$, then $\nu_p \to \infty$ as $p \to \infty$.*

A consequence of Proposition 2.5 is that no distribution with support on the positive half-line is truly negatively skewed. Similarly, no distribution with support on the negative half-line is truly positively skewed.

2.2. **Examples of true skewness.** True positive skewness has already been shown in [8] for several distributions: exponential, gamma, beta (with the mode in the left half-interval), log-normal, and Pareto. Using Lemma 1.10, we establish true positive skewness of Lévy distribution, and identify the parameter regions for which Weibull and skew-normal distributions are truly skewed. Recall that Lévy distribution has undefined Pearson moment coefficient of skewness because it has no finite integer moments. Yet, as we already mentioned in Remark 1.5, finiteness of integer moments is not required for true skewness. Thus, to the authors' knowledge, Lévy distribution's positive skewness is formally established for the first time in this work.

Definition 2.6. The **Lévy distribution** with location parameter $\mu \in \mathbb{R}$ and scale parameter $\lambda > 0$ is a continuous probability distribution, denoted by Lévy(μ, λ), with the density function

$$f(x;\mu,\lambda) := \sqrt{\frac{\lambda}{2\pi}}\,\frac{e^{-\frac{\lambda}{2(x-\mu)}}}{(x-\mu)^{3/2}}, \qquad x > \mu. \tag{2.1}$$

Definition 2.7. The **Weibull distribution** with shape paramter $k > 0$ and scale parameter $\lambda > 0$ is a continuous probability distribution, denoted by Weibull(k, λ), with the density function

$$f(x;k,\lambda) := \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}, \qquad x > 0. \tag{2.2}$$

Definition 2.8. The **skew-normal distribution** with shape parameter $\alpha \in \mathbb{R}$ is a continuous probability distribution, denoted by SkewNormal(α), with the density function

$$f(x;\alpha) = 2\phi(x)\Phi(\alpha x), \qquad x \in \mathbb{R}, \tag{2.3}$$

where $\phi(x) = (2\pi)^{-1/2}e^{-x^2/2}$ and $\Phi(x) = \int_{-\infty}^{x} \phi(t)dt$ are respectively the density and distribution functions of the standard normal Gaussian distribution.

A common strategy for showing the true skewness of the preceding distributions employs the following observation.

Lemma 2.9. *Suppose there exists a constant C such that, for every $p \in \mathcal{D} \setminus \{0\}$, $\nu(\cdot)$ is increasing at p whenever $\nu(p) > C$. If $\nu(p') > C$ for some $p' \in \mathcal{D} \setminus \{0\}$, then $\nu(\cdot)$ is increasing on $\mathcal{D} \cap [p', \infty)$.*

When considering distribution families, we can always assume the location parameter is 0 and the scale parameter is 1, since they do not affect the direction of skewness (see Proposition 2.3).

Theorem 2.10. *The* Lévy(μ, λ) *distribution is truly positively skewed.*

Theorem 2.11. *The* Weibull(k, λ) *distribution is truly positively skewed if and only if* $0 < k < \frac{1}{1-\log 2}$. *Moreover, it is never truly negatively skewed.*

The Lévy and Weibull distributions are supported only on the half-line, but similar techniques apply if the distribution is supported on the entire line.

Theorem 2.12. *The* SkewNormal(α) *distribution is truly positively skewed if $\alpha > 0$, truly negatively skewed if $\alpha < 0$, and symmetric if $\alpha = 0$.*

2.3. **True skewness under limits in distribution.** It is reasonable to conjecture that true skewness is preserved under uniform limits of distribution functions since Lemma 1.10 implies that true skewness is, essentially, a feature of a continuous random variable's density function. However, one must be somewhat careful: let $X_n \sim \text{Gamma}(n, \lambda)$ be a sequence of independent gamma random variables, for some fixed λ, such that each X_n can be expressed as a sum of n i.i.d. exponential random variables. We know from [8] that the X_n's are truly positively skewed, but the central limit theorem implies that their limit in distribution is Gaussian and thus symmetric.

Therefore, we introduce the notion of *true non-negative skewness* to refer to a random variable whose p-means are non-decreasing, i.e., $d\nu_p/dp \geq 0$, as opposed to the strict increase required by true positive skewness. Notice that truly positively skewed as well as symmetric distributions are truly non-negatively skewed.

Theorem 2.13. *Let F_n, F be the distribution functions of X_n, X respectively. Suppose that $F_n \to F$ uniformly and that $\sup_n E\big[|X_n|^{p+\epsilon}\big] < \infty$ holds for some $\epsilon > 0$ and every $p \in \mathcal{D}_X$. If the X_n's are truly non-negatively skewed, then X is truly non-negatively skewed.*

If, moreover, there exists a constant $C > 0$ such that $\frac{d}{dp}\nu_p^{X_n} > C$ holds for every n and every $p \in \mathcal{D}_X$, then X is truly positively skewed.

The condition that the distribution functions F_n converge uniformly is satisfied often in practice, since uniform convergence follows from pointwise convergence if F is continuous (see, e.g., [5, Exercise 3.2.9]), which holds if X has no point masses. Alternatively, uniform convergence of the distribution functions holds if the characteristic functions converge uniformly.

Theorem 2.13 can be used when considering the parameter regions of true skewness for certain distribution families. As an example, consider a sequence $X_n \sim \text{Weibull}(k_n, 1)$ of independent Weibull random variables, where $k_n \uparrow \frac{1}{1-\log 2}$; these are truly positively skewed by Theorem 2.11. It is clear that the distribution functions of the X_n's converge uniformly to the distribution function of $X \sim \text{Weibull}(\frac{1}{1-\log 2}, 1)$, and one can show directly that the p-th moments of the X_n's are uniformly bounded for every $p \geq 1$ (see, e.g., [12, Eq. 2.63d]). Theorem 2.13 then implies that X is truly non-negatively skewed.

2.4. **Additional criteria for true skewness.** In this section, we establish two criteria for true positive skewness, one based upon a stochastic representation and the other based upon numerically verifiable conditions.

Theorem 2.14. *Let X be a continuous random variable with density function decreasing on its support. If $u : \mathbb{R} \to \mathbb{R}$ is convex and strictly increasing on the support of X, then $u(X)$ is truly positively skewed.*

One immediate application of this theorem is when $X \sim \text{Exp}(\lambda)$ is exponentially distributed. It is well-known that $ke^X \sim \text{Pareto}(k, \lambda)$ and that $X^2 \sim \text{Weibull}(\frac{1}{2}, \frac{1}{\lambda^2})$, so Theorem 2.14 immediately yields the true positive skewness of the $\text{Pareto}(k, \lambda)$ and $\text{Weibull}(\frac{1}{2}, \frac{1}{\lambda^2})$ distributions.

Our second criteria for true positive skewness does not rely on the p-means of a distribution other than its mode and median, provided that it has a density function supported on the half-line which is twice continuously differentiable. It also does not require knowledge of the density f expressed in terms of elementary functions, which has conveniently been provided in each of the specific distributions previously examined; instead, it requires certain bounds on the logarithmic derivative within certain intervals. This theorem may have applications in numerically checking true positive skewness for one-sided stable distributions, for which very little descriptive information is known in general, given specific parameter values.

Theorem 2.15. *Let X have support on $(0,\infty)$ with continuous unimodal density $f \in C^2(0,\infty)$. Suppose f has exactly two positive inflection points θ_1, θ_2 such that $\theta_1 < \nu_0 < \theta_2$, and*

(1) $f'/f > 1/\nu_0$ on $(0,\theta_1)$,
(2) $f'/f > -1/\nu_0$ on (θ_2,∞).

If $\nu_1 > (\nu_0+\theta_2)/2$, then X is truly positively skewed.

Corollary 2.16. *Let X have support on $(0,\infty)$ with continuous unimodal density $f \in C^2(0,\infty)$. Suppose f has exactly one positive inflection point $\theta > \nu_0$. If $\nu_1 > (\nu_0+\theta)/2$, then X is truly positively skewed.*

Remark 2.17. The conditions of Theorem 2.15 can be relaxed, at the potential cost of practical ease. In particular, the following require one to compute c_1, as defined in (3.13).

(a) The condition $\nu_1 > (\nu_0+\theta_2)/2$ serves to guarantee that f is convex at $\nu_p + c_p$ for all p, but the required convexity can certainly be achieved for a weaker lower bound on ν_1. Indeed, one can see in the proof in Section 3 that $\nu_1 + c_1 > \theta_2$ is sufficient; we obtain (3.16) via Lemma 3.4. Note that this is not *always* a weaker lower bound.
(b) The quantity ν_0 in conditions (1) and (2) can be replaced by $\nu_1 - c_1$.* As we show later, $\nu_1 - c_1 < \nu_0$, so actually this replacement creates a stronger condition on the lower bound of f'/f to the left of the mode and a weaker condition on the lower bound of f'/f to the right of the mode. This replacement is useful when the density has a steeper right tail.
(c) Similarly, condition (2) only needs to hold on $(\nu_1 + c_1, \infty)$.

Example 2.18 (Log-logistic distribution). The log-logistic distribution with shape parameter $\beta > 0$ has density function

$$f(x;\beta) := \frac{\beta x^{\beta-1}}{(1+x^\beta)^2}, \qquad x > 0.$$

If $0 < \beta \le 1$, then f strictly decreasing and true positive skewness follows from Proposition 1.11. Suppose $\beta > 1$. One can verify that f is unimodal with mode

$$\nu_0 = \left(\frac{\beta-1}{\beta+1}\right)^{1/\beta},$$

* This makes the proof significantly lengthier; in fact, Lemmas 3.2, 3.3, and 3.4 are otherwise unnecessary. For the proof of Theorem 2.15 in Section 3, we present the most general argument.

median $\nu_1 = 1$, and inflection points

$$\theta^{\pm} = \left(\frac{2\beta^2 - 2 \pm \beta\sqrt{3\beta^2 - 3}}{\beta^2 + 3\beta + 2} \right)^{1/\beta}.$$

Straightforward computations show that $\theta^- \leq 0$ holds if and only if $\beta \leq 2$, and $\theta^+ > \nu_0$ holds if and only if $\beta > 1$. Moreover, $\theta^+ < 1$ holds if and only if $1 \leq \beta < 2$. Therefore, the log-logistic distribution is truly positively skewed if $0 < \beta < 2$, by Corollary 2.16.

3. Proofs

3.1. Proofs of results for Section 1.

Proof of Proposition 1.12. Clearly L is finite, otherwise f could not be a non-increasing density function. Notice that $\nu_p < (L+R)/2$ for all $p \in \mathcal{D}$. Otherwise (1.3) fails to hold since $f(\nu_p + y) \leq f(\nu_p - y)$ and $R - \nu_p < \nu_p - L$ by assumption.

Now the existence of $y_1 < y_2$ such that $f(y_1) > f(y_2)$ implies that there exists a non-singleton interval in $(0, \nu_p - L)$ on which $f(\nu_p + y) < f(\nu_p - y)$. Then strict stochastic dominance of $\frac{1}{H_p} y^{p-1} f(\nu_p + y) \mathbf{1}_{(0, R-\nu_p)}(y)$ over $\frac{1}{H_p} y^{p-1} f(\nu_p - y) \mathbf{1}_{(0, \nu_p - L)}(y)$ follows by integrating each density, applying monotonicity of the integral, and using equation (1.3). □

3.2. Proofs of results for Section 2.1.

Proof of Proposition 2.1. If $\nu_p \leq L$, then $E[(\nu_p - X)_+^{p-1}] = 0$ while $E[(X - \nu_p)_+^{p-1}] > 0$, contradicting (1.1). Analogous if $\nu_p \geq R$. □

Proof of Proposition 2.2. Consider the function $\Phi : \mathbb{R} \times \mathcal{D} \cap (1, \infty) \to \mathbb{R}$ given by

$$\Phi(a, p) := E[(X - a)_+^{p-1}] - E[(a - X)_+^{p-1}]$$

Both integrands $(X - a)_+^{p-1}$ and $(a - X)_+^{p-1}$ are continuously differentiable functions of a and p within their support and are dominated by an integrable function since $E|X|^{p-1} < \infty$. By the Leibniz integral rule, observe that

$$\frac{\partial \Phi}{\partial a}(a, p) = -(p-1)\Big(E[(X - a)_+^{p-2}] + E[(a - X)_+^{p-2}] \Big),$$

which is strictly negative and finite for all $a \in \mathbb{R}$ and $p \in \mathcal{D} \cap (1, \infty)$. The map $p \mapsto (\nu_p, p)$ is the zero level curve of Φ and so $p \mapsto \nu_p$ is continuously differentiable by the implicit function theorem. □

Proof of Proposition 2.3. Let $\mathcal{D} := \mathcal{D}_{cX+s}$. It is easy to see that $\mathcal{D}_X = \mathcal{D}$. For every $p \in \mathcal{D}$,

$$\begin{aligned} E\left[\Big((cX + s) - (c\nu_p^X + s) \Big)_+^{p-1} \right] &= c^{p-1} E[(X - \nu_p^X)_+^{p-1}] \\ &= c^{p-1} E[(\nu_p^X - X)_+^{p-1}] = E\left[\Big((c\nu_p^X + s) - (cX + s) \Big)_+^{p-1} \right] \end{aligned}$$

holds, and the result follows. □

Proof of Proposition 2.4. By Proposition 2.3, it suffices to show that if X has support on $(0, 1)$, then $\nu_p \to 1/2$.

Let $\epsilon \in (0, 1/2)$ be given. Suppose for the sake of contradiction that there exists a subsequence ν_{p_k} such that $\nu_{p_k} < 1/2 - \epsilon$ for all $k \in \mathbb{N}$. For $p_k > 1$, we have

$$\begin{aligned} E[(X - \nu_{p_k})^{p_k-1}\mathbf{1}\{X > \nu_{p_k}\}] &> E[(X - 1/2 + \epsilon)^{p_k-1}\mathbf{1}\{X > 1 - \epsilon\}] \\ &> E[2^{-(p_k-1)}\mathbf{1}\{X > 1 - \epsilon\}] \\ &= 2^{-(p_k-1)}P(X > 1 - \epsilon) \end{aligned} \tag{3.1}$$

and

$$\begin{aligned} E[(\nu_{p_k} - X)^{p_k-1}\mathbf{1}\{X < \nu_{p_k}\}] &< E[(1/2 - \epsilon)^{p_k-1}\mathbf{1}\{X < 1/2 - \epsilon\}] \\ &= (1/2 - \epsilon)^{p_k-1}P(X < 1/2 - \epsilon). \end{aligned} \tag{3.2}$$

From (3.1) and (3.2), we have

$$\frac{E[(\nu_{p_k} - X)^{p_k-1}\mathbf{1}\{X < \nu_{p_k}\}]}{E[(X - \nu_{p_k})^{p_k-1}\mathbf{1}\{X > \nu_{p_k}\}]} < \frac{P(X < 1/2 - \epsilon)}{P(X > 1 - \epsilon)} \cdot (1 - 2\epsilon)^{p_k-1} \to 0$$

as $k \to \infty$. This contradicts (1.1). Thus, every subsequence ν_{p_k} has only finitely many points that lie below $1/2 - \epsilon$. We can similarly show that every subsequence has only finitely many points that lie above $1/2 + \epsilon$. It follows by taking $\epsilon \downarrow 0$ that every subsequence converges to $1/2$, which completes the proof. □

Proof of Proposition 2.5. We may assume $L = 0$. Suppose for the sake of contradiction that there exists $M > 0$ such that $\nu_p < M$ for all $p \geq 1$. By Proposition 2.3, we may assume $M = 1/2$. Note that $\nu_p - X < 1/2$ on $\{X < \nu_p\}$, so by the bounded convergence theorem $E[(\nu_p - X)_+^{p-1}] \to 0$ as $p \to \infty$. On the other hand,

$$E[(X - \nu_p)_+^{p-1}] \geq E[(X - \nu_p)^{p-1}\mathbf{1}\{X > \nu_p + 1\}] \geq P(X > 3/2) > 0,$$

hence (1.1) fails to hold for large p, contradiction. □

3.3. **Proofs of results for Section 2.2.**

Proof of Lemma 2.9. Suppose for the sake of contradiction that the set

$$A := \{p \in \mathcal{D} \cap [p', \infty) : \nu_p \leq C\}$$

is non-empty. Since $p \mapsto \nu_p$ is continuous by Proposition 2.2, then A is closed and thus contains its infimum $q > p'$. By construction, $\nu_p > C$ holds for every $p \in [p', q)$, hence $\nu(\cdot)$ is increasing on $[p', q)$ and so $\nu_q \geq \nu_{p'} > C$, contradicting the fact that $q \in A$. □

Proof of Theorem 2.10. By Proposition 2.3, it suffices to consider the Lévy$(0, 1)$ distribution. Let f be its density function, i.e., $f(x) \equiv f(x; 0, 1)$ as in (2.1).

Fix $p \in \mathcal{D} \setminus \{0\}$. To show that $\nu(\cdot)$ is increasing at p, it suffices to show that the log density ratio of the left and right parts

$$R_p(x) := \log\left(\frac{f(\nu_p - x)}{f(\nu_p + x)}\right) = \frac{3}{2}(\log(\nu_p + x) - \log(\nu_p - x)) - \frac{1}{2(\nu_p - x)} + \frac{1}{2(\nu_p + x)},$$

for $x \in [0, \nu_p)$, has exactly one positive critical point. Indeed, observe that $R_p(x) \to -\infty$ as $x \to \nu_p-$ and that $R_p(0) = 0$. Moreover, (1.3) implies that $R_p(\cdot)$ cannot be non-positive for all $x > 0$. Therefore, if $R_p(\cdot)$ has a single positive critical point, then it must be a maximum, and it follows that the conditions of Lemma 1.10 are satisfied and so $\nu(\cdot)$ is increasing at p.

To identify the positive critical points, observe that

$$R_p'(x) = \frac{3}{2(\nu_p - x)} + \frac{3}{2(\nu_p + x)} - \frac{1}{2(\nu_p - x)^2} - \frac{1}{2(\nu_p + x)^2},$$

so the critical points are solutions to the equation

$$3(\nu_p - x)(\nu_p + x)^2 + 3(\nu_p + x)(\nu_p - x)^2 - (\nu_p + x)^2 - (\nu_p - x)^2 = 0.$$

Further simplification yields a quadratic equation in x,

$$(3\nu_p + 2)x^2 + \nu_p^2(3\nu_p - 2) = 0$$

which has exactly one root in $(0, \nu_p)$ if $\nu_p > 2/3$. This holds for arbitrary $p \in \mathcal{D}$, so we may conclude true positive skewness if $\nu_1 > 2/3$ and $\nu_1 > \nu_0$ hold, by Lemma 2.9. The mode and median of the Lévy distribution are well-known (see, e.g., [11]) and can be computed directly as $\nu_0 = 1/3$ and $\nu_1 = \frac{1}{2}(1 - \mathrm{erfc}(\frac{1}{2}))^{-2} \approx 2.17$, which completes the proof. □

Proof of Theorem 2.11. By Proposition 2.3, it suffices to consider the Weibull$(k, 1)$ distribution. Let f be its density function, i.e., $f(x) \equiv f(x; k, 1)$ as in (2.2).

It is well-known (see, e.g., [12]) that the Weibull$(k, 1)$ distribution has finite moments of all orders and is unimodal with median and mode given by

$$\nu_0 = \left(\frac{k-1}{k}\right)^{1/k}, \qquad \nu_1 = (\log 2)^{1/k}. \tag{3.3}$$

Notice that $\nu_0 < \nu_1$ holds if and only if $k < (1 - \log 2)^{-1}$.

Since the Weibull distribution has support on the positive half-line, then the second part of the theorem follows from Proposition 2.5. For the first part of the theorem, the "only if" part follows from (3.3).

It remains to show the "if" part. For $k \leq 1$, f is strictly decreasing, so we are done by Proposition 1.11.

Suppose $k > 1$ and fix $p \geq 1$. As in the proof of Theorem 2.10, to show that $\nu(\cdot)$ is increasing at p, it suffices to show that the log density ratio of the left and right parts

$$R_p(x) := \log\left(\frac{f(\nu_p - x)}{f(\nu_p + x)}\right) = (k-1)\left[\log(\nu_p - x) - \log(\nu_p + x)\right] - (\nu_p - x)^k + (\nu_p + x)^k$$

for $x \in [0, \nu_p)$, has exactly one positive critical point, since $R_p(x) \to -\infty$ as $x \to \nu_p-$ and $R_p(0) = 0$. Observe that

$$R_p'(x) = k(\nu_p - x)^{k-1} + k(\nu_p + x)^{k-1} - \frac{k-1}{\nu_p - x} - \frac{k-1}{\nu_p + x},$$

so the critical points are solutions to the equation

$$k(\nu_p^2 - x^2)\left[(\nu_p - x)^{k-1} + (\nu_p + x)^{k-1}\right] - 2(k-1)\nu_p = 0. \tag{3.4}$$

Since $x \in [0, \nu_p)$, the binomial series for $(\nu_p - x)^{k-1}$ and $(\nu_p + x)^{k-1}$ converge, hence

$$(\nu_p - x)^{k-1} + (\nu_p + x)^{k-1} = 2\sum_{n=0}^{\infty} \binom{k-1}{2n} \nu_p^{k-2n-1} x^{2n},$$

in the notation of the generalized binomial coefficient. Substituting into (3.4) yields the equation

$$g(x) := 2k\sum_{n=1}^{\infty}\left(\binom{k-1}{2n} - \binom{k-1}{2n-2}\right)\nu_p^{k-2n+1}x^{2n} + 2\nu_p(k\nu_p^k - k + 1) = 0, \tag{3.5}$$

where

$$\binom{k-1}{2n} - \binom{k-1}{2n-2} = \frac{k(k-1)\dots(k-2n+2)(k+1-4n)}{(2n)!}. \tag{3.6}$$

We analyze the sign changes of the coefficients in (3.5) to determine the number of its positive roots by splitting into several cases for the value of k.

Case 1: $2 < k < 3$. Then $k, k-1, k-2$ are positive and $k-3, \dots, k-2n+2$ are negative. There are an even number (possibly zero, if $n = 1, 2$) of negative factors in $k(k-1)\dots(k-2n+2)$, so it is positive. Also note that $k+1-4n < 0$ for all $n \geq 1$, hence (3.6) is negative for all $n \geq 1$.

For the series expression $g(x) = \sum_{m=0}^{\infty} a_m x^m$, we have shown that $a_m < 0$ holds for non-zero even m and $a_m = 0$ holds for odd m. By Descartes' rule of signs for infinite series, g_p^* has no positive real roots if $a_0 \leq 0$ and has at most one positive real root if $a_0 > 0$.

Suppose $a_0 \leq 0$ such that g has no positive real roots. By extension, R_p' has no positive real roots, so R_p has no positive extrema and is strictly monotonic on $[0, \nu_p)$. Since $g_p(0) = 0$ and $\lim_{x\to\nu_p-} g_p(x) = -\infty$, then g_p is strictly negative on $[0, \nu_p)$, hence $f(\nu_p - x) < f(\nu_p + x)$ for all $x \in (0, \nu_p)$. By monotonicity of the integral, this contradicts (1.3). Thus $a_0 > 0$ holds, and g has exactly one positive root in $(0, \nu_p)$, which implies that $\nu(\cdot)$ is increasing at p.

This holds for arbitrary $p \geq 1$, so $\nu(\cdot)$ is increasing on $\mathcal{D} \setminus \{0\}$. Moreover, since $a_0 = 2\nu_p(k\nu_p^k - k + 1) > 0$ holds, then $\nu_p > \left(\frac{k-1}{k}\right)^{1/k} = \nu_0$ holds for every $p \geq 1$, and we are done.

Case 2: $1 < k < 2$. The inequality $k + 1 - 4n < 0$ still holds for all $n \geq 1$, but we now have $k, k-1 > 0$ and $k-2, \dots, k-2n+2 < 0$. If $n > 1$, then there are an odd number of negative factors in $k(k-1)\dots(k-2n+2)$, so the numerator of (3.6) is positive. If $n = 1$, then the numerator is simply $k(k-3) < 0$.

Thus our coefficients in $g(x) = \sum_{m=0}^{\infty} a_m x^m$ are positive for even $m > 2$ and zero for odd m, with $a_2 < 0$. By Descartes' rule of signs, g has at most two positive real roots if $a_0 > 0$ and at most one if $a_0 \leq 0$.

Suppose $a_0 \leq 0$. If g has a single positive root, then $\lim_{x\to\nu_p-} g(x) > 0$. But by continuity this limit tends to $-2(k-1)\nu_p < 0$. If instead g_p^* has no positive real roots, then by the same argument above we contradict (1.3). Thus $a_0 > 0$. In this case, if g has zero or two positive roots, then again $\lim_{x\to\nu_p-} g(x) > 0$ and we reach a contradiction. Hence g has exactly one positive real root. We may now conclude in the same fashion as in Case 1.

Case 3: $3 < k < \frac{1}{1-\log 2}$. We again look at the numerator of (3.6). If $n = 1$, the numerator is $k(k-3) > 0$. If $n = 2$, the numerator is $k(k-1)(k-2)(k-7) < 0$. If $n > 2$, then $k(k-1)\dots(k-2n+2)$ has positive factors $k, \dots, k-3$, and an odd number of negative factors $k-4,\ldots,k-2n+2$. Additionally, $k+1-4n<0$ when $n>2$, so the numerator of (3.6) is positive.

Our coefficients in $g(x)=\sum_{m=0}^{\infty} a_m x^m$ are zero for odd m, positive for even $m \geq 6$, negative for $m=4$, and positive for $m=2$. This yields two sign changes and hence at most two positive roots of g if $a_0>0$. Recall from Case 1 that $a_0>0$ holds if and only if $\nu_p>\nu_0$ holds. If the latter holds, then g must have exactly one positive root since g is negative at the right limit of its domain, and so $\nu(\cdot)$ is increasing at p. Now by Lemma 2.9, true positive skewness follows if $\nu_1>\nu_0$, but this holds immediately from (3.3) and our assumption $k<\frac{1}{1-\log 2}$.

Case 4: $k=2$. We can plug $k=2$ directly into (3.4) to obtain the roots

$$x^2=\nu_p^2-\frac{1}{2}=\nu_p^2-\nu_0^2.$$

The argument in Case 1 can be adapted to show that $\nu_p>\nu_0$ holds for every $p\geq 1$, hence exactly one positive root exists.

Case 5: $k=3$. Similarly, we plug $k=3$ into (3.4) and obtain the roots

$$x^4=\nu_p^4-\frac{2}{3}\nu_p=\nu_p(\nu_p^3-\nu_0^3)$$

Again, since $\nu_p>\nu_0$ holds for every $p\geq 1$, then exactly one positive root exists. □

Proof of Theorem 2.12. Recall that ϕ,Φ are the density and distribution functions of the standard Gaussian distribution. For simplicity, set

$$\Psi(x):=\sqrt{2\pi}\,\Phi(x)=\int_{-\infty}^{x} e^{-t^2/2}\,dt.$$

We show true positive skewness for positive shape parameter $\alpha>0$; the proof for true negative skewness for $\alpha<0$ is analogous, and clearly the skew-normal is simply the normal distribution, which is symmetric, when $\alpha=0$. Thus, fix $\alpha>0$ and let f be the density function of SkewNormal(α), i.e., $f(x)\equiv f(x;\alpha)$ as in (2.3).

Fix $p\in\mathcal{D}\setminus\{0\}$. To show that $\nu(\cdot)$ is increasing at p, it suffices to show that the log density ratio of the left and right parts

$$R_p(x):=\log\left(\frac{f(\nu_p-x)}{f(\nu_p+x)}\right)=2\nu_p x+\log\Psi(\alpha(\nu_p-x))-\log\Psi(\alpha(\nu_p+x))$$

has exactly one positive root x_0, satisfying $R_p(x)>0$ for $0<x<x_0$ and $R_p(x)<0$ for $x>x_0$. This condition holds if the following four conditions can be verified:

(I) $R_p(0)=0$;
(II) $R_p''(x)<0$ for all $x>0$.
(III) $R_p'(0)>0$; and
(IV) $R_p'(x)<0$ for some $x>0$.

Condition (I) holds trivially. To prove condition (II), define

$$\psi(x):=\log\Psi(x),$$

hence

$$\begin{aligned}R_p(x) &= 2\nu_p x + \psi(\alpha(\nu_p - x)) - \psi(\alpha(\nu_p + x));\\ R_p'(x) &= 2\nu_p - \alpha\psi'(\alpha(\nu_p - x)) - \alpha\psi'(\alpha(\nu_p + x));\\ R_p''(x) &= \alpha^2\psi''(\alpha(\nu_p - x)) - \alpha^2\psi''(\alpha(\nu_p + x)).\end{aligned} \tag{3.7}$$

Condition (II) holds if and only if $\psi''(\alpha(\nu_p - x)) < \psi''(\alpha(\nu_p + x))$ holds for every $x > 0$. It suffices then to show that ψ'' is monotonically increasing everywhere, i.e., that $\psi'''(x) > 0$ holds for every $x \in \mathbb{R}$.

The third logarithmic derivative of Ψ is given by

$$\begin{aligned}\psi'''(x) &= \frac{\Psi'''(x)}{\Psi(x)} - \frac{3\Psi'(x)\Psi''(x)}{\Psi(x)^2} + 2\left(\frac{\Psi'(x)}{\Psi(x)}\right)^3\\ &= \frac{x^2-1}{\Psi(x)}\,e^{-x^2/2} + \frac{3x}{\Psi(x)^2}\,e^{-x^2} + \frac{2}{\Psi(x)^3}\,e^{-3x^2/2}\\ &= \frac{e^{-3x^2/2}}{\Psi(x)^3}\Big[(x^2-1)e^{x^2}\Psi(x)^2 + 3xe^{x^2/2}\Psi(x) + 2\Big].\end{aligned} \tag{3.8}$$

One may compute directly

$$\begin{aligned}\psi'''(1) &= \frac{e^{-3/2}}{\Psi(1)^3}\Big[3e^{1/2}\Psi(1) + 2\Big] > 0;\\ \psi'''(-1) &= \frac{e^{-3/2}}{\Psi(-1)^3}\Big[-3e^{1/2}\Psi(-1) + 2\Big] \approx 0.12 > 0,\end{aligned}$$

so if ψ''' has no zeroes, then continuity implies that ψ''' is positive everywhere.

Define $u \equiv u(x) := (\psi'(x))^{-1} = e^{x^2/2}\Psi(x)$, noticing that u is strictly positive. By (3.8), $\psi'''(x) = 0$ holds if and only if

$$(x^2-1)u^2 + 3xu + 2 = 0,$$

which, by the quadratic formula, holds exactly when

$$u = \frac{-3x \pm \sqrt{x^2+8}}{2(x^2-1)} = \mp\frac{4}{\sqrt{x^2+8} \pm 3x} =: Q^{\pm}(x).$$

In other words, the zeroes of h''' are solutions to the equations

$$u(x) = Q^{\pm}(x), \tag{3.9}$$

with special care needed for the asymptotic values $x = \pm 1$ for $Q^{\pm}$. We handled these cases in (3.8), so it suffices to consider $x \neq \pm 1$.

Observe that $Q^{\pm}(x) < 0 < u(x)$ holds if $x > 1$, so solutions to (3.9) must fall in $(-\infty, 1)$. If $x < -1$, then both $Q^+(x)$ and $Q^-(x)$ are positive and, one can easily verify the inequalities $0 < Q^-(x) < Q^+(x)$. On the other hand, if $-1 < x < 1$, then $Q^+(x) < 0 < u(x)$ whereas $Q^-(x) > 0$. Therefore, it suffices to show that $u(x) < Q^-(x)$ for every $x < 1$, $x \neq -1$.

The key observation is that $u(x)$ is simply the reflection of the Mills' ratio of the standard Gaussian, i.e.,

$$u(-x) = M(x) := \frac{1-\Phi(x)}{\phi(x)} = e^{x^2/2}\int_x^\infty e^{-t^2/2}\,dt,$$

for which very precise bounds in terms of elementary functions exist. In particular, we use the known bound (see, e.g., [1, Eq. 7.8.4], or [14, 16])

$$M(x) < \frac{4}{\sqrt{x^2+8}+3x}, \qquad \forall\, x > -1,$$

which yields

$$u(x) < \frac{4}{\sqrt{x^2+8}-3x} = Q^-(x), \qquad \forall\, x < 1,\ x \neq -1.$$

We have now shown that (3.9) has no solutions and thus that ψ''' has no zeroes, proving condition (II).

Since $R_p(\cdot)$ cannot be non-positive on the entire half-line, as a consequence of (1.3), then condition (II) implies condition (III). For condition (IV), observe that $\psi'(x) = e^{-x^2/2}\,\Psi(x)^{-1}$ is decreasing and tends to infinity as $x \to -\infty$, so (3.7) implies that $R_p'(x) \to -\infty$ as $x \to \infty$.

It remains to show that $\nu_p > \nu_0$ for every $p \in \mathcal{D}\setminus\{0\}$. Notice that $R_p'(0) = 2\nu_p - 2\alpha\psi'(\alpha\nu_p)$ is increasing as a function of ν_p given the decrease of ψ'. The mode ν_0 is the unique value satisfying $f'(\nu_0) = 0$, which is equivalent to the equality $R_0'(0) = 2\nu_0 - 2\alpha\psi'(\alpha\nu_0) = 0$. Condition (III) and the increase of $\nu_p \mapsto R_p'(0)$ now yield $\nu_0 < \nu_p$ for every $p \in \mathcal{D}\setminus\{0\}$. □

3.4. **Proofs of results for Section 2.3.**

Proof of Theorem 2.13. To prove the first part of the theorem, fix $p \in \mathcal{D}_X$ and define

$$f_n(a) := E|X_n - a|^p, \quad a \in \mathbb{R}; \qquad \nu_n := \arg\min_{a\in\mathbb{R}} f_n(a);$$

$$f(a) := E|X - a|^p, \quad a \in \mathbb{R}; \qquad \nu := \arg\min_{a\in\mathbb{R}} f(a);$$

so that ν_n and ν are the p-means of X_n and X respectively by Definition 1.2.

Let a_n be any sequence of numbers converging to a. Observe that the random variable $X_n - a_n$ has distribution function $\tilde{F}_n(x) = F_n(x + a_n)$. Similarly, $X - a$ has distribution function $\tilde{F}(x) = F(x+a)$, so the uniform convergence $F_n \to F$ implies the pointwise convergence $\tilde{F}_n(x) \to \tilde{F}(x)$. Therefore, $X_n - a_n$ converges to $X - a$ in distribution.

By Jensen's inequality, we have

$$\sup_n E|X_n - a_n|^{p+\epsilon} \le 2^{p+\epsilon}\sup_n (E|X_n|^{p+\epsilon} + |a_n|^{p+\epsilon}) < \infty,$$

where finiteness holds by the convergence of $a_n \to a$ and by assumption on the X_n's. It follows from, e.g., [5, Exercise 3.2.5], that

$$f_n(a_n) = E|X_n - a_n|^p \to E|X - a|^p = f(a). \tag{3.10}$$

Suppose that the ν_n's are contained in a compact interval. Then every subsequence ν_{n_k} has a limit point ν_*. Since $f_{n_k}(\nu_{n_k}) \le f_{n_k}(a)$ for all $a \in \mathbb{R}$, then by taking $k \to \infty$ we obtain $f(\nu_*) \le f(a)$ holds for every $a \in \mathbb{R}$ via (3.10). Then ν_* minimizes f, and since the minimizer of f is unique (see Definition 1.2), we obtain $\nu_* = \nu$. Every subsequence of ν_n converges to ν, hence $\nu_n \to \nu$.

Consider ν_n and ν as functions of p, so ν_n converges pointwise to ν. True non-negative skewness of X_n implies ν_n is non-decreasing, and the pointwise limit of monotone functions is monotone, hence ν is non-decreasing.

It remains to show that the ν_n's are contained in a compact interval. Suppose otherwise, so we have a subsequence $\nu_{n_k} \to \infty$ as $k \to \infty$. (The argument is similar if $\nu_{n_k} \to -\infty$.) We have from (3.10) the pointwise convergence $f_n \to f$, so for fixed $\delta > 0$, we may choose large enough K such that $\nu_{n_k} > \nu$ and $|f_{n_k}(\nu) - f(\nu)| < \delta$ for all $k \geq K$. Since f_{n_k} is strictly convex with minimizer ν_{n_k}, then $f_{n_k}(a) < f(\nu) + \delta$ for all $a \in [\nu, \nu_{n_k}]$.

Clearly $f \to \infty$ as $a \to \infty$, so choose $x > \nu_{n_K} > \nu$ large enough such that $f(x) > f(\nu)+2\delta$. For any $N > 0$, there exists $k > K$ large enough such that $n_k > N$ and $\nu_{n_k} > x$ such that $f_{n_k}(x) < f(\nu) + \delta < f(x) - \delta$ as shown above. This contradicts the pointwise convergence $f_{n_k}(x) \to f(x)$, and so concludes the proof of the first part.

The second part of the theorem now easily follows from the fact that $\nu_n(p) \to \nu_n(p)$. Indeed, for $p > p'$, we have by assumption that $\nu_n(p) - \nu_n(p') > C(p - p')$ holds for every $n \in \mathbb{N}$, so taking limits yields $\nu(p)-\nu(p') \geq C(p-p') > 0$, hence $\nu(\cdot)$ is strictly increasing. □

3.5. **Proofs of results for Section 2.4.**

Proof of Theorem 2.14. Let X be a continuous random variable with density function f_X decreasing on its support $(0, R)$ for possibly infinite R. Let $Y = u(X)$ where u is convex and strictly increasing on the support of X. We define $w = u^{-1}$. Note that $\mathcal{D}_X = \mathcal{D}_Y = [1, \infty)$. We write $\mathcal{D}$ for this domain. Let f_Y be the density of Y. It suffices to show that

$$\log\left(\frac{f_Y(\nu_p - y)}{f_Y(\nu_p + y)}\right) \tag{3.11}$$

changes sign exactly once on the interval $(0, \infty)$, for all $p \in \mathcal{D}$. First note that as y increases, $\nu_p - y$ approaches 0. Once $y > \nu_p$, $f_Y(\nu_p - y) = 0$. Thus, on the interval $y \in (\nu_p, \infty)$,

$$\log\left(\frac{f_Y(\nu_p - y)}{f_Y(\nu_p + y)}\right) = -\infty$$

On the interval $u \in (0, \nu_p]$, we have $\{\nu_p - y\} \in [0, \nu_p)$, and $\{\nu_p + y\} \in [\nu_p, 2\nu_p)$. We expand f_Y:

$$\begin{aligned}\log\left(\frac{f_Y(\nu_p - y)}{f_Y(\nu_p + y)}\right) &= \log\left(\frac{-f_X[w(\nu_p - y)]w'(\nu_p - y)}{-f_X[w(\nu_p + y)]w'(\nu_p + y)}\right) \\ &= \log\left(\frac{f_X[w(\nu_p - y)]}{f_X[w(\nu_p + y)]}\right) + \log\left(\frac{w'(\nu_p - y)}{w'(\nu_p + y)}\right)\end{aligned} \tag{3.12}$$

Because $w = u^{-1}$ is an increasing function and f_X is a decreasing function,

$$f_X[w(\nu_p - y)] > f_X[w(\nu_p + y)]$$

$$\log\left(\frac{f_X[w(\nu_p - y)]}{f_X[w(\nu_p + y)]}\right) > 0$$

The function u is convex and increasing, w is concave, and therefore w' is decreasing. It follows that

$$w'(\nu_p - y) > w'(\nu_p + y)$$
$$\log\left(\frac{w'(\nu_p - y)}{w'(\nu_p + y)}\right) > 0$$

which implies that (3.12) is strictly positive on $(0, \nu_p]$. Since (3.11) is strictly positive on $(0, \nu_p]$ and strictly negative on (ν_p, ∞), it changes sign exactly once on the interval $(0, \infty)$. This satisfies Lemma 1.10 for all $p \in \mathcal{D}$, so we obtain true positive skewness. □

Proof of Theorem 2.15. Let f denote the density function of X. For $p \in \mathcal{D} \setminus \{0\}$, define

$$h_p(c) := f(\nu_p + c) - f(\nu_p - c),\ c \in [0, \nu_p]$$

and

$$\mathcal{S}_p := \{c > 0 : h_p(c) > 0\}.$$

If $\mathcal{S}_p$ is non-empty, then its infimum

$$c_p := \inf \mathcal{S}_p \tag{3.13}$$

exists and is non-negative. Note that if f is continuous, then so is h_p. Then $\mathcal{S}_p$ is the preimage of an open set under h_p, so $\mathcal{S}_p$ is also open and $c_p \notin \mathcal{S}_p$. Since $h_p(0) = 0$, then continuity implies

$$h_p(c_p) = 0. \tag{3.14}$$

Similarly, if f is differentiable, then so is h_p. Because $h_p(c) \leq 0$ for all $c < c_p$, then

$$h_p'(c_p) \geq 0. \tag{3.15}$$

To prove the theorem, we will require several lemmas.

Lemma 3.1. *The density f is convex on $(0, \theta_1) \cup (\theta_2, \infty)$, concave on (θ_1, θ_2), strictly increasing on $(0, \nu_0)$, and strictly decreasing on (ν_0, ∞).*

Proof. Since f is positive on its support, then it must be increasing on $(0, \nu_0)$. If f is concave on $(0, \theta_1)$, then it is convex on (θ_1, θ_2), contradicting the fact that $\nu_0 \in (\theta_1, \theta_2)$. The convexity part of the lemma follows, and since $\theta_2 > \nu_0$ and f is unimodal, then f is decreasing on (ν_0, ∞). Moreover, since $f' > 0$ near 0 and $f' < 0$ near infinity, then f' has an odd number of zeroes. Integrability of f implies $f' \to 0$, hence f' is non-zero everywhere on $(0, \infty)$ except at ν_0, otherwise f would have at least three inflection points. □

Since f is C^2 on $(0, \infty)$, then (3.14) and (3.15) hold. Moreover, h_p is C^2. Let

$$\mathcal{D}^* := \{p \in \mathcal{D} \setminus \{0\} : \nu_p > (\nu_0 + \theta_2)/2\}.^\dagger$$

By assumption $1 \in \mathcal{D}^*$. We show that membership in $\mathcal{D}^*$ is sufficient to make (3.15) a strict inequality.

Lemma 3.2. *$h_p'(c_p) > 0$ for all $p \in \mathcal{D}^*$.*

† If, as in Remark 2.17, we use the condition $\nu_1 + c_1 > \theta_2$ rather than $\nu_1 > (\nu_0 + \theta_2)/2$, then instead let $\mathcal{D}^* := \{p \in \mathcal{D} : \nu_p \geq \nu_1 + c_1\}$. The arguments that follow still apply with trivial modifications.

Proof. First we locate $\nu_p + c_p$ and $\nu_p - c_p$. Note that $f(\nu_p - c)$ is increasing and $f(\nu_p + c)$ is decreasing for $c \in (0, \nu_p - \nu_0)$. The fact that h_p is negative near 0 and $h_p(c_p) = 0$ implies $c_p > \nu_p - \nu_0$. By assumption $\nu_p > (\nu_0 + \theta_2)/2$, so it follows that

$$\nu_p + c_p > \theta_2. \tag{3.16}$$

On the other hand, we easily have $\nu_p - c_p > 0$, otherwise $h_p(c_p) = f(\nu_p + c_p) > 0$. If $\nu_p - c_p > \nu_0$, then $f(\nu_p - c_p) > f(\nu_p + c_p)$ by Lemma 3.1, again contradicting $h_p(c_p) = 0$. Thus

$$0 < \nu_p - c_p < \nu_0. \tag{3.17}$$

Now suppose for the sake of contradiction that $h_p'(c_p) = 0$. Either $\nu_p - c_p \in (0, \theta_1)$ or $\nu_p - c_p \in [\theta_1, \nu_0)$ by (3.17). If the latter holds, then $f''(\nu_p - c_p) < 0$, and (3.16) implies $f''(\nu_p + c_p) > 0$. Note that $h_p''(c_p) = f''(\nu_p + c_p) - f''(\nu_p - c_p)$, hence $h_p''(c_p) > 0$. By continuity of h_p'' and (3.14), h_p is strictly convex and thus positive in a neighborhood of c_p (excluding c_p itself), contradicting the minimality of c_p in $\mathcal{S}_p$.

Suppose instead that $\nu_p - c_p \in (0, \theta_1)$. Then from conditions (1) and (2) and equations (3.16) and (3.17), we have the inequalities $\nu_0 f'(\nu_p - c_p) > f(\nu_p - c_p)$ and $\nu_0 f'(\nu_p + c_p) > -f(\nu_p + c_p)$. It follows that $\nu_0 h_p'(c_p) > -h_p(c_p) = 0$ and we obtain a contradiction. □

Next, we prove some properties of c_p.

Lemma 3.3. *The map $p \mapsto c_p$ is continuously differentiable in $\mathcal{D}^*$.*

Proof. Define $\psi : \mathcal{D}^* \times \mathbb{R}_+$ by $\psi(p, c) := h_p(c)$. By Proposition 2.2, ψ is differentiable and has partial derivatives

$$\frac{\partial}{\partial c}\psi(p, c) = f'(\nu_p + c) + f'(\nu_p - c) = h_p'(c),$$
$$\frac{\partial}{\partial p}\psi(p, c) = (f'(\nu_p + c) - f'(\nu_p - c))\frac{d\nu_p}{dp},$$

both of which are jointly continuous in p and c. Thus ψ is continuously differentiable. By Lemma 3.2, $\frac{\partial}{\partial c}\psi(p, c_p) > 0$ for all $p \in \mathcal{D}^*$ and so the continuous differentiability of $p \mapsto c_p$ follows from (3.14) and the implicit function theorem. □

Lemma 3.4. *For any $p \in \mathcal{D}^*$, $\frac{dc_p}{dp}$ and $\frac{d\nu_p}{dp}$ have the same sign, and $\left|\frac{dc_p}{dp}\right| > \left|\frac{d\nu_p}{dp}\right|$ if they are non-zero.*

Proof. By (3.14), we have $f(\nu_p + c_p) - f(\nu_p - c_p) = 0$. The left side is differentiable in p by Lemma 3.3, so taking derivatives, we obtain

$$f'(\nu_p + c_p)\left(\frac{d\nu_p}{dp} + \frac{dc_p}{dp}\right) - f'(\nu_p - c_p)\left(\frac{d\nu_p}{dp} - \frac{dc_p}{dp}\right) = 0. \tag{3.18}$$

Rearranging yields

$$\frac{dc_p}{dp} = \frac{d\nu_p}{dp}\left(\frac{f'(\nu_p - c_p) - f'(\nu_p + c_p)}{f'(\nu_p - c_p) + f'(\nu_p + c_p)}\right),$$

where the fraction is well-defined with positive denominator by Lemma 3.2. The numerator is positive by Lemma 3.1, (3.16), and (3.17), so $\frac{dc_p}{dp}$ has the same sign as $\frac{d\nu_p}{dp}$. If $\frac{d\nu_p}{dp}, \frac{dc_p}{dp} > 0$, then one can see immediately from (3.18) that $\frac{dc_p}{dp} > \frac{d\nu_p}{dp}$. The reverse also follows. □

Our final lemma concerns a criterion for pointwise increasingness of ν_p.

Lemma 3.5. *Fix* $p \in \mathcal{D} \setminus \{0\}$. *If*

$$(\nu_p - c_p)\frac{f'(\nu_p + c_p)}{f(\nu_p + c_p)} > -1, \tag{3.19}$$

then ν *is increasing at* p.

Proof. Since $f(\nu_p + c_p) = f(\nu_p - c_p)$, we may rearrange to obtain

$$-\frac{f(\nu_p - c_p)}{\nu_p - c_p} < f'(\nu_p + c_p).$$

Define the line

$$\ell(x) := -\frac{x f(\nu_p - c_p)}{\nu_p - c_p} + \frac{2\nu_p f(\nu_p - c_p)}{\nu_p - c_p},\ x \geq \nu_p + c_p,$$

and note that

$$\begin{aligned} \ell(\nu_p + c_p) = f(\nu_p - c_p) = f(\nu_p + c_p), \\ \ell(2\nu_p) = 0 = f(0). \end{aligned} \tag{3.20}$$

Note that (3.19) implies $f'(\nu_p + c_p) = \ell'(\nu_p + c_p)$, so by convexity of f on (θ_2, ∞), we have $f' > \ell'$ on $(\nu_p + c_p, \infty)$. Via integration we obtain

$$f(\nu_p + c) > \ell(\nu_p + c) \tag{3.21}$$

for $c > c_p$. We now split into two cases to show $f(\nu_p + c) > f(\nu_p - c)$ for $c > c_p$.

Case 1. Suppose $\nu_p - c_p \in (0, \theta_1]$. By (3.20) and the convexity of f on this interval, we have $\ell(x) > f(2\nu_p - x)$ for $x \in (\nu_p + c_p, 2\nu_p)$, making use of the fact that convexity is preserved under reflection and translation. A change of variables gives the inequality $\ell(\nu_p + c) > f(\nu_p - c)$ for $c \in (c_p, \nu_p)$. Combining with (3.21) yields $f(\nu_p + c) > f(\nu_p - c)$ for $c \in (c_p, \nu_p)$. If $c \geq \nu_p$, then $f(\nu_p + c) > 0 = f(\nu_p - c)$ and we are done.

Case 2. Suppose instead $\nu_p - c_p \in (\theta_1, \nu_0)$. If it happens that $\ell(\nu_p + c) > f(\nu_p - c)$ for all $c \in (c_p, \nu_p)$, then the argument in Case 1 applies and we are done. Otherwise, let

$$\tilde{\ell}(x) := -(x - \nu_p - c_p) f'(\nu_p - c_p) + f(\nu_p + c_p),\ x \geq \nu_p + c_p,$$

be the line such that

$$\tilde{\ell}(\nu_p + c_p) = f(\nu_p + c_p)$$

and

$$\tilde{\ell}(\nu_p + c_p + f(\nu_p + c_p)/f'(\nu_p - c_p)) = 0.$$

Lemma 3.2 and convexity imply $f'(\nu_p + c) > \tilde{\ell}'(\nu_p + c)$ and thus $f(\nu_p + c) > \tilde{\ell}(\nu_p + c)$ for all $c > c_p$. By concavity of f near $\nu_p - c_p$, one can easily see that $\tilde{\ell}(\nu_p + c) > f(\nu_p - c)$ for $c \in (c_p, \nu_p - \theta_1]$. It remains to show $\tilde{\ell}(\nu_p + c) > f(\nu_p - c)$ for $c \in (\nu_p - \theta_1, \nu_p)$.

We have by assumption that $f(\nu_p - c) > \ell(\nu_p + c)$ for c in a right neighborhood of c_p. We also have by (3.20) that $f(\nu_p - c_p) = \ell(\nu_p + c_p)$. It follows that

$$-f'(\nu_p - c_p) > \ell'(\nu_p + c_p) = -\frac{f(\nu_p - c_p)}{\nu_p - c_p}.$$

Substituting with $f(\nu_p - c_p)$ with $f(\nu_p + c_p)$ and rearranging yields

$$\nu_p + c_p + \frac{f(\nu_p + c_p)}{f'(\nu_p - c_p)} > 2\nu_p.$$

The left side is precisely the root of $\tilde{\ell}$ whereas the right side is a root of f. We showed previously that $\tilde{\ell}(2\nu_p - \theta_1) > f(\theta_1)$. Convexity of f on $(0, \theta_1)$ implies $\tilde{\ell}(\nu_p + c) > f(\nu_p - c)$ for $c \in (\nu_p - \theta_1, \nu_p)$, and we are done.

We have now shown that, in general, $f(\nu_p + c) > f(\nu_p - c)$ for $c > c_p$. We also know by definition of c_p that $f(\nu_p + c) \leq f(\nu_p - c)$ for $c < c_p$. Thus the conditions of Lemma 1.10 are satisfied and so ν is increasing at the point p. □

To prove the main theorem, it suffices to show (3.19) for all $p \in \mathcal{D}^*$ and that $\mathcal{D}^* = \mathcal{D} \setminus \{0\}$. If we have conditions (1) and (2) as they are written, then (3.19) holds for all $p \in \mathcal{D}^*$ immediately by (3.16) and (3.17). Suppose instead we only have the weaker condition as stated in Remark 2.17(b):

$$f'/f > 1/(\nu_1 - c_1) \text{ on } (0, \theta_1),$$
$$f'/f > -1/(\nu_1 - c_1) \text{ on } (\nu_1 + c_1, \infty).$$

Let $u(p) := \nu_p - c_p$ on $\mathcal{D}^*$, so u is differentiable on its domain. Also note that $u > 0$ by (3.17). Suppose for some $p' \in \mathcal{D}^*$ that $u(p') \leq \nu_1 - c_1$. Then

$$(\nu_{p'} - c_{p'})\frac{f'(\nu_{p'} + c_{p'})}{f(\nu_{p'} + c_{p'})} > -1,$$

i.e., (3.19) holds for p' and so ν_p is increasing at p'. Lemma 3.4 implies u is decreasing at p'. It follows that for all $p \geq p'$, u is decreasing and so $u(p) \leq \nu_1 - c_1$. In particular, ν_p is increasing for all $p \geq p'$.

By assumption, $1 \in \mathcal{D}^*$ and trivially $u(1) \leq \nu_1 - c_1$. Since $\mathcal{D}^* \in [1, \infty)$, then ν_p is increasing for all $p \in \mathcal{D}^*$. Recall that $p \in \mathcal{D}^*$ if $\nu_p > (\nu_0 + \theta_2)/2$. Since ν_p is increasing at $p = 1$ and for all $p \in \mathcal{D}^*$, then $\nu_p > (\nu_0 + \theta_2)/2$ for all $p \in \mathcal{D} \setminus \{0\}$, hence $\mathcal{D} \setminus \{0\} = \mathcal{D}^*$.[‡] □

Proof of Corollary 2.16. The proof directly extends to the case where f'' only has a single root $\theta > \nu_0$ by setting $\theta_1 = 0$. In fact, conditions (1) and (2) are not necessary. Indeed, we use them once in the proof of Lemma 3.2 in the case $\nu_p - c_p \in (0, \theta_1)$, but this is no longer relevant if f'' has only one root. The only other time we use the conditions is to prove that (3.19) holds for all $p \in \mathcal{D}^*$. However, if f'' has only one root then, (3.19) holds automatically. Indeed, note that f is concave and increasing near $\nu_p - c_p$ for all $p \in \mathcal{D}^*$, hence

$$\frac{f(\nu_p - c_p)}{\nu_p - c_p} > f'(\nu_p - c_p).$$

Since $f'(\nu_p + c_p) < 0$, then by substituting $f(\nu_p + c_p) = f(\nu_p - c_p)$, we have

$$(\nu_p - c_p)\frac{f'(\nu_p + c_p)}{f(\nu_p + c_p)} > \frac{f'(\nu_p + c_p)}{f'(\nu_p - c_p)}.$$

[‡] Clearly the argument for $\mathcal{D}^* = \mathcal{D} \setminus \{0\}$ still applies if we use the alternative definition for $\mathcal{D}^*$.

The right side dominates -1 as a consequence of Lemma 3.2, so we arrive at (3.19). This proves the corollary. □

4. Discussion

This work attempts to broaden our understanding of true skewness by demonstrating the true skewness of several additional distributions and by establishing simpler criteria for which one can conclude a distribution is truly skewed. Theorems 2.10–2.12, which establish the parameter regions for which the Lévy, Weibull, and skew-normal distributions are truly positively skewed, all rely crucially on showing that the conditions of Lemma 1.10 hold for every relevant value of p. This is done by analyzing what we have called the "log density ratio of the left and right parts," which for given p is

$$R_p(x) = \log\left(\frac{f(\nu_p - x)}{f(\nu_p + x)}\right), \qquad x \in [0, \nu_p),$$

where f is the density function of the distribution in question. The conditions of Lemma 1.10 hold if one can show that $R_p(\cdot)$ has exactly one positive root x_0, satisfying $R_p(x) > 0$ for $0 < x < x_0$ and $R_p(x) < 0$ for $x > x_0$. In Theorems 2.10 and 2.11, this can be shown in quite a straightforward manner simply by finding the critical points of $R_p(\cdot)$. Such an approach may likely be applied to other visibly skewed distributions for which a closed-form expression for its density function exists.

However, for the skew-normal distribution, Theorem 2.12, the absence of a closed-form expression for its density function makes the computation of the critical points of $R_p(\cdot)$ intractible. Instead, we exploit the fact that the log derivative of a distribution function Ψ is simply ψ/Ψ, where ψ is the corresponding density function. The ratio ψ/Ψ is the reciprocal of what is commonly referred to as the *hazard rate* of the distribution Ψ, which has been studied with some detail for most well-known distributions. In the proof of Theorem 2.12, we encounter the reciprocal of the hazard rate of the standard Gaussian, which goes by the special name of the *Mills' ratio*. The Mills' ratio, as one may expect, has been very well-studied, allowing us to employ a closed-form expression, which very tightly bounds the quantities we care about, to prove the required properties of the skew-normal distribution's log density ratio of the left and right parts.

Indeed, the proof of Theorem 2.12 identifies a key yet unsurprising connection between the notion of true skewness, which is fundamentally a comparison of the rate of decay of the two sides of a distribution, and the hazard rate. It suggests that the notion of true skewness is an accurate reflection, and in a much more rigorous fashion than Pearson's coefficients of skewness, of the essence of what we imagine *skewness* to mean. More practically, the method by which we prove Theorem 2.12 should work for a larger variety of distributions; a reasonable first direction would be skewed versions of other symmetric distributions, as introduced in, for example, Chapter 1 of [3].

The remainder of this section discusses sums and products of truly skewed random variables; examines whether true skewness extends to the discrete case; and provides an interpretation of true skewness for multivariate distributions.

4.1. **Sums and products of truly skewed random variables.** The goal of this paper and this discussion is to present tools for establishing a larger class of distributions which are truly skewed. A natural question is whether or not true skewness is preserved under certain "transformations" of truly skewed random variables.

Theorem 2.14 gives an affirmative answer to this question, but with limitations: transforming a random variable with a strictly decaying distribution by an increasing convex function preserves its true positive skewness. One might also expect that sums of truly positively skewed random variables preserve true positive skewness, but the rather strong requirement that the p-means of a truly positively skewed random variable be *strictly* increasing allows one to construct rather straightforward counterexamples. Indeed, we may even take the two summands to be identically distributed. For $1/2 < \lambda < 1$, let X and Y be i.i.d. with density

$$f(x) := \begin{cases} \lambda & 0 < x \leq 1 \\ 1-\lambda & 1 \leq x < 2 \\ 0 & \text{otherwise.} \end{cases} \tag{4.1}$$

Since f is decreasing on its support, X and Y are truly positively skewed by Proposition 1.12. One can easily compute the density function f_Z of $Z = X + Y$, the convolution of f with itself and see that, depending on the value of λ, the p-means of Z are not always increasing. For instance, for $\lambda = 3/5$, computing numerically the median ν_1^Z and the integrals in (1.4) yields $\nu_1^Z \approx 1.786$ and

$$\int_0^{4-\nu_1} \log y \ f_Z(\nu_1 + y)dy - \int_0^{\nu_1} \log y \ f_Z(\nu_1 - y)dy \approx -0.000699 < 0,$$

which implies that $p \mapsto \nu_p^Z$ is decreasing at $p = 1$, and so Z is not truly positively skewed.[§]

On the other hand, it may be more fruitful to start with random variables with monotone density functions. One class of truly positively skewed densities that is closed under summation is the class of "decreasing linear densities," of the form

$$f(x) = h - \frac{h^2 x}{2}, \qquad x \in [0, 2/h],$$

for arbitrary $h > 0$.

4.2. **True skewness of discrete distributions.** In the discrete case, even the simplest sums of truly skewed random variables fail to remain truly skewed. Take X and Y to be independent Bernoulli(1/3) random variables; one can show that these are truly positively skewed quite easily. Their sum is Binomial$(2, 1/3)$, which has median 1 and mean $2/3$, hence $\nu_2 < \nu_1$, and true positive skewness fails to hold. Indeed, the notion of true skewness is finnicky for discrete distributions, at least in the way it has been formulated here and in [8]: this could be due to the interpretation of discrete distributions as limits of successfully sharper (continuous) bump functions, in which case distributions like the Binomial$(n, 1/3)$ are no longer unimodal but in fact have n modes.

The proof techniques used in this paper often involve taking the log of a ratio of the density function of a distribution. In particular, the logarithm of a product of two entities

[§] Numerical computations were carried out using Wolfram Mathematica, version 12.3.0.0.

is equal to the sum of the individual logarithms of those entities. For this reason, it seems natural to examine how true skewness behaves under the products of random variables and whether it is preserved. In particular, it seems that taking the logarithm of the density of product of two random variables would yield a sum which could then imply true skewness depending on how one defines the two random variables.

4.3. **True skewness of continuous multivariate distributions.** In multivariate setting, the definition of p**-mean** (1.1) extends naturally to multivariate distributions as follows. For a continuously distributed random vector $\mathbf{X} = (X_1, \ldots, X_k)$, let $\boldsymbol{\nu}_p = \left(\nu_p^{(1)}, \ldots, \nu_p^{(k)}\right)$, where $\nu_p^{(j)}$ is solves

$$E\big[(X_j - \nu_p^{(j)})_+ \,\|\mathbf{X} - \boldsymbol{\nu}_p\|^{p-2}\big] = E\big[(\nu_p^{(j)} - X_j)_+ \,\|\mathbf{X} - \boldsymbol{\nu}_p\|^{p-2}\big] \qquad (j = 1, \ldots, k), \qquad (4.2)$$

where $\|\cdot\|$ is the usual Euclidean norm. Notice that, analogously to (1.2),

$$\boldsymbol{\nu}_p = \underset{\mathbf{a}\in\mathbb{R}^k}{\arg\min}\, E\big[\|\mathbf{X} - \mathbf{a}\|^p\big],$$

whenever p-th moment of $\mathbf{X}$ is finite, i.e., $\boldsymbol{\nu}_p$ is the **Fréchet** p**-mean**. It is compelling to consider the trajectory of $\boldsymbol{\nu}_p$ as a characterization of multidimensional skewness provided (4.2) has a unique solution. Additionally, we are working on applications of the above defined multidimensional p-mean in nonlinear regression analysis.

In the univariate setting, true skewness corresponds to the sign of $d\nu_p/dp$ representing the direction of trajectory $\boldsymbol{\nu}_p$. Following [8], we adjust true skewness accordingly. We let

$$\boldsymbol{\tau}_p := \frac{d\boldsymbol{\nu}_p}{dp} \Big/ \left\| \frac{d\boldsymbol{\nu}_p}{dp} \right\|$$

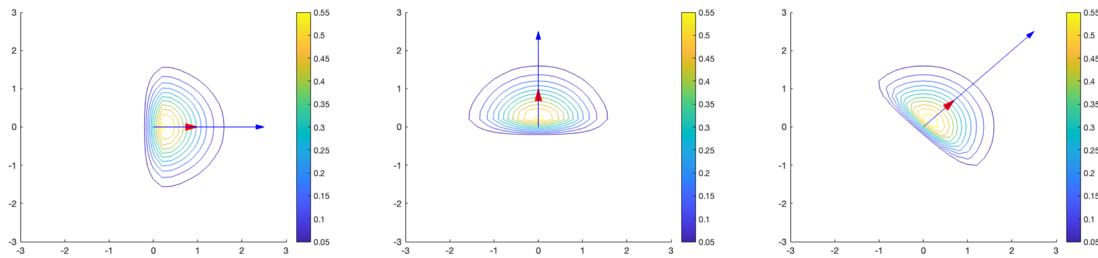
denote the unit tangent vector for trajectory of $\boldsymbol{\nu}_p$ in $\mathbb{R}^k$. It was hypothesized in [8] that the trajectory of $\boldsymbol{\nu}_p$ as a function of p can be interpreted as a backbone of the multivariate distribution and that the limiting direction vector $\lim\limits_{p\to\sup\mathcal{D}} \boldsymbol{\tau}_p$ with p increasing to the rightmost bound in its domain $\mathcal{D}$ may be interpreted as the direction of true skewness.

We will illustrate true skewness in a multivariate setting by means of an example. Considered a *multivariate skew normal distribution* defined in [9], the probability density function of a multivariate skew-normal random vector $\mathbf{Y} \in \mathbb{R}^k$ is

$$f(\mathbf{y}) = 2\phi_k(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma})\Phi_1(\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1/2}(\mathbf{y} - \boldsymbol{\mu})) \qquad \mathbf{y} \in \mathbb{R}^k$$

where $\phi_k(\cdot\,; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the density function of the k-variate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, and $\Phi_1(\cdot)$ is the cumulative distribution function of the univariate standard normal distribution. Taking $\boldsymbol{\lambda} = 0$ recovers the standard multivariate normal distribution with density $\phi_k(\cdot; \boldsymbol{\mu}, \boldsymbol{\Sigma})$. We refer to $\boldsymbol{\lambda}$ as the *skewness parameter* vector.

The arrows plotted in Figure 4.1 motivate the interpretation of multivariate true skewness via direction vectors $\boldsymbol{\tau}_p$, which are naturally co-linear with the skewness parameter vector $\boldsymbol{\lambda}$.



(A) $\boldsymbol{\lambda} = (5, 0)^\top$ (B) $\boldsymbol{\lambda} = (0, 5)^\top$ (C) $\boldsymbol{\lambda} = (5, 5)^\top$

FIGURE 4.1. Contour plots of 2-variate skew-normal distributions. Red arrow is $\boldsymbol{\tau}_p$, which is the same for all p, and blue arrow is $\boldsymbol{\lambda}$.

ACKNOWLEDGMENTS

The majority of this research was done as a part of a Research Experience for Undergraduates (REU) program at the Oregon State University Department of Mathematics funded by NSF grant DMS-1757995. We would like to thank Javier Rojo of Indiana University for his helpful feedback on the concepts considered in this paper. We would also like to acknowledge professor Holly Swisher of Oregon State University who as the director of the REU program provided invaluable perspective throughout the summer.

REFERENCES

[1] *NIST Digital Library of Mathematical Functions*. http://dlmf.nist.gov/, Release 1.1.4 of 2022-01-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.
[2] ABADIR, K. M. The mean-median-mode inequality: counterexamples. *Econometric Theory 21*, 2 (2005), 477–482.
[3] AZZALINI, A. *The skew-normal and related families*, vol. 3. Cambridge University Press, 2013.
[4] DAS, S., MANDAL, P. K., AND GHOSH, D. On homogeneous skewness of unimodal distributions. *Sankhyā: The Indian Journal of Statistics, Series B (2008–)* (2009), 187–205.
[5] DURRETT, R. *Probability: Theory and Examples*, 5th ed. Cambridge University Press, 2019.
[6] GROENEVELD, R. A. Skewness for the Weibull family. *Statistica Neerlandica 40*, 3 (1986), 135–140.
[7] KOPA, M., AND PETROVÁ, B. Strong and weak multivariate first-order stochastic dominance. SSRN 3144058 (2018).
[8] KOVCHEGOV, Y. A new life of Pearson's skewness. *Journal of Theoretical Probability 35*, 4 (2022), 2896–2915. doi:10.1007/s10959-021-01149-7.
[9] LACHOS, V. H., AND LABRA, F. V. Multivariate skew-normal/independent distributions: properties and inference. *Pro Mathematica 28*, 56 (2014), 11–53.
[10] MÜLLER, A., AND STOYAN, D. *Comparison Methods for Stochastic Models and Risks*. Wiley, 2002.
[11] NOLAN, J. P. *Univariate stable distributions*. Springer, 2020.
[12] RINNE, H. *The Weibull distribution: a handbook*. CRC Press, 2008.
[13] ROJO, J. Heavy-tailed densities. *Wiley Interdisciplinary Reviews: Computational Statistics 5*, 1 (2013), 30–40.
[14] SAMPFORD, M. R. Some inequalities on Mill's ratio and related functions. *The Annals of Mathematical Statistics 24*, 1 (1953), 130–132.
[15] SATO, K.-I. Convolution of unimodal distributions can produce any number of modes. *The Annals of Probability* (1993), 1543–1549.

[16] Shenton, L. Inequalities for the normal integral including a new continued fraction. *Biometrika 41*, 1/2 (1954), 177–189.

[17] Yu, Y. Stochastic ordering of exponential family distributions and their mixtures. *Journal of Applied Probability 46*, 1 (2009), 244–254.

Oregon State University
Email address: kovchegy@math.oregonstate.edu

Illinois Institute of Technology
Email address: alexnegron18@gmail.com

Cornell University
Email address: cep87@cornell.edu

Columbia University
Email address: cyw2124@columbia.edu