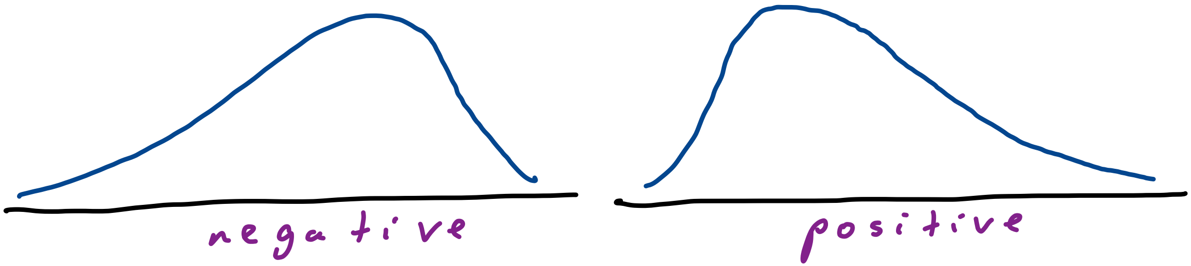
A new life of Pearson's skewness

Yevgeniy Kovchegov
Oregon State University

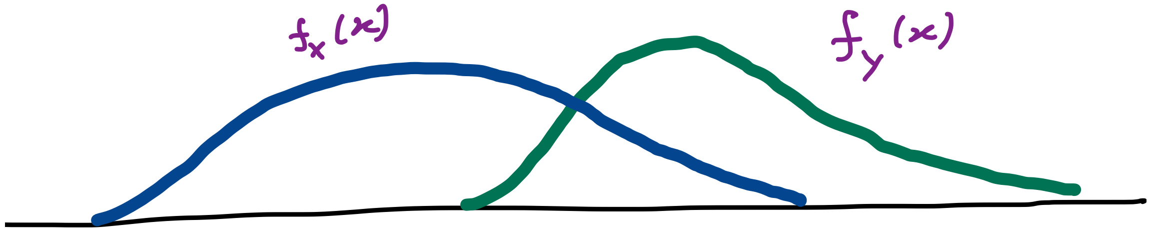
Skewness and Stochastic Dominance.

In this talk, we will connect the following two concepts.

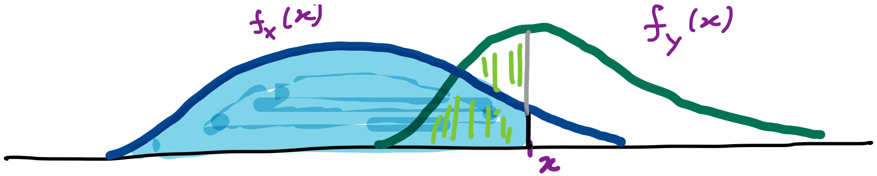
- **Skewness**



- **Stochastic Dominance**



Stochastic Dominance.



Consider random variables X and Y. If the cumulative distribution functions F_X and F_Y satisfy

$$F_X(x) \geq F_Y(x) \qquad \forall x \in \mathbb{R},$$

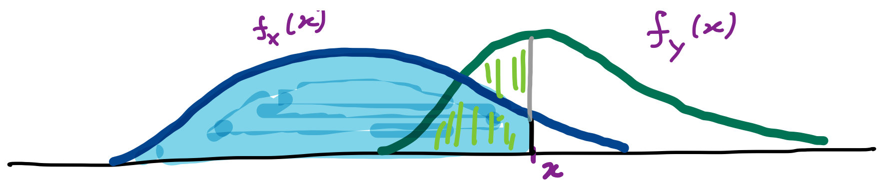
then Y exhibits **stochastic dominance** over X.

Name explained: Suppose for simplicity that both, F_X and F_Y are strictly increasing. Then, we can construct a coupling by letting $U \sim \text{Unif}(0,1)$,

$$X = F_X^{-1}(U) \quad \text{and} \quad Y = F_Y^{-1}(U).$$

Indeed, Y dominates X: $X \leq Y$.

Stochastic Dominance.



Consider random variables X and Y. If the cumulative distribution functions F_X and F_Y satisfy

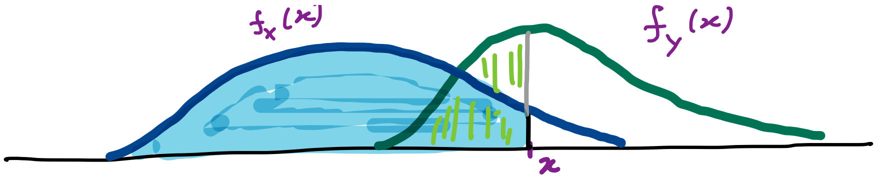
$$F_X(x) \geq F_Y(x) \qquad \forall x \in \mathbb{R},$$

then Y exhibits **stochastic dominance** over X.

Furthermore, if $F_X \not\equiv F_Y$, then, Y exhibits **strict stochastic dominance** over X.

Lemma. If Y exhibits stochastic dominance over X, then, for any increasing function $h : \mathbb{R} \to \mathbb{R}$ we have $E[h(Y)] \geq E[h(X)]$. Moreover, if Y exhibits strict stochastic dominance over X, and if $h(x)$ is strictly increasing, then $E[h(Y)] > E[h(X)]$.

Stochastic Dominance.



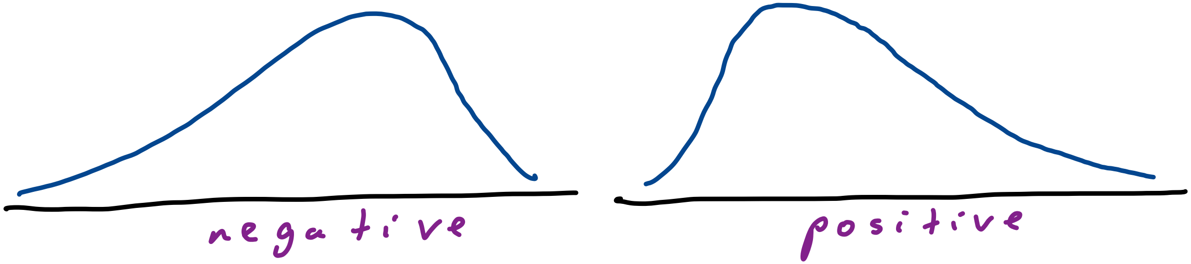
Lemma. If Y exhibits stochastic dominance over X, then, for any increasing function $h : \mathbb{R} \to \mathbb{R}$ we have $E[h(Y)] \geq E[h(X)]$. Moreover, if Y exhibits strict stochastic dominance over X, and if $h(x)$ is strictly increasing, then $E[h(Y)] > E[h(X)]$.

The above lemma is usually proved via coupling argument. For continuous random variables, sometimes can use integration by parts:

$$\int_a^b h(x) f_Y(x) dx = h(b) - \int_a^b h'(x) F_Y(x) dx \geq h(b) - \int_a^b h'(x) F_X(x) dx = \int_a^b h(x) f_X(x) dx$$

Skewness.

Let $\mu = E[X]$ and $\sigma = \sqrt{Var(X)} = \sqrt{E[(X-\mu)^2]}$.



Pearson's moment coefficient of skewness

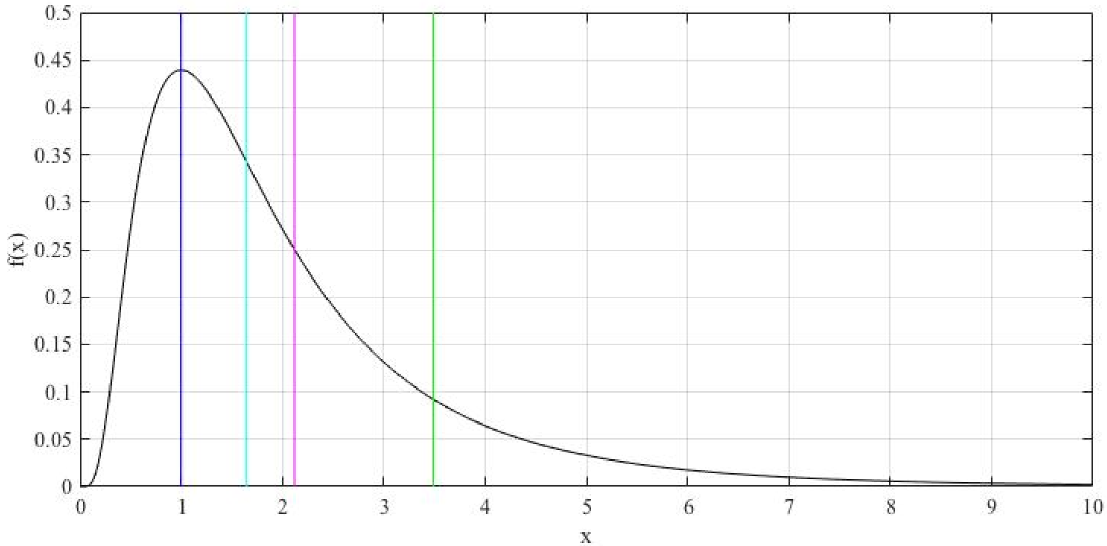
$$\gamma = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$$

The sign of γ determines positive/negative skewness.

Positive skewness $\Rightarrow$ **mean-median-mode inequality:**

mode $<$ median $<$ mean.

Skewness.



Positive skewness $\Rightarrow$ **mean-median-mode inequality:**

mode $<$ median $<$ mean.

Negative skewness $\Rightarrow$ mean $<$ median $<$ mode.

Skewness.

Pearson's moment coefficient of skewness

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

Positive skewness $\Rightarrow$ **mean-median-mode inequality:**

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

Pearson's first skewness coefficient (mode skewness):

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

Pearson's second skewness coefficient (median skewness):

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

Fréchet p-mean.

For $p \in [1, \infty)$ and a random variable X with the finite $(p-1)$-st moment, $E[|X|^{p-1}] < \infty$, the p-**mean** ν_p is the unique solution of

$$E\left[(X-\nu_p)^{p-1}\,\mathbf{1}_{X>\nu_p}\right] = E\left[(\nu_p - X)^{p-1}\,\mathbf{1}_{X<\nu_p}\right].$$

The above defined p-**mean** is extending the notion of the **Fréchet** p-**mean**

$$\nu_p = \mathrm{argmin}_{a\in\mathbb{R}} E\left[|X-a|^p\right]$$

which required finiteness of p-th moment, $E[|X|^p] < \infty$.

Notice that ν_1 is the **median**:

$$P(X > \nu_1) = P(X < \nu_1).$$

Fréchet p-mean.

For $p \in [1, \infty)$ and a random variable X with the finite $(p-1)$-st moment, $E[|X|^{p-1}] < \infty$, the p-**mean** ν_p is the unique solution of

$$E\big[(X-\nu_p)^{p-1}\,\mathbf{1}_{X>\nu_p}\big] = E\big[(\nu_p - X)^{p-1}\,\mathbf{1}_{X<\nu_p}\big].$$

Notice that ν_1 is the **median**:

$$P(X>\nu_1) = P(X<\nu_1).$$

Similarly, ν_2 is the **mean**:

$$E\big[(X-\nu_2)\,\mathbf{1}_{X>\nu_2}\big] = E\big[(\nu_2-X)\,\mathbf{1}_{X<\nu_2}\big] \quad \Leftrightarrow \quad E[X] = \nu_2.$$

Fréchet p-mean.

$$E\left[(X-\nu_p)^{p-1}\,\mathbf{1}_{X>\nu_p}\right] = E\left[(\nu_p-X)^{p-1}\,\mathbf{1}_{X<\nu_p}\right]$$

Notice that ν_1 is the **median**:

$$P(X>\nu_1) = P(X<\nu_1).$$

Similarly, ν_2 is the **mean**:

$$E\left[(X-\nu_2)\,\mathbf{1}_{X>\nu_2}\right] = E\left[(\nu_2-X)\,\mathbf{1}_{X<\nu_2}\right] \quad \Leftrightarrow \quad E[X] = \nu_2.$$

Proposition. Consider a random variable X with $E[|X|^3] < \infty$. The Pearson's moment coefficient of skewness $\gamma > 0$ if and only if $\nu_4 > \nu_2$.

Proof.

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

Skewness.

Pearson's moment coefficient of skewness

$$\gamma = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] > 0 \quad \Leftrightarrow \quad \nu_2 < \nu_4$$

In the **unimodal case**, let ν_0 denote the **mode**.

Pearson's first skewness coefficient (mode skewness)

$$\frac{\nu_2 - \nu_0}{\sigma} > 0 \quad \Leftrightarrow \quad \nu_0 < \nu_2$$

Pearson's second skewness coefficient (median skewness)

$$\frac{3(\nu_2 - \nu_1)}{\sigma} > 0 \quad \Leftrightarrow \quad \nu_1 < \nu_2$$

True positive skewness.

Let $\mathcal{D} = \{p \in [1, \infty) : E[|X|^{p-1}] < \infty\}$ and $\mathcal{D}_0 = \{0\} \cup \mathcal{D}$.

Definition. A continuous random variable X (equivalently, its distribution) is truly positively skewed if and only if ν_p is an **increasing function** of $p \in \mathcal{D}$. Analogously, for true negative skewness.

It is truly mode positively skewed if and only if ν_p, is an **increasing function** of $p \in \mathcal{D}_0$. Analogously, for true mode negative skewness.

True positive skewness insures

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

while truly mode positively skewed insures

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

i.e., all Pearson's skewness criteria and mean-median-mode inequality.

Exponential distribution.

Exponential random variable X: $f(x) = \lambda e^{-\lambda x} \mathbf{1}_{(0,\infty)}(x)$.
W.l.o.g. let $\lambda = 1$. Equation

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

simplifies to

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

Differentiating $\frac{d}{dp}$ yields

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

Exponential distribution.

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

Since $\frac{1}{\Gamma(p)} x^{p-1}\, e^{-x} \mathbf{1}_{(0,\infty)}(x)$ stochastically dominates $\frac{1}{\Gamma(p)} x^{p-1}\, e^{x} \mathbf{1}_{(0,\nu_p)}(x)$,

$$\nu_p^{p-1} e^{\nu_p} \frac{d\nu_p}{dp} = \int_0^{\infty} x^{p-1}\, e^{-x} \log x\,dx - \int_0^{\nu_p} x^{p-1} e^{x} \log x\,dx \; > 0$$

by

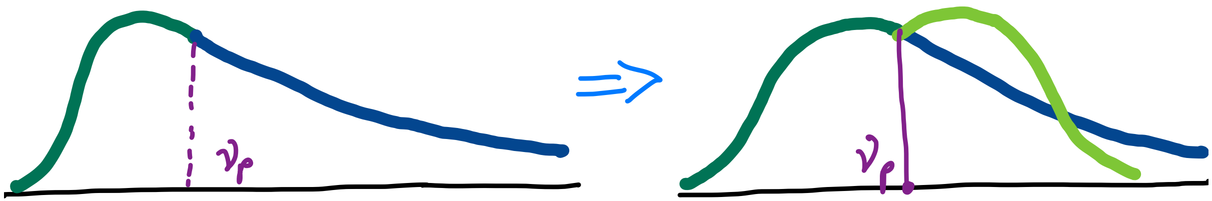
Lemma. If Y exhibits stochastic dominance over X, then, for any increasing function $h : \mathbb{R} \to \mathbb{R}$ we have $E[h(Y)] \geq E[h(X)]$. Moreover, if Y exhibits strict stochastic dominance over X, and if $h(x)$ is strictly increasing, then $E[h(Y)] > E[h(X)]$.

Skewness and stochastic dominance.

Let X be a continuous random variable with density $f(x)$ and $supp(f) = (L, R)$. Then

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

Positive skewness: the left half is "spreading short" and the right half is "spreading longer".



Interpretation: $\frac{1}{H_p} x^{p-1} f(\nu_p + x) \mathbf{1}_{(0,R-\nu_p)}(x)$ to exhibit strict stochastic dominance over $\frac{1}{H_p} x^{p-1} f(\nu_p - x) \mathbf{1}_{(0,\nu_p - L)}(x)$.

Skewness and stochastic dominance.

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

Positive skewness: the left half is "spreading short" and the right half is "spreading longer".

Interpretation: $\frac{1}{H_p} x^{p-1} f(\nu_p + x)\mathbf{1}_{(0,R-\nu_p)}(x)$ to exhibit strict stochastic dominance over $\frac{1}{H_p} x^{p-1} f(\nu_p - x)\mathbf{1}_{(0,\nu_p - L)}(x)$.

Theorem. If $\frac{1}{H_p} x^{p-1} f(\nu_p + x)\mathbf{1}_{(0,R-\nu_p)}(x)$ exhibits strict stochastic dominance over $\frac{1}{H_p} x^{p-1} f(\nu_p - x)\mathbf{1}_{(0,\nu_p - L)}(x)$, 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

True positive skewness: examples.

• **Gamma random variable**: $f(x) = \frac{1}{\Gamma(\alpha)}\lambda^{\alpha}x^{\alpha-1}e^{-\lambda x}$ with parameters $\alpha > 0$ and $\lambda > 0$ is truly mode positively skewed.

• **Beta random variable**: $f(x) = \frac{1}{\mathcal{B}(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}$ with parameters $\beta > \alpha > 1$ (and mode $\nu_0 = \frac{\alpha-1}{\alpha+\beta-2} < \frac{1}{2}$) is truly mode positively skewed.

• **Pareto random variable**: $f(x) = \dfrac{\alpha}{x^{\alpha+1}}$, $x \in [1, \infty)$, with parameter $\alpha > 0$ is truly mode positively skewed.

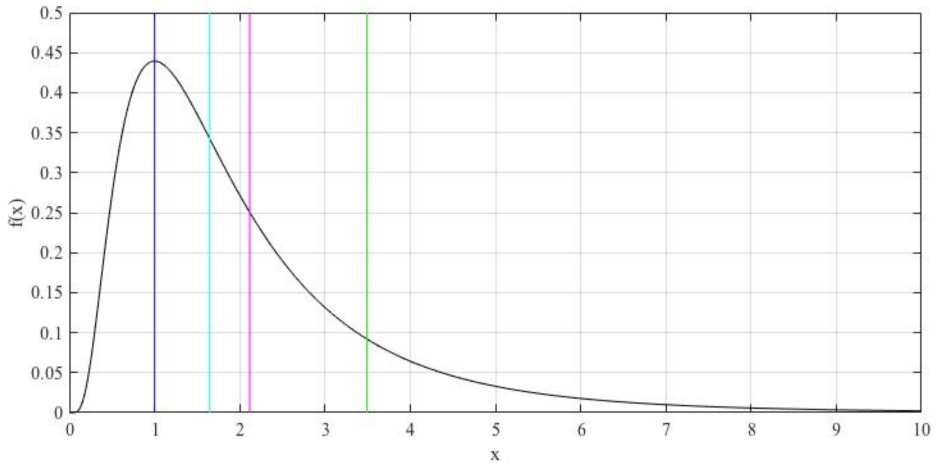
Notice that for $\alpha \in (0, 1)$, the quantities $\nu_2 = E[X]$, σ, and γ do not exist.

- **Log-normal random variable:**

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

over $(L, R) = (0, \infty)$, with parameters μ and σ^2.

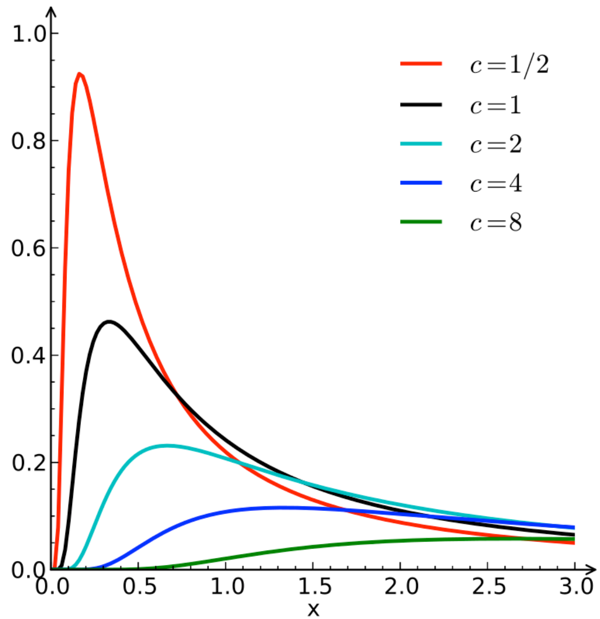
$$\nu_0 = \exp\left\{\mu - \sigma^2\right\}$$



Theorem.

$$\nu_p = \exp\left\{\mu + \frac{p-1}{2}\sigma^2\right\} \quad \text{for all} \quad p \in (0, \infty)$$

Lévy distribution.



For scale parameter $c > 0$,

$$f(x) = \sqrt{\frac{c}{2\pi}} \frac{1}{x^{3/2}} e^{-c/(2x)}, \ x > 0$$

Mean: ∞

Skewness: undefined

My **students**,
Alex Negrón, Clarice Pertel, and Christopher Wang

have shown that

- **Lévy distribution** is truly positively skewed:

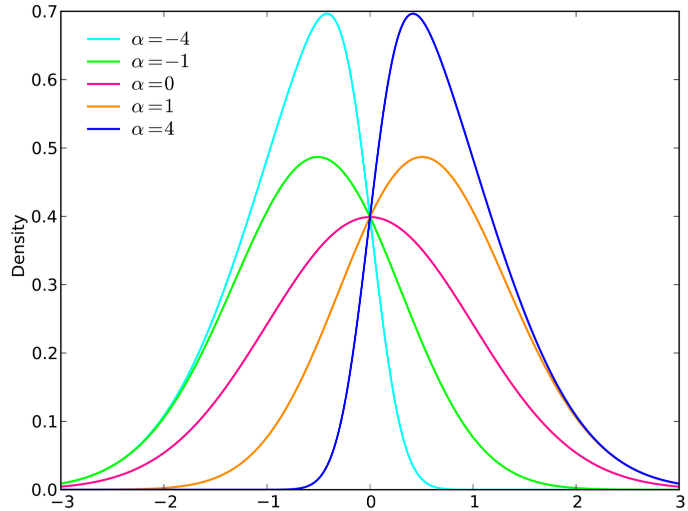
$$\nu_p \uparrow \quad \text{for } p \in \mathcal{D} = [1, 3/2).$$

Skew-normal distribution.

For the parameter $\alpha \in \mathbb{R}$,

$f(x) = 2\varphi(x)\Phi(\alpha x)$, $x \in \mathbb{R}$

where $\varphi(x)$ and $\Phi(x)$ are the standard normal p.d.f. and c.d.f.



My **students**,
Alex Negrón, Clarice Pertel, and Christopher Wang

have shown that

• **Skew-normal distribution** is truly positively skewed if and only if the shape parameter

$$\alpha > 0.$$

Multivariate setting.

The definition of p-**mean** extends naturally to multivariate distributions as follows.

For a continuously distributed random vector $\mathbf{X} = (X_1, \ldots, X_k)$, let $\boldsymbol{\nu}_p = \left(\nu_p^{(1)}, \ldots, \nu_p^{(k)}\right)$ with $\nu_p^{(j)}$ solving

$$E\left[(X_j - \nu_p^{(j)})_+ \, \|\mathbf{X} - \boldsymbol{\nu}_p\|^{p-2}\right] = E\left[(\nu_p^{(j)} - X_j)_+ \, \|\mathbf{X} - \boldsymbol{\nu}_p\|^{p-2}\right],$$

where $x_+ = \max\{0, x\}$ and $\|\cdot\|$ denotes the usual Euclidean norm.

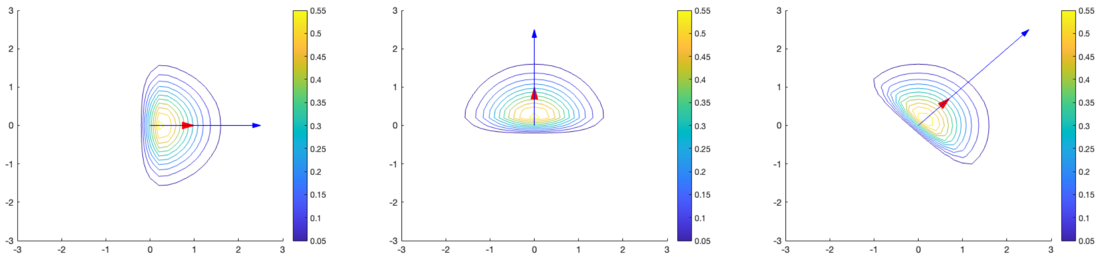
Analogously to 1D,

$$\boldsymbol{\nu}_p = \mathrm{argmin}_{\mathbf{a}\in\mathbb{R}^k} E\left[\|\mathbf{X} - \mathbf{a}\|^p\right],$$

whenever p-th moment of $\mathbf{X}$ is finite, i.e., $\boldsymbol{\nu}_p$ is the **Fréchet** p-**mean**.

Potential applications of multidimensional p-mean in nonlinear regression analysis are being considered.

Multivariate setting.



(a) $\boldsymbol{\lambda} = (5,0)^\top$ (b) $\boldsymbol{\lambda} = (0,5)^\top$ (c) $\boldsymbol{\lambda} = (5,5)^\top$

Consider the multivariate skew normal p.d.f.

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

where $\varphi_k(\cdot\,;\boldsymbol{\mu},\boldsymbol{\Sigma})$ is k-variate normal p.d.f. with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, and $\Phi_1(\cdot)$ is the univariate standard normal distribution.

Vector $\boldsymbol{\lambda}$ is sometimes called the skewness vector.

Bibliography.

• **YK**, "A new life of Pearson's skewness" , *Journal of Theoretical Probability* (2022)

• **YK, A. Negrón, C. Pertel, C. Wang** "Extensions of true skewness for unimodal distributions" , *Mathematical Methods of Statistics* (2024)