An Overview of Sinh-arcsinh Distributions and Their Properties

Arthur Pewsey

University of Extremadura, Spain

Grant: GR24055

Skew 2026, Padua, 7th-9th January 2026

Cofinanciado por la Unión Europea

Fondos Europeos

GOBIERNO DE ESPAÑA

JUNTA DE EXTREMADURA
Consejería de Educación, Ciencia y Formación Profesional

Overview

Skewing approach and goals

Skewing approach

- Perturbation: de Helguero (1908, 1909), Azzalini (1985, ...), ...

$$f(x; \underline{\theta}_1, \underline{\theta}_2) = g(x; \underline{\theta}_2) f_0(x; \underline{\theta}_1). \text{ For example, } f(x; \lambda) = 2\Phi(\lambda x)\phi(x).$$

- Transformation: Edgeworth (1898), Kapteyn (1903), ...

$$Z \sim f_0(z; \underline{\theta}_1), \quad X = h(Z; \underline{\theta}_2).$$

Goals

Develop models:

- that are unimodal;
- have four parameters controlling location, dispersion, skewness and tail-weight;
- which represent departures from a base random variable with varying degrees of skewness and tail-weight.
- with appealing properties.

SAS transformation

The SAS transform and its inverse

- Let Z denote a standardised absolutely continuous base random variable that is symmetric about the origin.
- The canonical sinh-arcsinh (SAS) counterpart of Z, $X_{\varepsilon,\delta}$, is defined through

$$Z = S_{\varepsilon,\delta}(X_{\varepsilon,\delta}),$$

where

$$S_{\varepsilon,\delta}(x) = \sinh\{\delta \sinh^{-1}(x) - \varepsilon\}$$

is the SAS transformation, $\varepsilon \in \mathbb{R}$, and $\delta > 0$.

- Inverting,

$$X_{\varepsilon,\delta} = S^{-1}_{\varepsilon,\delta}(Z) = S_{-\varepsilon/\delta,1/\delta}(Z) = \sinh\{\tfrac{1}{\delta}\sinh^{-1}(Z) + \tfrac{\varepsilon}{\delta}\}.$$

- $X_{\varepsilon,\delta}$ is skew to the left (right) if $\varepsilon < 0$ ($\varepsilon > 0$) and has tails that are heavier (lighter) than those of the base random variable Z if $\delta < 1$ ($\delta > 1$).

Hyberbolic functions

$$\sinh(x) = \frac{e^x - e^{-x}}{2}, \quad \cosh(x) = \frac{e^x + e^{-x}}{2}, \quad \tanh(x) = \frac{\sinh(x)}{\cosh(x)}.$$

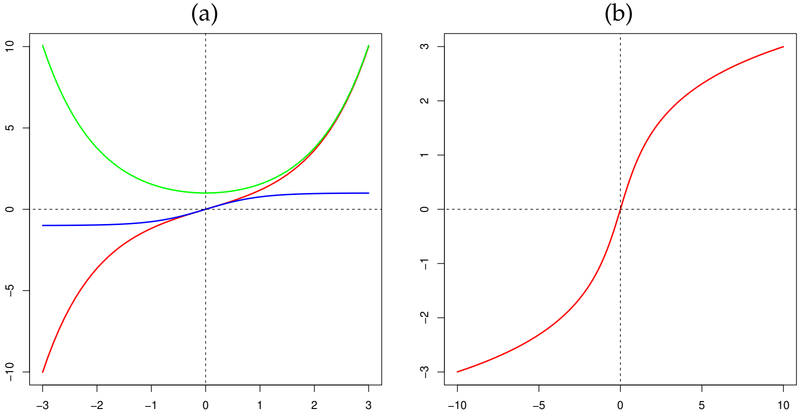

Fig. 1: (a) $\sinh(x)$ (red), $\cosh(x)$ (green) and $\tanh(x)$ (blue) against x for $x \in [-3, 3]$; (b) $\sinh^{-1}(x)$ (red) against x for $x \in [-10, 10]$.

Application to $Z \sim N(0,1)$

$$X_{\varepsilon,\delta} = \sinh\{\tfrac{1}{\delta}\sinh^{-1}(Z) + \tfrac{\varepsilon}{\delta}\}$$

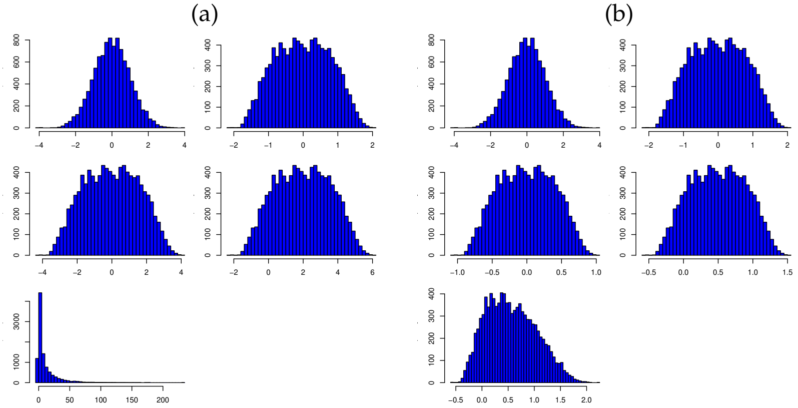

Fig. 2: The five component transformations of $S^{-1}_{\varepsilon,\delta}(Z)$ applied to 10,000 $Z \sim N(0,1)$ random variates using $\varepsilon = 1$ and: (a) $\delta = 1/2$; (b) $\delta = 2$.

General properties of SAS distributions

Distribution and quantile functions

- For the base random variable Z, let $F_Z(z)$ denote its distribution function, $Q_Z(u) = F_Z^{-1}(u), 0 < u < 1$, its quantile function, and $f_Z(z)$ its density function.
- The distribution function of $X_{\varepsilon,\delta}$ is related to that of Z through

$$\begin{aligned} F_{\varepsilon,\delta}(x) &= P(X_{\varepsilon,\delta} \leq x) = P(S_{-\varepsilon/\delta,1/\delta}(Z) \leq x) = P(Z \leq S^{-1}_{-\varepsilon/\delta,1/\delta}(x)) \\ &= P(Z \leq S_{\varepsilon,\delta}(x)) = F_Z(S_{\varepsilon,\delta}(x)). \end{aligned}$$

- Its quantile function is given by

$$Q_{\varepsilon,\delta}(u) = F^{-1}_{\varepsilon,\delta}(u) = S^{-1}_{\varepsilon,\delta}(F_Z^{-1}(u)) = S_{-\varepsilon/\delta,1/\delta}(Q_Z(u)), \;\; 0 < u < 1.$$

- As Z is symmetric about 0, the median of $X_{\varepsilon,\delta}$ is

$$Q_{\varepsilon,\delta}(1/2) = S_{-\varepsilon/\delta,1/\delta}(Q_Z(1/2)) = S_{-\varepsilon/\delta,1/\delta}(0) = \sinh(\varepsilon/\delta).$$

General properties of SAS distributions

Density function

- Differentiating the distribution function $F_{\varepsilon,\delta}(x)$ with respect to x, the density of $X_{\varepsilon,\delta}$ is given by

$$f_{\varepsilon,\delta}(x) = F'_{\varepsilon,\delta}(x) = F'_Z(S_{\varepsilon,\delta}(x))S'_{\varepsilon,\delta}(x) = f_Z(S_{\varepsilon,\delta}(x))S'_{\varepsilon,\delta}(x)$$
$$= \delta \left\{ \frac{1 + S^2_{\varepsilon,\delta}(x)}{1 + x^2} \right\}^{1/2} f_Z(S_{\varepsilon,\delta}(x)).$$

- It is easily shown that $f_{-\varepsilon,\delta}(x) = f_{\varepsilon,\delta}(-x)$.

Properties of SAS-normal distributions

Simulation, distribution, quantile and density functions

- For $Z \sim \mathrm{N}(0,1)$, $F_Z(z) = \Phi(z)$, $Q_Z(u) = \Phi^{-1}(u)$, and $f_Z(z) = \phi(z)$.
- $X_{\varepsilon,\delta} = \sinh\{\frac{1}{\delta}\sinh^{-1}(Z) + \frac{\varepsilon}{\delta}\}$
- $F_{\varepsilon,\delta}(x) = \Phi(S_{\varepsilon,\delta}(x))$
- $Q_{\varepsilon,\delta}(u) = S_{-\varepsilon/\delta,1/\delta}(\Phi^{-1}(u))$
- $f_{\varepsilon,\delta}(x) = \delta\left\{\frac{1+S^2_{\varepsilon,\delta}(x)}{1+x^2}\right\}^{1/2}\phi(S_{\varepsilon,\delta}(x))$
- All SAS-normal densities are unimodal.
- The symmetric models obtained using the transform $S_{0,\delta}(x)$ behave very much like the S_U models of Johnson (1949) when $\delta < 1$, and like the sinh-normal models of Rieck & Nedelman (1991) when $\delta > 1$.

Examples of SAS-normal densities

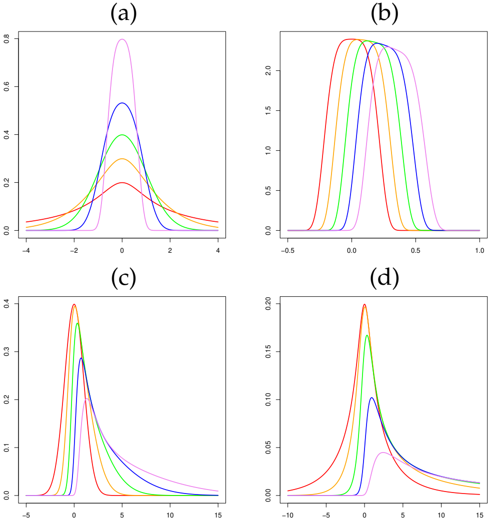

Fig. 3: SAS-normal densities with: (a) $\varepsilon = 0$ and $\delta = \frac{1}{2}, \frac{3}{4}, 1, \frac{4}{3}, 2$; (b) $\delta = 6$ and $\varepsilon = 0, \frac{1}{2}, 1, \frac{3}{2}, 2$; (c) $\delta = 1$ and $\varepsilon = 0, \frac{1}{2}, 1, \frac{3}{2}, 2$; (d) $\delta = \frac{1}{2}$ and $\varepsilon = 0, \frac{1}{2}, 1, \frac{3}{2}, 2$.

Properties of SAS-normal distributions

Moments

$$E(X_{\varepsilon,\delta}) = \sinh(\varepsilon/\delta)P_{1/\delta},$$

$$E(X_{\varepsilon,\delta}^2) = \tfrac{1}{2}\{\cosh(2\varepsilon/\delta)P_{2/\delta} - 1\},$$

$$E(X_{\varepsilon,\delta}^3) = \tfrac{1}{4}\{\sinh(3\varepsilon/\delta)P_{3/\delta} - 3\sinh(\varepsilon/\delta)P_{1/\delta}\},$$

$$E(X_{\varepsilon,\delta}^4) = \tfrac{1}{8}\{\cosh(4\varepsilon/\delta)P_{4/\delta} - 4\cosh(2\varepsilon/\delta)P_{2/\delta} + 3\},$$

where

$$P_q = E[\{Z^2 + (Z^2+1)^{1/2}\}^q] = \frac{e^{1/4}}{(8\pi)^{1/2}}\{K_{(q+1)/2}(1/4) + K_{(q-1)/2}(1/4)\}$$

and K is the modified Bessel function of the second kind.

Skewness-invariant measures of kurtosis

The classic measure of kurtosis, α_4

Moment-based measure of kurtosis:

$$\alpha_4 = \frac{\mu_4}{\sigma^4},$$

where $\mu_k = E[(X-\mu)^k]$, $\mu = E(X)$ and $\sigma^2 = \mu_2$ (Thiele, 1889; Pearson, 1905).

Problems with α_4

- Basically, doesn't measure what it is supposed to measure ("peakedness"): especially, but not exclusively, for asymmetric distributions.
- Undefined for heavy-tailed distributions.
- $\alpha_4 \geq \alpha_3^2 + 1$, where $\alpha_3 = \mu_3/\sigma^3$ (Pearson, 1916). So, higher skewness (as measured by α_3) implies higher kurtosis (as measured by α_4).

Alternative definition of “kurtosis”

Kurtosis: “Location- and scale-free movement of probability mass from the shoulders of a distribution to its centre and tails.” Balanda & MacGillivray (1988).

Quantile-based measures of kurtosis as in JRP (2011)

$$t(p) = \frac{Q(\frac{1}{2}+p) - Q(\frac{1}{2}-p)}{Q(\frac{3}{4}) - Q(\frac{1}{4})}, \quad 0 < p < 1/2,$$ Balanda & MacGillivray (1988).

$$M = \frac{Q(\frac{7}{8}) - Q(\frac{5}{8}) + Q(\frac{3}{8}) - Q(\frac{1}{8})}{Q(\frac{3}{4}) - Q(\frac{1}{4})},$$ Moors (1988).

$$J = \frac{Q(\frac{4}{5}) - 3Q(\frac{3}{5}) + 3Q(\frac{2}{5}) - Q(\frac{1}{5})}{Q(\frac{4}{5}) - Q(\frac{1}{5})},$$ JRP (2011).

$$\tau_4 = \frac{\int_0^1 P_3^*(u)Q(u)du}{\int_0^1 P_1^*(u)Q(u)du},$$ Hosking (1990, 1992),

where $P_1^*(u) = 2u - 1$ and $P_3^*(u) = 20u^3 - 30u^2 + 12u - 1$.

Skewness-invariant measures of kurtosis

For SAS distributions

For all SAS distributions, such quantile-based measures of kurtosis are skewness invariant because

$$S_{-\varepsilon/\delta,1/\delta}(z) - S_{-\varepsilon/\delta,1/\delta}(-z) = 2\cosh(\varepsilon/\delta)\sinh(\sinh^{-1}(z)/\delta).$$

For the SAS-normal model, using $z(p)$ to denote $\Phi^{-1}(\frac{1}{2}+p)$ for $0 < p < 1/2$:

$$t(p) = \frac{\sinh(\sinh^{-1}(z(p))/\delta)}{\sinh(\sinh^{-1}(z(\frac{1}{4}))/\delta)}.$$

$$M = \frac{\sinh(\sinh^{-1}(z(\frac{3}{8}))/\delta) - \sinh(\sinh^{-1}(z(\frac{1}{8}))/\delta)}{\sinh(\sinh^{-1}(z(\frac{1}{4}))/\delta)}.$$

$$J = 1 - 3\frac{\sinh(\sinh^{-1}(z(\frac{1}{10}))/\delta)}{\sinh(\sinh^{-1}(z(\frac{3}{10}))/\delta)}.$$

$$\tau_4 = \frac{\int_0^{1/2}\sinh(\sinh^{-1}(\Phi^{-1}(u))/\delta)P_3^*(u)du}{\int_0^{1/2}\sinh(\sinh^{-1}(\Phi^{-1}(u))/\delta)P_1^*(u)du}.$$

SAS-t distributions

Original SAS-t distribution of RJP (2011)

- As a Student-t random variable with $\nu > 0$ degrees of freedom, T_ν, already has a tail-weight parameter, set $\delta = 1$ in the SAS transformation and define a SAS-t random variable, $T_{\varepsilon,\nu}$, through

$$T_\nu = S_\varepsilon(T_{\varepsilon,\nu}), \tag{1}$$

where $S_\varepsilon(x) = \sinh\{\sinh^{-1}(x) - \varepsilon\}$, with $\varepsilon \in \mathbb{R}$.

- Inverting,

$$T_{\varepsilon,\nu} = S_\varepsilon^{-1}(T_\nu) = S_{-\varepsilon}(T_\nu). \tag{2}$$

- Density of $T_{\varepsilon,\nu}$ is unimodal if $\nu \geq 0.35$. However, for $\nu < 0.35$ there are values of ε for which the density is bimodal.

SAS-t distributions

Eschewed SAS-t distribution of JP (2026)

Define, instead, the Eschewed SAS-t (ESAS-t) random variable $T_{\varepsilon,\nu}$ through

$$T_{\varepsilon,\nu} = \sqrt{\nu}\, S_{-\varepsilon}(T_\nu/\sqrt{\nu}).$$

Denoting the distribution and quantile functions of T_ν by F_ν and Q_ν, respectively, and those of $T_{\varepsilon,\nu}$ by $F_{\varepsilon,\nu}$ and $Q_{\varepsilon,\nu}$, respectively:

- $F_{\varepsilon,\nu}(x) = F_\nu \left\{\sqrt{\nu} S_\varepsilon \left(x/\sqrt{\nu}\right)\right\}$;
- $Q_{\varepsilon,\nu}(u) = \sqrt{\nu} S_{-\varepsilon}\{Q_\nu(u)/\sqrt{\nu}\}$;
- The median of $T_{\varepsilon,\nu}$ is $\sqrt{\nu}\sinh(\varepsilon)$;
- The density of $T_{\varepsilon,\nu}$ is

$$f_{\varepsilon,\nu}(x) = \frac{1}{\sqrt{\nu} B(\frac{\nu}{2}, \frac{1}{2})\{1 + (x^2/\nu)\}^{1/2} \left\{1 + S_\varepsilon^2(x/\sqrt{\nu})\right\}^{\nu/2}},$$

  where $B(\cdot, \cdot)$ denotes the beta function.
- Now, the density of $T_{\varepsilon,\nu}$ is always unimodal.

ESAS-t and SAS-t densities

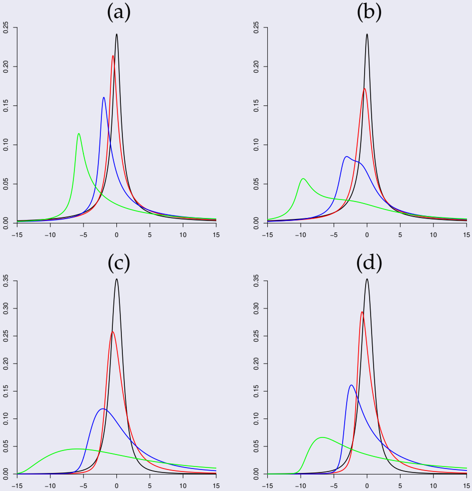

Fig. 4: ESAS-t (left) and SAS-t (right) densities for $\nu = 0.35$ (top) and $\nu = 2$ (bottom). In each panel, $\epsilon = 0$ (black), $\epsilon = 1$ (red), $\epsilon = 2$ (blue), $\epsilon = 3$ (green) and the densities have been re-centred to have zero median.

Moments of ESAS-t distributions

$$E(X_{\varepsilon,\nu}) = \frac{B(\frac{\nu+1}{2}, \frac{\nu-1}{2})}{B(\frac{\nu}{2}, \frac{\nu}{2})} \sqrt{\nu}\sinh(\varepsilon);$$

$$E(X^2_{\varepsilon,\nu}) = \frac{\nu}{\nu-2}\{1 + \nu\sinh^2(\varepsilon)\};$$

$$E(X^3_{\varepsilon,\nu}) = \frac{B\left(\frac{\nu+1}{2}, \frac{\nu-1}{2}\right)}{B\left(\frac{\nu}{2}, \frac{\nu}{2}\right)} \frac{\nu^{3/2}}{(\nu-3)} \sinh(\varepsilon)\{3 + (\nu+1)\sinh^2(\varepsilon)\};$$

and

$$E(X^4_{\varepsilon,\nu}) = \frac{\nu^2}{(\nu-2)(\nu-4)}\{3 + 6\nu\sinh^2(\varepsilon) + \nu(\nu+2)\sinh^4(\varepsilon)\}.$$

Point estimation for SAS distributions

Location-scale extension

- In applied work, one will generally be interested in inference for the parameters of the location-scale extension of $X_{\varepsilon,\delta}$, namely

$$X_{\xi,\eta,\varepsilon,\delta} = \xi + \eta X_{\varepsilon,\delta},$$

  with location and scale parameters, $\xi \in \mathbb{R}$ and $\eta > 0$, respectively.
- $X_{\xi,\eta,\varepsilon,\delta}$ has density $\eta^{-1} f_{\varepsilon,\delta}((x-\xi)/\eta)$.

Maximum likelihood point estimation

- Conducted numerically.
- Approach involving: (i) standardisation of the data; (ii) grid search to identify starting point; (iii) optimization routine; (iv) back transformation.

Goodness-of-fit testing for SAS distributions

Testing set up

- In parametric goodness-of-fit testing we observe a random sample $\mathbf{X}_n = \{X_1, X_2, ..., X_n\}$ drawn from some unknown distribution function, G_X say, and test the composite null hypothesis

$$H_0 \colon G_X \in \mathcal{F} \text{ against } H_1 \colon G_X \notin \mathcal{F},$$

  where $\mathcal{F}$ denotes some specified class of distribution functions.

- Here we suppose that $\mathcal{F} = \{F_X(\cdot;\theta) : \theta \in \Theta\}$ is a specified class of SAS distribution functions (SAS-normal, SAS-t, ...), θ denoting its associated parameter vector.

Goodness-of-fit testing for SAS distributions

Parametric bootstrap algorithm

Let T denote the test statistic of a chosen edf-based goodness-of-fit test.

1. For the original random sample $\mathbf{X}_n = \{X_1, X_2, ..., X_n\}$:
    1. Calculate the ML estimate $\hat{\theta}$, and hence identify $F_X(\cdot;\hat{\theta})$, the distribution function of the best-fitting member of $\mathcal{F}$.
    2. Evaluate T. Denote the value obtained by T_0.
2. Generate B bootstrap samples from $F_X(\cdot;\hat{\theta})$, $\{\mathbf{X}^*_{nj} = \{X^*_{1j}, X^*_{2j}, ..., X^*_{nj}\}, j = 1, 2, ..., B\}$.
3. For the jth bootstrap sample, $\mathbf{X}^*_{nj}$, $j = 1, 2, .., B$:
    1. Calculate the ML estimate $\hat{\theta}^*_j$, and hence identify $F_X(\cdot;\hat{\theta}^*_j)$, the distribution function of the best-fitting member of $\mathcal{F}$.
    2. Evaluate T. Denote the value obtained by T^*_j.
4. A conservative estimate of the p-value of the test is

$$\hat{p} = \frac{(\text{Number of the } T^*_j \geq T_0) + 1}{B + 1}.$$

Goodness-of-fit testing for SAS distributions

Simulation-based results when $\mathcal{F}$ is the SAS-normal class

Edf-based test statistic, T: (i) D (Kolmogorov-Smirnov); (ii) V (Kuiper); (iii) W^2 (Cramér-von Mises); (iv) U^2 (Watson); (v) A^2 (Anderson-Darling).

Size

1. All 5 tests maintain the nominal significance level well when $n = 100$.
2. For $n \leq 50$, tests tend to be conservative.
3. Overall, D-based test maintains the nominal level best.

Power

1. Have relatively low power against alternatives which can be closely approximated by some member of the SAS-normal class.
2. Can be very powerful for samples of suitable size drawn from alternatives with distributional shapes that are very different to those that SAS-normal distributions can adopt.
3. For $n \leq 100$, generally little difference in the power of the tests.
4. For $n \geq 200$, the D-based test generally has relatively low power, and the A^2- and U^2-based tests are often the most powerful.

References

1. JONES, M.C. & PEWSEY, A. (2009) Sinh-arcsinh distributions. *Biometrika*, **96**, 761–780.
2. JONES, M.C., ROSCO, J.F. & PEWSEY, A. (2011) Skewness-invariant measures of kurtosis. *The American Statistician*, **65**, 89–95.
3. ROSCO, J.F., JONES, M.C. & PEWSEY, A. (2011) Skew *t* distributions via the sinh-arcsinh transformation. *Test*, **20**, 630–652.
4. PEWSEY, A. (2018) Parametric bootstrap edf-based goodness-of-fit testing for sinh-arcsinh distributions. *Test*, **27**, 147–172.
5. JONES, M.C. & PEWSEY, A. (2026) The eschewed sinh-arcsinh *t* distribution. *Statistics & Probability Letters*, **228**, 110560.