The L Distribution and Skew Generalizations

Matthias Fischer[1]

Department of Statistics and Econometrics,

University of Erlangen-Nuremberg, Germany

Abstract: Leptokurtic or platykurtic distributions can, for example, be generated by applying certain non-linear transformations to a Gaussian random variable. Within this work we focus on the class of so-called power transformations which are determined by their generator function. Examples are the $H-$transformation of Tukey (1960), the $J-$transformation of Fischer and Klein (2004) and the $L-$transformation which is derived from Johnson's inverse hyperbolic sine transformation. It is shown that generator functions themselves which meet certain requirements can be used to construct both probability densities and cumulative distribution functions. For the $J-$transformation, we recover the logistic distribution. Using the $L-$transformation, a new class of densities is derived, discussed and generalized.

Keywords: Power kurtosis transformation; leptokurtosis; (skew) $L-$distribution.

[1]Department of Statistics and Econometrics, Dr. Matthias Fischer, Lange Gasse 20, D-90403 Nürnberg, Germany, e-mail address: *Matthias.Fischer@wiso.uni-erlangen.de*

1 Introduction

Flexible distribution families which accommodate, for instance, leptokurtosis can be generated, for example, if we transform a Gaussian random variable with certain non-linear transformations. Examples are the $H-$transformation of Tukey (1960), the $K-$transformation of Haynes et al. (1997) or the $J-$transformation of Fischer and Klein (2004). All of these transformations can be embedded in so-called *power transformations* $T(x) = p(x)^r$, where $p(x)$ is the generator function and r can be understood as kurtosis parameter in the sense of preserving the kurtosis ordering of van Zwet (1964). Within this work we restrict these generator functions to the class of so-called density generator functions. It will be verified that the $H-$ and the $K-$generator functions are not member of this class, whereas the $J-$generator function is. Additionally, we propose the $L-$transformation as a special power transformation which is closely related to the inverse hyperbolic sine transformation of Johnson (1949) and show that its generator function is also a density generator function. Moreover, we demonstrate how to derive both a probability density and a cumulative distribution function by means of a density generator. In particular, the logistic density is recovered for the $J-$transformation. Applying the mechanism to the $L-$transformation, a new class of densities is derived. These so-called $L-$distributions are shown to be symmetric and heavy-tailed, with non-existing mean. Finally, we present some skew $L-$distributions derived using the fact that the cumulative distribution is available in closed form.

2 Power Kurtosis Transformations

Let Z be a random variable which is symmetric around the median 0 and which has a continuous distribution function. Define

$$Y = Z \cdot W(Z) \tag{1}$$

where W is a suitable kurtosis transformation. Hoaglin (1983) postulated some plausible requirements to a suitable transformation W of kurtosis. Firstly, W should preserve symmetry, i.e. $\mathrm{W}(z) = \mathrm{W}(-z)$ for $z \in \mathbb{R}$. Hence, we can restrict discussion of W only to the positive axis. Secondly, the initial distribution Z should hardly be transformed in the centre, i.e. $\mathrm{W}(z) \approx 1$ for $z \approx 0$. Finally, in order to increase the tails of the distribution, we have to assure that W is accelerated strictly monotone increasing for positive $z > 0$, i.e. $W'(z) > 0$ and $W''(z) > 0$ for $z > 0$. Consequently, W is strictly monotone increasing and convex for $z > 0$. Conversely, a shortening of the tails takes place, either if W is strictly monotone increasing with negative second derivation or if W is not monotone but concave for $z > 0$. Differentiability and monotonicity imply that $W'(0) = 0$.

Example 2.1 *Kurtosis transformations which satisfy the aforementioned conditions are:*

1. *The $H-$transformation of Tukey (1960):* $H(z) = \exp(1/2z^2)^h$ *for* $h \in \mathbb{R}$,
2. *The $J-$transformation of Fischer and Klein (2004):* $J(z) = \cosh(z)^j$ *for* $j \in \mathbb{R}$ *with* $\cosh(z) = 0.5(e^z + e^{-z})$,
3. *The $K-$transformation of Haynes et al. (1997):* $K(z) = (1 + z^2)^k$ *for* $k \in \mathbb{R}$.

The $H-$, $J-$ and $K-$transformation can be embedded in so-called *power transformations* which are defined next.

Definition 2.1 (Power transformation) *A kurtosis transformation is called a power transformation if it admits a representation*

$$W(z;p) = (p(z))^r, \quad r \in \mathbb{R},$$

where $p(z) = p(-z)$*,* $p(0) = 1$*,* $p'(z) \geq 0$ *for* $z > 0$ *and* $p''(z) > 0$ *for* $z > 0$*. The function* p *will be termed as the generating function of the power transform.*

Next, we introduce the $L-$transformation which will play the leading part within this work.

Example 2.2 (L-transformation) *Originally, Johnson (1949) recommends using the inverse hyperbolic sine (IHS) transformation* $Y = \sinh(Z/l)$ *instead of* $Z \cdot W(Z)$ *from (1), where* l *serves as kurtosis parameter. Setting* $l = 1$*, we can rewrite this equation as*

$$Y = Z \cdot L(Z) \textit{ with } L(Z) = \frac{\sinh(Z)}{Z} \textit{ and } \sinh(Z) = 0.5(e^Z - e^{-Z}).$$

The transformation $L(z) = (\sinh(z)/z)^l$ *will be called the* $L-$*transformation in the following.*

3 Distribution Generating Functions and corresponding Distributions

In a first step, we now restrict the class of generating functions to those functions which dominate their first derivative, but coincide with the first derivate in the limit.

Definition 3.1 (Distribution generating function) *Assume $W(x) = p(x)^r$ is a power kurtosis transformation with generating function $p(x)$. The function $p(x)$ is called a distribution generating function (dgf) if the following three properties are satisfied:*

(D1) $p(x) \in C^2(\mathbb{R})$,

(D2) $p'(x) \leq p(x), \quad x \geq 0$,

(D3) $\lim_{x\to\infty} \frac{p'(x)}{p(x)} = 1.$

Example 3.1 *Revisiting the examples from the previous section, we obtain the following results:*

1. *$H-$transformation: The generating function is given by $p_H(x) = \exp(0.5x^2)$. Hence, $p'_H(x) = x\exp(0.5x^2)$ and (D2) is violated for $x > 1$, i.e. $p_H(x)$ is no dgt.*

2. *$J-$transformation: The generating function is given by $p_J(x) = \cosh(x) = \frac{e^x+e^{-x}}{2}$. Because of $p'_J(x) = \sinh(x) = 0.5(e^x - e^{-x})$, (D2) and (D3) are satisfied and $p_J(x)$ is a dgt. Note that (D1) is trivial.*

3. *$K-$transformation: From $p_K(x) = 1+x^2$ we conclude that $p'_K(x) = 2x$. It follows immediately that (D3) is not valid and $p_K(x)$ is no dgt.*

4. *$L-$transformation: Using $p_L(z) = \frac{\sinh(x)}{x}$, we obtain $p'_L(x) = \frac{\cosh(x)x-\sinh(x)}{x^2}$. Hence, for $x > 0$,*

$$
\begin{aligned}
p_L(x) - p'_L(x) &= \frac{\sinh(x)}{x} - \frac{\cosh(x)x - \sinh(x)}{x^2} \\
&= \frac{x\sinh(x) - \cosh(x)\,x + \sinh(x)}{x^2} \\
&= \frac{(x+1)\sinh(x) - x\cosh(x)}{x^2} > 0
\end{aligned}
$$

because $\exp(x) - (2x+1)\exp(-x) > 0$. *Moreover,*

$$\lim_{x\to\infty} \frac{p_L'(x)}{p_L(x)} = \lim_{x\to\infty} \frac{\cosh(x)\,x - \sinh(x)}{x\sinh(x)} = \lim_{x\to\infty} (\coth(x) - 1/x) = 1.$$

Thus (D1), (D2) and (D3) are satisfied and $p_L(x)$ *is a dgt.*

Definition 3.2 *Assume that* $p(x)$ *is distribution generating function. Then we define*

$$F(x;p) = \frac{1}{2}\left(\frac{d\log(p(x))}{dx} + 1\right). \tag{2}$$

Note that $F(x;p) = \frac{1}{2}\left(\frac{p'(x)}{p(x)} + 1\right)$. The next lemma verifies that $F(x;p)$ is a cumulative distribution function which is "generated" by p. This explains where the name "distribution generating function" comes from.

Lemma 3.1 $F(x;p)$ *is a cumulative distribution function on* $\mathbb{R}$ *for every distribution generating function* p.

Proof: Using the symmetry of $p(x)$, we can concentrate on $[0,\infty)$. From $\lim\limits_{x\to\infty} \frac{p'(x)}{p(x)} = 1$ we conclude that $\lim\limits_{x\to\infty} F(x;p) = 0.5(1+1) = 1$, too. $F(x;p)$ is strictly increasing because of $(D1), (D2), (D3)$ and the symmetry of F. $\square$

Definition 3.3 (LDGF Distribution) *Assume that* $p(x)$ *is distribution generating transformation (dgt). The distribution associated with Lemma 3.1 will be called a LDGF distribution in the sequel. According to (D1), the corresponding density is well-defined and given by*

$$f(x;p) = \frac{dF(x;p)}{dx} = \frac{1}{2}\frac{d^2\log(p(x))}{dx^2}. \tag{3}$$

If we consider the density generating function of the $J-$transformation, the corresponding LDGF distribution is identical to the logistic distribution.

Example 3.2 (Logistic distribution) *From example 3.1 we knew that the generating function of the $J-$transformation is a distribution generating function. Plugging $p_J(x) = \cosh(x)$ into (3), we get*

$$F(x;p) = \frac{1}{2}\left(\frac{d\log(\cosh(x))}{dx} + 1\right) = \frac{\cosh(x) + \sinh(x)}{2\cosh(x)}$$

with corresponding density given by

$$f(x;p) = \frac{1}{2\left(\cosh(x)\right)^2} = \frac{2}{(\exp(x) + \exp(-x))^2} = \frac{2\exp(2x)}{(\exp(2x) + 1)^2},$$

i.e. we recover the logistic density.

4 The L Distribution: Definition and Properties

The focus of this section is on the $L-$transform from example 2.1 with distribution generating function $p(x) = p_L(x) = \sinh(x)/x$. Let us first derive some properties of the function $p(x)$:

Lemma 4.1 *Assume that* $p(x) = \sinh(x)/x$.

1. *The function* $p(x)$ *is continuous on* $\mathbb{R}$ *with* $\lim\limits_{x \to 0\pm} p(x) = 1$.

2. *The power series representation is given by*

$$p(x) = 1 + \frac{x^2}{3!} + \frac{x^4}{5!} + \ldots$$

3. *The function* $p(x)$ *is symmetric:* $p(x) = p(-x)$ *for all* $x \in \mathbb{R}$.

4. *For all* $x \in \mathbb{R}$ *holds the inequality* $\cosh(x) \geq p(x)$.

5. *The first derivative is given by* $p'(x) = \frac{\cosh(x) - p(x)}{x} > 0$ *for* $x \geq 0$.

Proof: The first property follows direct from the rule of l'Hospital with $\sinh(x)' = \cosh(x)$ and $\cosh(0) = 1$. To derive property 2, divide the power series representation $\sinh(x) = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots$ by x. Hence, property 3 is obvious. Property 4 follows from $\cosh(x) = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots$ which itself implies property 5.□

A plot of $p(x)$ and $p'(x)$ is given in figure 6, below.

Figure 1 to be inserted here

Definition 4.1 (L distribution) *Plugging $p(x)$ into (3), we get the density*

$$f(x) = \frac{x^2 + \left(1 - (\cosh(x))^2\right)}{2x^2\left(1 - (\cosh(x))^2\right)} = \frac{\sinh(x)^2 - x^2}{2x^2 \sinh(x)^2} = \frac{p(x)^2 - 1}{2\sinh(x)^2}, \quad x \in \mathbb{R}.$$

The distribution belonging to this symmetric density will be called a L distribution in the sequel.

According to (2), the corresponding cumulative distribution function is given by

$$F(x) = \frac{\cosh(x)\,x - \sinh(x) + x\sinh(x)}{2x\sinh(x)}, \quad x \in \mathbb{R}. \tag{4}$$

Tedious but straightforward calculations allow us to determine the score function which is plotted in figure 2(b). The decreasing behavior for $x \to \pm\infty$ stems from the heavier tails of the distribution.

Lemma 4.2 (Score function) *The score function of the L distribution is given by*

$$\psi(z) = -\frac{f'(x)}{f(x)} = 2\,\frac{\cosh(x)\,x^3 + \sinh(x) - \sinh(x)(\cosh(x))^2}{x\sinh(x)\left(x^2 + 1 - (\cosh(x))^2\right)}.$$

Figure 2 to be inserted here

Lemma 4.3 (Tail function) *The tail function $T(x) = 1 - F(x)$ is given by*

$$T(x) = \frac{1}{2z} - \frac{\exp(-x)}{\exp(x) - \exp(-x)} \approx \frac{1}{2x} \text{ for large } x.$$

Proof: Using (4),

$$\begin{aligned} T(x) &= 1 - \frac{\cosh(x)\,x - \sinh(x) + x\sinh(x)}{2x\sinh(x)} \\ &= 1 - \frac{1}{2}\coth(x) + \frac{1}{2x} - \frac{1}{2} = \frac{1}{2z} - \frac{\exp(-x)}{\exp(x) - \exp(-x)} \quad \square \end{aligned}$$

Lemma 4.4 *The expectation value of a $L-$distribution doesn't exist.*

Proof: Consider the integral

$$\int_0^\infty xf(x)dx = \int_0^\infty x\frac{\sinh(x)^2 - x^2}{2x^2\sinh(x)^2}dx = \int_0^\infty \frac{1}{2x}dx + \int_0^\infty \frac{x}{2\sinh(x)^2}dx \to \infty,$$

because the first integral tends to ∞ and the second is $\pi^2/8$. $\square$

5 Skew L Distributions

The L distribution is a unimodal, symmetric distribution family. Exploiting the fact that the cumulative distribution function is available in closed-form, we finally introduce skewness by means of techniques summarized in Ferreira and Steel (2004). All methods use weighting functions to incorporate skewness into an originally symmetric density.

Following the proposal of Jones (2004), a new density can be obtained using order statistics via $f(x;\beta_1,\beta_2) = \frac{1}{B(\beta_1,\beta_2)}f(x)F(x)_1^\beta(1-F(x))^{\beta_2}$. Applied to the L distribution, the corresponding density of the first skew L (SL1) distribution is given by

$$f_{SL1}(x;\beta_1,\beta_2) = \frac{\left(\frac{\cosh(x)x-\sinh(x)+x\sinh(x)}{2x\sinh(x)}\right)^{\beta_1}\left(1-\frac{\cosh(x)x-\sinh(x)+x\sinh(x)}{2x\sinh(x)}\right)^{\beta_2}}{\frac{2B(\beta_1,\beta_2)x^2\sinh(x)^2}{(\sinh(x)^2-x^2)}}$$

with symmetry for $\beta_1 = \beta_2$. Secondly, hidden truncation models initiated by Azzalini (1985) with building rule $f(x;\lambda) = 2f(x)F(\lambda x)$ lead to the SL2 distribution with pdf

$$f_{SL2}(x;\lambda) = \frac{\left((\cosh(z))^2 - 1 - z^2\right)\left(\cosh(\lambda z)\lambda z - \sinh(\lambda z) + \lambda z\sinh(\lambda z)\right)}{2z^3\left((\cosh(z))^2 - 1\right)\lambda\sinh(\lambda z)}.$$

Note that symmetry is now achieved for $\lambda = 0$. Following the suggestion of Fernández and Steel (1998), skewness can be introduced using inverse scale factors in the positive and negative orthant. Thus, the SL_3 density is defined by

$$f_{SL3}(x;\gamma) = \frac{2}{\gamma+\frac{1}{\gamma}} f(x\gamma^{-\mathrm{sign(x)}}) = \frac{2}{\gamma+\frac{1}{\gamma}} \frac{p_L(x\gamma^{-\mathrm{sign(x)}})^2-1}{2\sinh(x\gamma^{-\mathrm{sign(x)}})^2}, \quad \gamma \in (-1,1),$$

where symmetry corresponds to the case $\gamma = 1$. Finally, the scaling factor approach used, for example, by Hansen (1994) results in the SL_4 density

$$f(x;\eta) = f(x/(1+\mathrm{sign(x)}\eta)) = \frac{\mathrm{p_L(x/(1+sign(x)}\eta))^2-1}{2\sinh(\mathrm{x/(1+sign(x)}\eta))^2}, \quad \eta > 0$$

with symmetry for $\eta = 1$. Note that both applications and further properties of the L distribution and their skew counterparts are factored out to future research.

6 Summary

Power transformations – like the $H-$transformation of Tukey (1960) or the $J-$transformation of Fischer and Klein (2004) – are used to generate leptokurtic distributions by means of variable transformation. They are characterized by so-called generator functions. We show that generator functions which meet certain requirements can be used to construct both probability densities and cumulative distribution functions. For the $J-$transformation, we recover the logistic distribution. After introducing the $L-$transformation, a new class of densities is derived. The $L-$distributions are symmetric, heavy-tailed with non-existing expectation value. Moreover, skew versions of the L distributions are introduced which exploit the closed-form of the cumulative distribution function.

References

[1] Azzalini, A. (1985), *Kurtosis and Spread*, Scandinavian Journal of Statistics **12**, 171-178.

[2] Balanda, K. P., MacGillivray, H. L. (1990), *Kurtosis and Spread*, The Canadian Journal of Statistics **18**(1), 17-30.

[3] Ferreira, J. T. A. S., Steel, M. F. J. (2004), *A Constructive Representation of Univariate Distributions*, Working Paper, Department of Statistics, University of Warwick.

[4] Fernández, C., Steel, M. F. J. (1998), *On Bayesian Modelling of Fat Tails and Skewness*, Journal of American Statistical Association **93**, 359-371.

[5] Fischer, M., Klein, I. (2004), *Kurtosis Modelling by means of the J-Transformation*, Allgemeines Statistisches Archiv **88**(1), 35-50.

[6] Hansen, B. E. (1994), *Autoregressive Conditional Density Estimation*, International Economic Review **35**(3), 705-730.

[7] Haynes, M. A., MacGillivray, H. L., Mengersen, K. L. (1997), *Robustness of Ranking and Selection Rules using Generalized g-and-k Distributions*, Journal of Statistical Planning and Interference **65**, 45-66.

[8] Hoaglin, D. C. (1983), *Summarizing shape numerically: The $g-$and$-h$ distributions*, in: D. C. Hoaglin and F. Mosteller and J. W. Tukey, eds., Data analysis for tables, trends, and shapes (Wiley, New York), 461-513.

[9] Johnson, N. L. (1949), *Systems of Frequency Curves Generazed by Methods of Translation*, Biometrika **36**, 149-176.

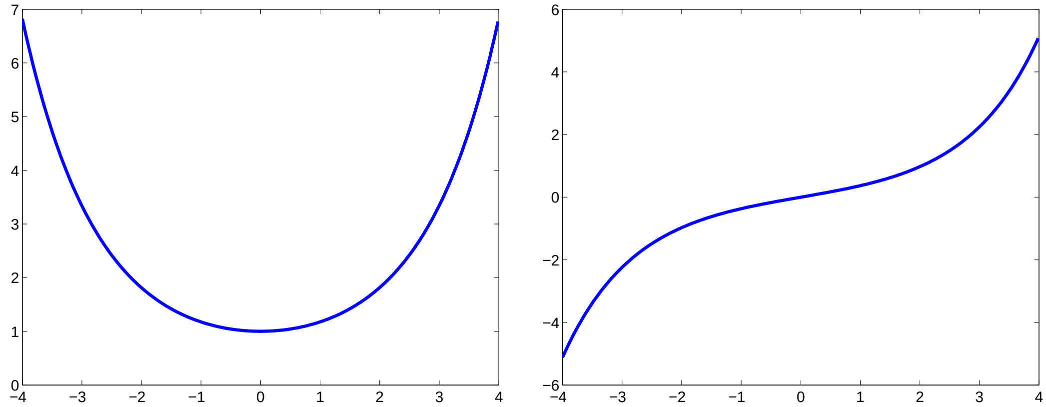
[10] Jones, M. C. (2004), *Families of Distributions arising from Distributions of Order Statistics*, Test, to appear.

[11] Rayner, G. D., MacGillivray, H. L. (2002), *Weighted quantile-based estimation for a class of transformation distributions*, Computational Statistics & Data Analysis **39**, 401-433.

[12] Tukey, J. W. (1960), *The Practical Relationship between the Common Transformations of Counts of Amounts*, Princeton University Statistical Techniques Research Group, Technical Report No. 36.

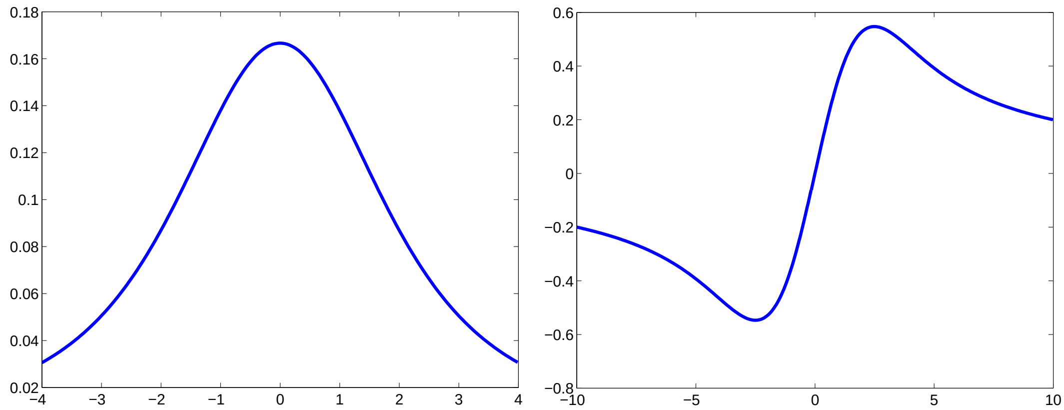
[13] Van Zwet, W. R. (1964), *Convex Transformations of Random Variables*, Mathematical Centre Tracts No. 7. Mathematical Centre, Amsterdam.

Figure 1: The generating function $p(x)$ and its first derivative



(a) $p(x)$ (b) $p'(x)$

Figure 2: L distribution



(a) Density function (b) Score function