Kurtosis modelling by means of the J-transformation

Matthias Fischer and Ingo Klein
Department of Statistics and Econometrics,
University of Erlangen-Nuremberg

Abstract: The $H-$family of distributions or $H-$distributions, introduced by Tukey (1960, 1977), are generated by a single transformation of the standard normal distribution and allow for leptokurtosis represented by the parameter h. Alternatively, Haynes, MacGillivray and Mengersen (1997) generated leptokurtic distributions by applying the $K-$transformation to the normal distribution. In this study we propose a third transformation – the so-called $J-$transformation – and derive some properties of this transformation. Moreover, so-called elongation generating functions (EGF's) are introduced. By means of EGF's we are able to visualize the strength of tail elongation and to construct new transformations. Finally, we compare the three transformations towards their goodness-of-fit in the context of financial return data.

Keywords: kurtosis; variable transformation; normal transformation; tail elongation.

1 Introduction

Using the Gaussian distribution as a statistical model for data sets is widespread, especially in practice. However, departure from normality seems to be more the rule than the exception. Take, for instance, the distribution of continuous returns (i.e. differences of consecutive log-prices) of financial data which displays more kurtosis than that permitted under the assumption of normality (cf. Fama, 1965). Roughly the same phenomenon can be observed for the mass-size distribution of aeolian sand deposits (cf. Barndorff-Nielsen, 1977). In order to construct distributions which are more leptokurtic than the normal distribution, several methods have been developed in the statistical literature. So-called normal-variance mixtures are very popular, where the scale-parameter of a Gaussian distribution itself is assumed to follow a distribution on the positive axis. For example, mixing the zero-mean normal distribution with the generalized inverse Gaussian distribution leads to the symmetric hyperbolic distribution. Alternatively, a non-linear transformation can be applied to a standard normal distribution to obtain a more flexible distribution family. This approach dates back to Tukey (1960, 1977), who introduced the $H-$transformation, where a parameter h controls the amount of kurtosis and elongation, respectively. One property of the $H-$transformed normal distribution ("$H-$distribution") is that moments exist only up to a certain order (see also MacGillivray, 1981, MacGillivray and Belanda, 1988 and Martinez and Iglewicz, 1984). Haynes et al. (1997) proposed another transformation, the so-called $K-$transformation, which exhibits similar properties than the $H-$transformation, but ensures that all moments of the $K-$transformed normal distribution ("$K-$distribution") exist. However, empirical studies of leptokurtic data show (cf. Fischer et al., 2003) that the fit of $K-$distributions is worse than that of the $H-$distribution, especially in the tails. The aim of this paper is "to bridge this gap", i.e. to introduce a transformation – we call it $J-$transformation – that induces a distribution with existing moments but with a similar goodness-of-fit than the $H-$distribution. By means of so-called elongation generating functions we show that the strength of tail elongation of the $J-$transformation is less than that of the $H-$transformation but higher than that of the $K-$transformation.

2 Elongation versus kurtosis

According to Hoaglin (1984, p. 148) and probably also to Tukey, *elongation* is closely related to the notion of ”tail strength”. Investigating the elongation of data or distributions means comparing the tail strength of empirical or theoretical distributions with the tail strength of the Gaussian or normal distribution. I.e. while tail strength is an absolute concept, elongation is (through the comparison with the normal distribution) a relative concept. The normal distribution has an undefined tail strength, but a neutral elongation which is assigned to zero by a suitable elongation measure. In general, transformations which shorten the tails can be considered, too. However, within this work, we focus on tail-increasing transformations or, equivalently, on elongation measures which are required to be positive. Note that elongation is one component of the shape of distributions which is independent of location and scale.

On the other hand, the notion of *kurtosis* is not uniquely defined in the literature. Originally, kurtosis was identified with the fourth standardized moment which should serve as a measure for the ”sharpedness” or the ”peakedness” of a distribution (see, for example, Oja (1981), p. 165). Kaplansky (1945) has already exemplified that the fourth standardized moment does not preserve a peakedness order. In the sense of Finuncan (1964), the fourth standardized moment is a measure for ”a prominent peak and a prominent tail”, whereas Ali (1974) reduces this notion to a measure of tail strength. Darlington (1970) even speaks of a measure of bi-modality. At the latest in the work of Oja (1981) kurtosis is discussed apart from the notion of the fourth standardized moment. Oja discusses a kurtosis model, introduces a kurtosis ordering and finally shows that the fourth standardized moment preserves that ordering under certain conditions and therefore can be seen as a specific kurtosis measure. The kurtosis model of Oja (1981) is based on van Zwet (1964), who introduced a partial ordering of kurtosis $\preceq_S$ on the set of symmetric distribution functions $\mathcal{F}^s$. Let $F, G \in \mathcal{F}^s$ and μ_F denote the location of symmetry of F, then $\preceq_S$ is defined by

$$\textbf{(A)} \quad F \preceq_S G :\Longleftrightarrow \; G^{-1}(F(x)) \; \text{ is convex for } x > \mu_F$$

and means that G has higher kurtosis than F. Balanda and MacGillivray (1990) generalized this partial ordering of van Zwet by using so-called spread functions defined as symmetric differences of quantiles:

$$S_F(u) = F^{-1}(u) - F^{-1}(1-u), \;\; u \geq 0.5.$$

In the sense of Balanda and MacGillivray (1990), an arbitrary continuous, monotone increasing distribution function F has less kurtosis than an equal distribution function G if

$$\textbf{(B)} \quad F \preceq_{S^*} G :\Longleftrightarrow \; S_G(S_F^{-1}(x)) \quad \text{ is convex for } x > F^{-1}(0.5).$$

If F is symmetric, $F^{-1}(u) = -F^{-1}(1-u)$ for $u > 0.5$, so that $S_F(u) = 2F^{-1}(u) \;\; u \geq 0.5$. This means that the spread function essentially coincides with the quantile function. It can be shown that (A) and (B) coincide in this case. Furthermore, Balanda und MacGillivray (1992, p. 1234) use kurtosis (in a very broad sense) as tail strength, peakedness or similar concepts.

Groeneveld (1998) states a whole class of quantile-based kurtosis measures which preserves the kurtosis ordering mentioned above. It is generally accepted that kurtosis cannot be characterized only by the fourth standardized moments. One component of this concept is the tail strength which is also denoted by elongation, if a comparison is based on the tail strength of the normal distribution. Therefore, elongation measures are specific kurtosis measures.

3 Elongation transformations: A review

Let Z be a standard normal variate. Note that most of the results can be also derived for a random variable which is symmetric around the median 0 and which has continuous distribution function. Define

$$X \equiv \mathrm{T}(Z) = Z \cdot W(Z) \tag{1}$$

where T is a suitable elongation transformation. Hoaglin (1983) postulated some plausible requirements to T. Firstly, T should preserve symmetry, i.e. $\mathrm{T}(z) = \mathrm{T}(-z)$ for $z \in \mathbb{R}$ and we therefore have to discuss T only on the positive axis. Secondly, the initial distribution T should hardly be transformed in the centre, i.e. $\mathrm{T}(z) = z + O(z^2)$ for $z \approx 0$. Finally, in order to increase the tails of the distribution, we have to assure that T is accelerated strictly monotone increasing for positive $z > 0$, i.e. $\mathrm{T}'(z) > 0$ and $\mathrm{T}''(z) > 0$ for $z > 0$. Consequently, T is strictly monotone increasing and convex for $z > 0$. Conversely, a shortening of the tails takes place, either if T is strictly monotone increasing with negative second derivation or if T is not monotone but concave for $z > 0$. Differentiability and monotony imply that $\mathrm{T}'(0) = 0$. An example which satisfies the aforementioned conditions is the $H-$transformation of Tukey (1960, 1977) given by

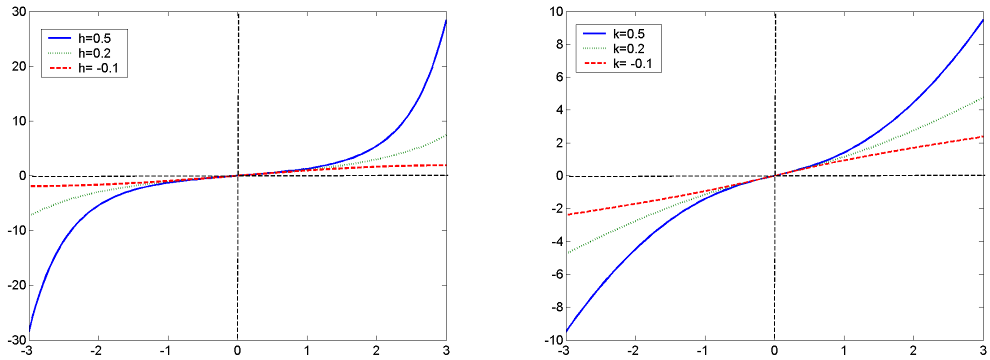
$$\mathrm{H}_h(z) \equiv z \exp(hz^2/2), \quad z, h \in \mathbb{R}. \tag{2}$$

The corresponding distribution of X from (1) is termed as *family of $H-$distributions*, or simply as *$H-$distribution*. $\mathrm{H}_h(Z)$ introduces elongation through the factor h: In the normal case, the distribution of X is leptokurtic for $h > 0$ and platykurtic for $h < 0$. The amount of kurtosis is determined by the parameter h. For $h < 0$, the support of the random variable X is a finite interval and the distribution of X is U-shaped (cf. Klein and Fischer, 2002). A special case of the $H-$distribution is the normal distribution ($h = 0$). Moreover, moments of X only exists up to order $n < 1/h$. Haynes et al. (1997) introduced another elongation transformation ("$K-$transformation") by

$$\mathrm{K}_k(z) \equiv z(1 + z^2)^k, \quad z, k \in \mathbb{R}, \tag{3}$$

where the elongation is governed by the parameter k. Different $H-$ and $K-$transformations are plotted in figure 1, below. It can be proved that all moments of $K-$transformed normal distributions exist.

Figure 1: Elongation transformations for different parameter values.



(a) $H-$transformation (b) $K-$transformation

4 J-transformation: Definition and properties

Basic elements of the $J-$transformation are the *hyperbolic cosine*, the *hyperbolic sine* and the *hyperbolic tangens* function which are given by

$$\cosh(z) \equiv \frac{e^z + e^{-z}}{2}, \ \sinh(z) \equiv \frac{e^z - e^{-z}}{2} \text{ and } \tanh(z) \equiv \frac{\sinh(z)}{\cosh(z)}.$$

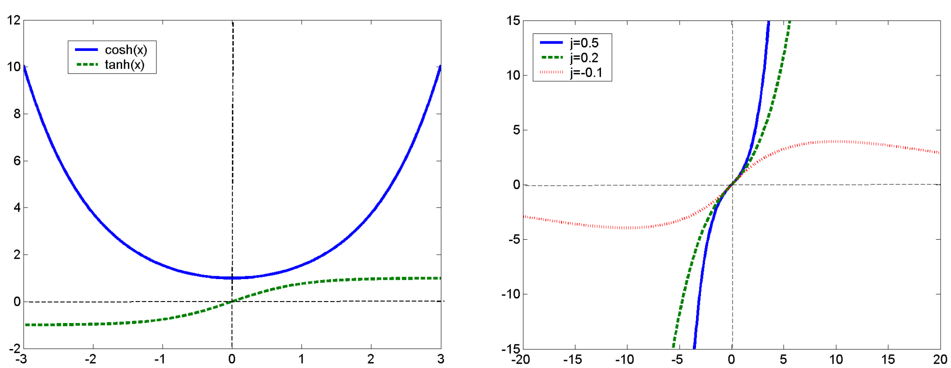
The corresponding graphs can be seen in figure 2(a), below. Note that $\cosh(z)' = \sinh(z)$ and $\tanh(z)' = 1 - \tanh(z)^2$. Next, the $J-$transformation will be defined by means of the hyperbolic cosine function.

Definition 1 ($J-$transformation) *For $z, j \in \mathbb{R}$, the $J-$transformation is defined by*

$$\mathrm{J}_j(z) \equiv z\cosh(z)^j = z\left(\frac{\exp(z) + \exp(-z)}{2}\right)^j. \tag{4}$$

For $j = 0$, $\mathrm{J}_0(z)$ coincides with the bisecting line. For $j > 0$, $\lim_{z\to\infty} \mathrm{J}_j(z) = \infty$ and $\lim_{z\to-\infty} \mathrm{J}_j(z) = -\infty$. On the contrary, for $j < 0$, $\lim_{z\to\infty} \mathrm{J}_j(z) = \lim_{z\to-\infty} \mathrm{J}_j(z) = 0$. Typical curves for $j = 0.2, 0.5, -0.1$ can be seen in figure 2(b), below.

Figure 2: Hyperbolic functions and $J-$transformations.



(a) Hyperbolic functions (b) $J-$transformation

Lemma 1 (Derivatives of the $J-$transformation) *The first two derivatives of* J_j *are*

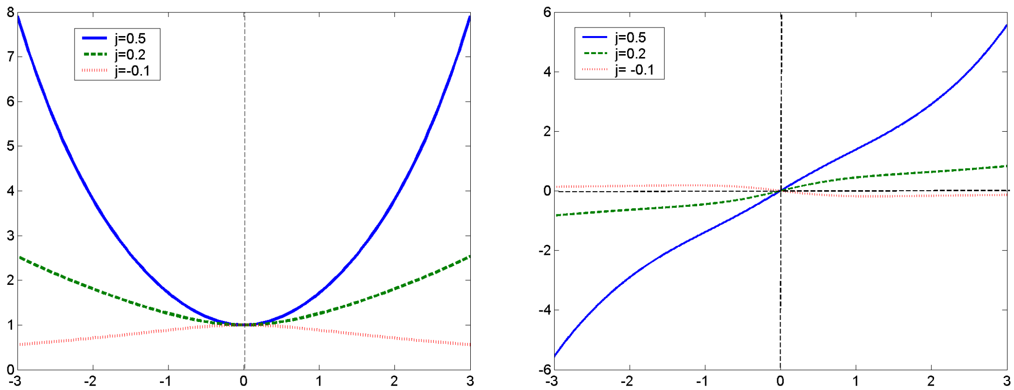
$$\mathrm{J}_j'(z) = \cosh(z)^j + jz\cosh(z)^j\tanh(z) = \cosh(z)^j(1 + jz\tanh(z)) \tag{5}$$

and

$$\mathrm{J}_j''(z) = j\cosh(z)^j\left(z + 2\tanh(z) + z\tanh(z)^2(j-1)\right). \tag{6}$$

Because $z\tanh(z)$ is non-negative for all $z \in \mathbb{R}$, $\mathrm{J}_j'(z) \geq 1$ for positive j. In this case, $\lim_{z\to\infty}\mathrm{J}_j'(z) = \lim_{z\to-\infty}\mathrm{J}_j'(z) = \infty$. Let z_p^* denote the positive root of $1 + jz\tanh(z)$. For $j < 0$, the first derivative is both positive iff $|z| < z_p^*$ and negative iff $|z| > z_p^*$. Now, $\lim_{z\to\infty}\mathrm{J}_j'(z) = \lim_{z\to-\infty}\mathrm{J}_j'(z) = 0$. Consequently, J_j isn't a one-to-one mapping with maximum at z_p^* and minimum at $-z_p^*$. Some curves of J_j' and J_j'' are illustrated in figure 3.

Figure 3: Derivatives of the $J-$transformation.



(a) 1st derivative (b) 2nd derivative

Note that the inverse mapping $\mathrm{J}_j^{-1}(x)$ of $\mathrm{J}_j(z)$, namely

$$\mathrm{J_j}^{-1}(x) \equiv \{x|f(x|z) = x\cosh(x)^j - z = 0\}$$

has no closed form and therefore be approximated numerically.

Lemma 2 ($\mathrm{J}-$ versus $\mathrm{H}-,\mathrm{K}-$ transformation) *Suppose* $j = h = k > 0$ *and* $c^* \approx 2.98$. *The following relations hold between the $J-$transformation on the one hand and the $H-/K-$transformation on the other hand:*

$$|\mathrm{J}_j(z)| \leq |\mathrm{H}_h(z)| \;\textit{ for }\; j = h \;\textit{ and }\; z \in \mathbb{R}. \tag{7}$$

$$|\mathrm{J}_j(z)| \begin{cases} \leq |\mathrm{K}_k(z)| & \textit{for } j = k \textit{ and } z \in [-c^*, c^*], \\ \geq |\mathrm{K}_k(z)| & \textit{for } j = k \textit{ and } |z| > c^*. \end{cases} \tag{8}$$

Proof: Suppose $z \geq 0$ and $h = j \equiv c > 0$. Then we have to show that $H_c(z) - J_c(z) \geq 0$. From (2) and (4), this difference is given by

$$z\exp(0.5cz^2) - z\cosh(z)^c = z\exp(0.5cz^2) - z\exp(c\ln(\cosh(z))).$$

It is sufficient to show that $D(z) \equiv 0.5z^2 - \ln(\cosh(z)) \geq 0$. This, however, follows from $D(0) = 0$, $\lim_{z\to\infty} D(z) \geq 0$ and $D'(z) = z - \tanh(z) \geq 0$ for $z \geq 0$.
Similarly, using $\mathrm{K}_k(z) = z\exp(k\ln(1+z^2))$, equation (8) can be verified. $\square$

5 Generating elongation transformations by means of elongation generating function

First we introduce the class of elongation generating function which can be used to compare the strength of elongation for different transformations and to construct new transformations.

Definition 2 (Elongation generating function) *A function $f : \mathbb{R} \to \mathbb{R}$ is called an elongation generating function (EGF) or of class Υ if the following requirements are satisfied:*

E1 *Smoothness: f is a C^2-function.*

E2 *Anti-symmetry: $f(-z) = -f(z)$.*

E3 *Positivity on $\mathbb{R}_+$: $f(z) > 0$ for $z > 0$.*

E4 *Tail elongation condition: $z\frac{f'(z)}{f(z)} \geq -2$ for $z > 0$.*

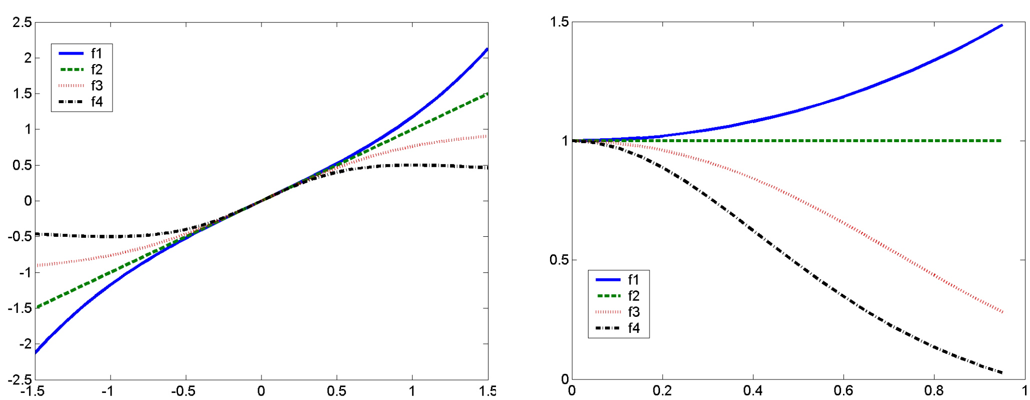
Note that condition **E4'** ($f'(z) > 0$ for $z > 0$) together with **E3** imply condition **E4** which ensures that the second derivation of the corresponding transformation will be positive (cp. Theorem 1). Because of $z\frac{f'(z)}{f(z)} = \frac{\frac{df}{dz}}{\frac{f}{z}}$, **E4** can be interpreted as an elasticity condition. Moreover, **E1** and **E2** imply that $f(0) = 0$.

Example 5.1 (EGF's) *Functions which belong to Υ are*

- $f_1(z) = \sinh(z)$ *("elongation generating function of exponential-type"),*
- $f_2(z) = z$ *("linear elongation generating function"),*
- $f_3(z) = \tanh(z)$ *("asymptotic constant elongation function") and*
- $f_4(z) = \frac{z}{1+z^2}$ *("asymptotic zero elongation function").*

The corresponding graphs of f_i and f_i', $i = 1, \ldots, 4$ can be seen in figure 4, below.

Figure 4: Different elongation generating functions.



(a) Different EGF's (b) Corresponding 1st derivative

Theorem 1 (Construction of elongation transformations) *Assume $f \in \Upsilon$. Then*

$$T_{\theta,f}(z) \equiv zW(z) \equiv z \exp\left(\theta \int_0^z f(u)\, du\right) \tag{9}$$

is an elongation transformation with parameter θ in the sense of Hoaglin (1983).

Proof: From equation (9), $T_{\theta,f}(z) = z + O(u^2)$ for $z \approx 0$, Moreover, $T_{\theta,f}$ is symmetric around the origin. Finally, for $z > 0$

$$T'_{\theta,f}(z) = W(z)(1 + z\theta f(z)) > 0 \quad \text{and}$$

$$T''_{\theta,f}(z) = W(z)\theta(2f(z) + zf'(z) + z\theta f(z)^2) > 0. \tag{10}$$

This follows from the assumptions on f. $\square$

Example 5.2 ($E-$, $H-$, $K-$ and $J-$transformation) *The elongation generating functions from example 5.1 correspond to the following transformations:*

1. *$E-$transformation:* $\mathrm{E}_e(z) \equiv z \exp(e \cosh(z))$.
2. *$H-$transformation of Tukey (1960):* $\mathrm{H}_h(z) = z \exp(hz^2/2)$.
3. *$J-$transformation:* $\mathrm{J}_j(z) = z \cosh(z)^j$.
4. *$K-$transformation of Haynes et al. (1997):* $\mathrm{K}^*_k(z) = z(1 + z^2)^{k/2}$.

By the end of this work we will focus on the $J-$transformation, because $H-$ and $K-$transformation have been extensively studied in the literature. Further discussion of the $E-$transformation is factored out to future research.

6 $J-$transformed symmetrical distributions: Density, quantiles, moments and kurtosis ordering

Let Z denote a standard normal distribution, for simplicity. Most of the following results can be applied to arbitrary symmetric distributions as well. From the previous section it follows that J_j is a kurtosis family in the sense of Hoaglin (1983). Let the random variable X be defined as

$$X \equiv \mu + \sigma \cdot \mathrm{J}_j(Z), \quad \mu, j \in \mathbb{R}, \, \sigma > 0. \tag{11}$$

Obviously, the properties of the distribution of X (which we simply call $J-$distribution) depend on the sign of j. For $j = 0$, X reduces to a normal distribution with mean μ and variance σ. In particular, for $j > 0$, $\mathrm{J}'_j(z) \geq 1$ and $\mathrm{J}''_j(z) > 0$ for $z > 0$. Therefore, J_j is strictly monotone increasing and convex for $z > 0$ and makes the tails of the distribution of X longer. Applying methods of variable transformations, the following theorem is easily obtained:

Theorem 2 (Density and quantiles of X) *Assume $j > 0$.*

1. *Let J_j^{-1} denote the inverse mapping of J_j. Then the probability density function $f_X(x)$ can be determined by*

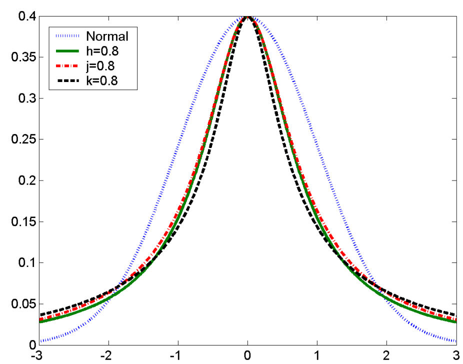
$$f_X(x;\mu,\sigma,j) = \frac{f_Z(\mathrm{J}_j^{-1}(\frac{x-\mu}{\sigma}))}{\mathrm{J}_j'(\mathrm{J}_j^{-1}(\frac{x-\mu}{\sigma}))}.$$

2. *The p-quantiles of X can be obtained from the p-quantiles of Z by means of*

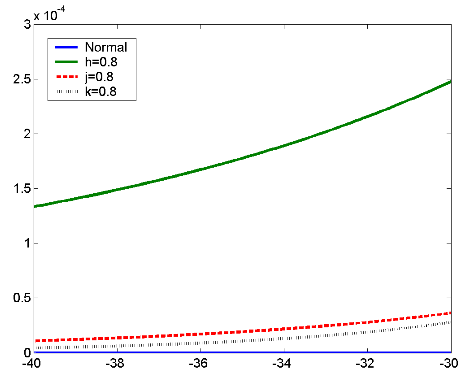
$$x_p = \mu + \sigma \cdot z_p \cosh(z_p)^j. \tag{12}$$

Different $H-$, $J-$, $K-$transformed Gaussian densities with identical parameter $h = j = k = 0.8$ are shown in figure 5. The inequalities of Lemma 2 are especially illustrated in figure 5(b).

Figure 5: Normal, $H-$, $J-$ and $K-$distributions.



(a) Transformed Gaussian densities



(b) Zoom of (a)

Note that for $j < 0$, J_j is not a one-to-one mapping. However, $\mathrm{J}_{j,1}(z) \equiv \mathrm{J}_j(z)$ for $|z| < z_p^*$ is strictly monotone increasing and $\mathrm{J}_{j,2}(z) \equiv \mathrm{J}_j(z)$ for $|z| > z_p^*$ is strictly monotone decreasing. Let $\mathrm{J}_{j,i}^{-1}$ denote the inverse function of $\mathrm{J}_{j,i}$, $i = 1, 2$. Then, the corresponding density is given by

$$f_X(x;\mu,\sigma,j) = \frac{f_Z(\mathrm{J}_{j,1}^{-1}(\frac{x-\mu}{\sigma}))}{\mathrm{J}_j'(\mathrm{J}_{j,1}^{-1}(\frac{x-\mu}{\sigma}))} + \frac{f_Z(\mathrm{J}_{j,2}^{-1}(\frac{x-\mu}{\sigma}))}{\mathrm{J}_j'(\mathrm{J}_{j,2}^{-1}(\frac{x-\mu}{\sigma}))}.$$

for $\mu + \sigma \mathrm{J}_j(z_p) < x < \mu + \sigma \mathrm{J}_j(z_p)$. The ambiguity of J_j for negative j makes the calculation of the quantiles of X slightly more complicated. Details are neglected within this work and we refer to Klein and Fischer (2002) for a similar discussion in the context of symmetrical $H-$distributions.

Theorem 3 (Existence of moments) *Let Z denote a Gaussian random variable and define $X_j \equiv \mathrm{J}_j(Z)$ for $j > 0$. Then all moments of X_j exist.*

Proof: By assumption, Z is symmetrically distributed around 0. Consequently, $X_j = J_j(Z) = Z(\frac{1}{2}e^Z + \frac{1}{2}e^{-Z})^j$ is also symmetrically distributed around 0 for $j > 0$. In particular,

$$E(X_j^k) = 2\int_0^\infty \mathrm{J}_j(z)^k f_Z(z)dz \leq E(X_{j'}^k), \quad \text{for } j < j',$$

provided that this integral exists. The last inequality can be derived from

$$j < j' \Rightarrow \mathrm{J}_j(z) \leq \mathrm{J}_{j'}(z) \quad \text{for all } z > 0.$$

It will be shown that the power moments of X_j for integer values of j exist. If j is not integer we can use the inequality

$$E(X_j^k) \leq E(X_{[j+1]}^k)$$

to prove the existence of the power moments of X_j for arbitrary $j > 0$. Let μ_i', $i = 1, 2, \ldots$ denote the power moments of Z. For $j \in \mathbb{N}$, using quadratic completion,

$$E(X_j^k) = \sum_{i=0}^{jk} \binom{jk}{i} e^{1/2(2i-jk)^2} \sum_{p=0}^{k} \binom{k}{p} \mu_i'(2i-jk)^{k-p}. \tag{13}$$

Note that because Z is standard normal,

$$\mu_i' = E(Z^i) = \begin{cases} (i-1)! & \text{for odd } i \\ 0 & \text{for even } i \end{cases}$$

for $i = 1, 2, \ldots$. All power moments of X_j exist because all sums in equation (13) are finite and all power moments of Z exist, by assumption. □

Some values of the fourth standardized moments for the $H-$, $J-$ and $K-$distribution are given in the table 1, below.

Table 1: Fourth standardized moments.

$h/j/k$	H	J	K
0	3.0000	3.0000	3.0000
0.01	3.1270	3.0593	3.0532
0.02	3.2694	3.1211	3.1079
0.05	3.8202	3.3222	3.2812
0.1	5.4417	3.7187	3.6039
0.2	11.3544	4.8265	4.3988
0.3	15.1050	6.5518	5.4438
0.4	15.6930	8.8264	6.7851
0.5	17.6393	10.8781	8.2289

Finally it will be shown that the $J-$distributions can be ordered in the sense of van Zwet (1964).

Theorem 4 (Kurtosis ordering) *Let $0 < j_1 < j_2$ and $X_j = \mathrm{J}_j(Z)$ for a symmetric random variable Z. Then, $F_{j_1} \preceq_S F_{j_2}$.*

Proof: According to condition (B) we have to show that $S_{F_{j_2}}(S^{-1}_{F_{j_1}}(x))$ is convex for $x > {F_{j_1}}^{-1}(0.5)$. Assuming that $S_{F_{j_2}}(S^{-1}_{F_{j_1}}(x))$ is twice differentiable, it is sufficient to verify that the second derivative is positive. Applying standard calculus and using $u \equiv F_{j_1}(x)$, the first derivative is given by

$$a(u) = \frac{S'_{F_{j_2}}(u)}{S'_{F_{j_1}}(u)} = \frac{[u\cosh(u)^{j_2}]'}{[u\cosh(u)^{j_1}]'} = \frac{\cosh(u)^{j_2}\,(1 + uj_2\tanh(u))}{\cosh(u)^{j_1}\,(1 + uj_1\tanh(u))} \tag{14}$$

From equation (14), the second derivative can be derived as

$$\begin{aligned} a'(u) = &\frac{[\cosh(u)^{j_2}\,(1 + uj_2\tanh(u))]' \cdot \cosh(u)^{j_1}\,(1 + uj_1\tanh(u))}{(\cosh(u)^{j_1}\,(1 + uj_1\tanh(u)))^2} \\ &-\frac{\cosh(u)^{j_2}\,(1 + uj_2\tanh(u)) \cdot [\cosh(u)^{j_1}\,(1 + uj_1\tanh(u))]'}{(\cosh(u)^{j_1}\,(1 + uj_1\tanh(u)))^2}. \end{aligned} \tag{15}$$

With

$$\begin{aligned} &\left[\cosh(u)^{j_i}\,(1 + uj_i\tanh(u))\right]' \\ =& \left(j_i\cosh(u)^{j_i-1}\sinh(u)\right)(1 + uj_i\tanh(u)) + \cosh(u)^{j_i}\left(j_i\tanh(u) + uj_i(1 - \tanh(u)^2)\right) \\ =& j_i\cosh(u)^{j_i}\Big[\tanh(u)\,(1 + uj_i\tanh(u)) + \tanh(u) + u(1 - \tanh(u)^2)\Big] \\ =& j_i\cosh(u)^{j_i}\Big[2\tanh(u) + u\tanh(u)^2(j_i - 1) + u\Big], \quad i = 1, 2, \end{aligned}$$

and equation (15), $a'(u) \cdot (\cosh(u)^{j_1}\,(1 + uj_1\tanh(u)))^2$ is given by

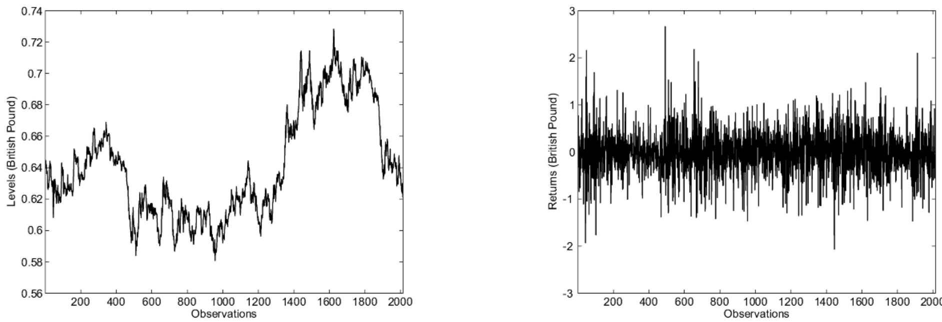
$$\begin{aligned} &j_2\cosh(u)^{j_2}\Big[2\tanh(u) + u\tanh(u)^2(j_2 - 1) + u\Big] \cdot \left(\cosh(u)^{j_1}\,(1 + uj_1\tanh(u))\right) \\ &-j_1\cosh(u)^{j_1}\Big[2\tanh(u) + u\tanh(u)^2(j_1 - 1) + u\Big] \cdot \left(\cosh(u)^{j_2}\,(1 + uj_2\tanh(u))\right) \\ =& \cosh(u)^{j_1+j_2}\Big[(j_2 - j_1)2\tanh(u) + (j_2(j_2 - 1) - j_1(j_1 - 1))u^2\tanh(u)^2 + (j_2 - j_1)u \\ &\quad +2j_2\tanh(u)uj_1\tanh(u) - 2j_1\tanh(u)uj_2\tanh(u) \\ &\quad j_2u^2j_1\tanh(u) - j_1u^2j_2\tanh(u) + (j_2 - j_1)u^2j_1j_2\tanh(u)^3\Big] \\ =& \cosh(u)^{j_1+j_2}\Big[(j_2 - j_1)2\tanh(u) + (j_2^2 - j_1^2 + (j_1 - j_2))u^2\tanh(u)^2 + (j_2 - j_1)u \\ &+(j_2 - j_1)u^2j_1j_2\tanh(u)^3\Big] \\ \geq& \cosh(u)^{j_1+j_2}\Big[(j_2 - j_1)2\tanh(u) + (j_1 - j_2)u^2\tanh(u)^2\Big] \\ =& \cosh(u)^{j_1+j_2}(j_2 - j_1)\Big[2\tanh(u) - u^2\tanh(u)^2\Big] \geq 0, \text{ for } 0 \leq u \leq 1, j_2 > j_1. \end{aligned}$$

Note that $\cosh(u) > 0$ and $\tanh(u) \geq 0$ for $0 \leq u \leq 1$. Using $\frac{du}{dx} > 0$, the result follows immediately. □

7 Application to financial return data

In order to compare results concerning the fit of the transformed distributions, we focus on the series of the US dollar exchange rate for the British pound from January 1995 to December 2002 ($N = 2014$ observations) which can be obtained from the PACIFIC (Policy Analysis Computing & Information Facility in Commerce) Exchange Rate Service of the University of British Columbia.[1] The series of prices and corresponding log-returns are given in figure 6.

Figure 6: Prices and log-returns of the British pound from 02-01-1995 to 31-12-2002.



(a) Prices (1 US$ in pounds) (b) Log-returns

The (sample) mean of the log-returns is -0.0014 with a (sample) standard deviation of 0.4779. Moreover, there seems to be no remarkable skewness in the data set (the skewness coefficient – measured by the third standardized moments – is given by by 0.0929), whereas the kurtosis coefficient – in terms of the fourth standardized moments – is 4.8122, reflecting the remarkable leptokurtosis of the data. This is the reason why we apply the elongation transformation to different symmetric distributions (i.e. Gaussian, logistic and Student distribution with 7 degrees of freedom) only.

Applying the Lagrange multiplier test of Engle (1982) to the data we come across the presence of ARCH-effects. To overcome this problem, we ”pre-whiten” the log-returns by fitting a GARCH(1,1) model and considering the GARCH residuals in addition to the log-returns. The mean of the residuals is -0.0062, the standard deviation is given by 1.0004. Moreover skewness and kurtosis coefficient are 0.0891 and 4.9661, respectively.

Four criteria have been employed to compare the goodness-of-fit of the different candidate distributions. The first is the *log-Likelihood value* ($\mathcal{LL}$) obtained from the Maximum-Likelihood estimation. The $\mathcal{LL}$-value can be considered as an ”overall measure of goodness-of-fit and allows us to judge which candidate is more likely to have generated the data”. As distributions with different numbers of parameters N_k are used, this is taken into account by calculating the *Akaike criterion* given by

$$AIC = -2 \cdot \mathcal{LL} + \frac{2N(N_k + 1)}{N - N_k - 2}.$$

The third criterion is the *Kolmogorov-Smirnov distance* as a measure of the distance between the estimated parametric cumulative distribution function, $\hat{F}$, and the empirical

[1] Download under `http://www.pacific.commerce.ubc.ca/`.

sample distribution, F_{emp}. It is usually defined by

$$\mathcal{K} = 100 \cdot \sup_{x\in\mathbb{R}} |F_{emp}(x) - \hat{F}(x)|. \tag{16}$$

Finally, *Anderson-Darling statistic* is calculated, which weights $|F_{emp}(x) - \hat{F}(x)|$ by the reciprocal of the standard deviation of F_{emp}, namely $\sqrt{\hat{F}(x)(1-\hat{F}(x))}$, that is

$$\mathcal{AD}_0 = \sup_{x\in\mathbb{R}} \frac{|F_{emp}(x) - \hat{F}(x)|}{\sqrt{\hat{F}(x)(1-\hat{F}(x))}}. \tag{17}$$

Instead of just the maximum discrepancy, the second and third largest value, which is commonly termed as $\mathcal{AD}_1$ and $\mathcal{AD}_2$, are also taken into consideration. Whereas $\mathcal{K}$ emphasizes deviations around the median of the fitted distribution, $\mathcal{AD}_0, \mathcal{AD}_1$ and $\mathcal{AD}_2$ allow discrepancies in the tails of the distribution to be appropriately weighted. The *results of the Maximum likelihood estimation* are summarized in table 2 and 3, below. Note that μ and δ denote the location and scale parameter, respectively.

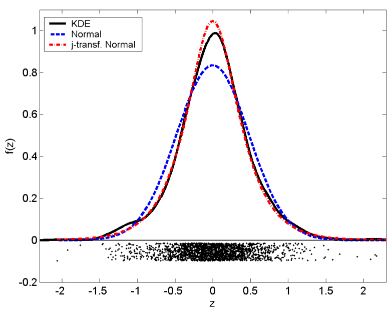
Table 2: Goodness-of-fit and estimated parameters: Log-returns

Type	$\mathcal{LL}$	$\mathcal{AIC}$	$\mathcal{K}$	$\mathcal{AD}_0$	$\mathcal{AD}_1$	$\mathcal{AD}_2$	$\widehat{\mu}$	$\widehat{\delta}$	$\widehat{h}/\widehat{j}/\widehat{k}$
Transformed Gaussian									
No	-1369.5	2745.0	4.153	4.480	0.838	0.833	-0.0014	0.4778	0.0000
h	-1310.8	2629.7	1.134	0.058	0.056	0.056	-0.0007	0.3947	0.1183
k	-1308.5	2625.1	1.068	0.089	0.088	0.074	-0.0002	0.3645	0.2090
j	-1307.9	2623.7	0.910	0.051	0.051	0.051	-0.0004	0.3806	0.2152
Transformed logistic									
No	-1314.1	2634.2	1.800	0.098	0.083	0.081	-0.0009	0.2579	0.0000
h	-1310.9	2629.9	1.244	0.063	0.062	0.061	-0.0008	0.4528	0.0252
k	-1308.2	2624.4	0.873	0.053	0.051	0.049	-0.0004	0.4345	0.0702
j	-1309.3	2626.7	1.029	0.055	0.054	0.053	-0.0006	0.4440	0.0627
Transformed Student-t with 7 degrees of freedom									
No	-1314.5	2636.9	1.802	0.081	0.080	0.077	-0.0010	0.3977	0.0000
h	-1313.2	2634.4	1.418	0.067	0.066	0.065	-0.0008	0.3292	0.0096
k	-1309.4	2626.8	0.910	0.056	0.052	0.050	-0.0004	0.3112	0.0550
j	-1311.0	2630.1	1.084	0.056	0.056	0.055	-0.0007	0.3204	0.0390

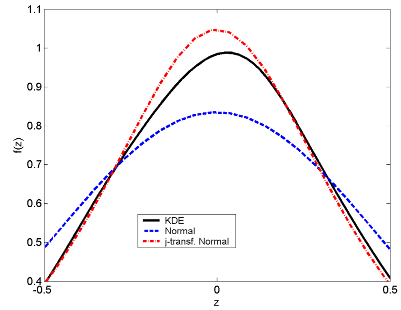
As expected, application of elongation transformations to different symmetric distributions leads to a significant improvement of all goodness-of-fit measures: The less the kurtosis of the original distribution the better the improvement. Transforming the Student-t(7) distribution only slightly improves the goodness-of-fit. Moreover, transformed Gaussian distributions provide a better fit than transformed logistic or transformed Student-t(7) distributions do. Consequently, for our data set we recommend applying the transforms to the Gaussian distribution (or distributions with similar kurtosis) only, at least for leptokurtic data. Within that class the $J-$transformation outperforms both the $K-$transformation and the $H-$transformation (concerning both the global fit and the fit of the tails).

The fit of the $J-$transformed Gaussian distribution is illustrated in figure 7, below.

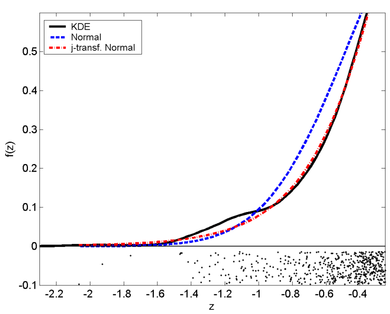
Figure 7: Kernel density estimation versus parametric fit.



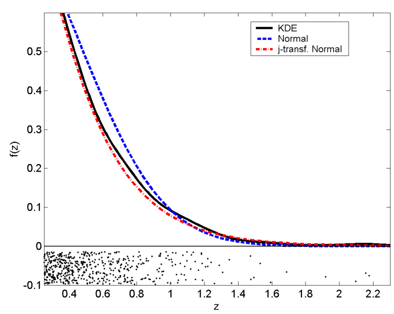
(a) $J-$transformed normal



(b) Zoom of figure(a): Centre



(c) Zoom of figure(a): Left tail



(d) Zoom of figure(a): Right tail

Note that the results are very similar for the GARCH(1,1)-residuals (see table 3, below). Again, combining the $J-$transformation with the normal distribution seems to be very promising. For a more detailed discussion concerning the goodness-of-fit in the context of financial return data we refer to Fischer et al. (2003).

Table 3: Goodness-of-fit and estimated parameters: GARCH(1,1)-residuals

Type	$\mathcal{LL}$	$\mathcal{AIC}$	$\mathcal{K}$	$\mathcal{AD}_0$	$\mathcal{AD}_1$	$\mathcal{AD}_2$	$\widehat{\mu}$	$\widehat{\delta}$	$\widehat{h}/\widehat{j}/\widehat{k}$
Transformed Gaussian									
No	-2855.7	5717.4	3.778	30.61	2.175	0.245	-0.0062	1.0004	0.0000
h	-2800.1	5608.1	0.891	0.033	0.032	0.031	-0.0034	0.8383	0.1083
k	-2801.2	5610.4	1.092	0.226	0.113	0.034	-0.0008	0.7793	0.1932
j	-2798.8	5605.7	0.884	0.079	0.057	0.027	-0.0024	0.8106	0.1989
Transformed logistic									
No	-2802.4	5610.9	1.576	0.045	0.045	0.045	-0.0041	0.5415	0.0000
h	-2799.9	5607.9	0.954	0.036	0.033	0.032	-0.0036	0.9564	0.0206
k	-2798.9	5605.8	0.836	0.070	0.050	0.029	-0.0026	0.9287	0.0537
j	-2799.2	5606.4	0.770	0.050	0.039	0.030	-0.0031	0.9427	0.0494
Transformed Student-t with 7 degrees of freedom									
No	-2802.1	5612.3	1.560	0.047	0.047	0.045	-0.0045	0.6974	0.0000
h	-2801.4	5610.8	1.126	0.038	0.037	0.037	-0.0041	0.6957	0.0067
k	-2799.5	5607.1	0.827	0.033	0.031	0.030	-0.0029	0.6683	0.0392
j	-2800.3	5608.6	0.876	0.034	0.033	0.032	-0.0035	0.6823	0.0277

8 Summary

Within this work we have proposed an alternative elongation transformation — the so-called $J-$transformation — and derived some basic properties of this transformation. By means of elongation generating functions we have shown that the $J-$transformation generates less elongation than the $H-$transformation but more elongation than the $K-$transformation. In particular, we have proved that all moments of $J-$transformed Gaussian distributions exist and that the parameter $j > 0$ of the $J-$transformation is a kurtosis parameter in the sense of van Zwet (1964). Finally, by means of forex data, we empirically investigated the influence of the elongation transformation on different symmetric distributions and demonstrated the excellent fit of $J-$transformed Gaussian distributions.

References

[1] Ali, M. M.: *Stochastic Ordering and Kurtosis Measure*. Journal of the American Statistical Association, **69**:543-545, 1974.

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

[3] Barndorff-Nielsen, O. E.: *Exponentially Decreasing Distributions for the Logarithm of Particle Size*. Proceedings of the Royal Statistical Society Series A, **353**:401-419, 1977.

[4] Darlington, R. B.: *Is Kurtosis really Peakedness?* American Statistician, **24**(2):19-20, 1970.

[5] Engle, R. F. *Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of the United Kingdom Inflation*. Econometrica, **50**(4):987-1007, 1982.

[6] Fama, E.: *The Behaviour of Stock Prices*. Journal of Business, **38**:34-105, 1965.

[7] Fischer, M., Horn, A. and I. Klein: *Tukey-type Distributions in the Context of Financial Return Data*. Diskussion paper 53, University of Erlangen-Nürnberg, 2003.

[8] Groeneveld, R. A.: *A Class of Quantile Measures for Kurtosis*. American Statistician, **52**:325-329, 1998.

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

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

[11] Kaplanski, I.: *A Common Error concerning Kurtosis*. Journal of the American Statistical Association, **40**:259, 1945.

[12] Klein, I. and M. Fischer: *Families of gh-Transformed Distributions*, in: S. Mittnik and I. Klein (ed.): Contribution to Modern Econometrics, Kluwer, 2002.

[13] MacGillivray, H. L.: *The Mean, Median, Mode inequality and Skewness for a Class of Densities*. Australian Journal of Statistics, **23**(2):247-250, 1981.

[14] MacGillivray, H. L. and K. P. Belanda: *The Relationship between Skewness and Kurtosis*. Australian Journal of Statistics, **30**(3):319-337, 1988.

[15] Martinez, J. and B. Iglewicz: *Some Properties of the Tukey g and h Family of Distributions*. Communication in Statistics (Theory and Methods), **13**:353-369, 1984.

[16] Oja, H.: *On Location, Scale, Skewness and Kurtosis of Univariate Distributions*. Scandinavian Journal of Statistics, **8**:154-168, 1981.

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

[18] Tukey, J. W.: *Exploratory Data Analysis*. Addison-Wesley, Reading, M.A., 1977.

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

Adresse der Autoren:

Dr. Matthias Fischer
Lehrstuhl für Statistik und Ökonometrie
Universität Erlangen-Nürnberg
Lange Gasse 20
D-90403 Nürnberg

Tel. +60 911 5320271
Fax +60 911 5320277
Elec. Mail: Matthias.Fischer@wiso.uni-erlangen.de
http://www.statistik.wiso.uni-erlangen.de

Prof. Dr. Ingo Klein
Lehrstuhl für Statistik und Ökonometrie
Universität Erlangen-Nürnberg
Lange Gasse 20
D-90403 Nürnberg

Tel. +60 911 5320290
Fax +60 911 5320277
Elec. Mail: Ingo.Klein@wiso.uni-erlangen.de
http://www.statistik.wiso.uni-erlangen.de