On Blest's measure of kurtosis adjusted for skewness

J.F. Rosco[a], Arthur Pewsey[a,*], M.C. Jones[b]

[a] *Department of Mathematics, Escuela Politécnica, University of Extremadura, 10003 Cáceres, Spain*
[b] *Department of Mathematics & Statistics, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK*

Abstract

We reconsider the derivation of Blest's (2003) skewness adjusted version of the classical moment-based coefficient of kurtosis and propose an adaptation of it which generally eliminates the effects of asymmetry more successfully. Lower bounds are provided for the two skewness adjusted kurtosis measures as functions of the classical coefficient of skewness. The results from a Monte Carlo experiment designed to investigate the sampling properties of numerous moment-based estimators of the two skewness adjusted kurtosis measures are used to identify those estimators with lowest mean squared error for small to medium sized samples drawn from distributions with varying levels of asymmetry and tailweight.

Keywords: Asymmetry, Estimation, Lower bounds, Moment-based measures, Sinh-arcsinh transformation

1. Introduction

The classical fourth moment-based coefficient $\alpha_4 = \mu_4/\sigma^4$, where $\mu_k = E[(X-\mu)^k]$, $\sigma^2 = \mu_2$, $\mu = E(X)$ and X denotes a random variable (Thiele, 1889; Pearson, 1905), remains the best known and most widely applied measure of kurtosis. This is in spite of the fact that the coefficient does not exist if the fourth moment does not exist, a major limitation on its use with heavy-tailed distributions. Moreover, even for symmetric distributions, its interpretation can be far from obvious. For asymmetric distributions, it has long been known (Pearson, 1916) that $\alpha_4 \geq \alpha_3^2 + 1$, where $\alpha_3 = \mu_3/\sigma^3$ is the classical moment-based coefficient of skewness. Thus, higher skewness (as measured by α_3) is inevitably accompanied by higher kurtosis (as measured by α_4). These unappealing features of α_4 have stimulated considerable debate within the literature regarding exactly what 'kurtosis' is, what it measures (or should measure), and how best to measure it. An excellent review of the extensive related literature is provided by Balanda and MacGillivray (1988); measures of kurtosis for use with asymmetric distributions are considered by Balanda and MacGillivray

*Corresponding author
Email addresses: juanfranrn@unex.es (J.F. Rosco), apewsey@unex.es (Arthur Pewsey), m.c.jones@open.ac.uk (M.C. Jones)

(1990). Many of the alternative kurtosis measures that have been proposed are based on quantiles, which exist and are unique if the distribution function is continuous and strictly monotone.

Flying somewhat in the face of these developments, and ignoring its potential non-existence, Blest (2003) advocated a skewness adjusted version of α_4. His proposal arose from consideration of what he termed the *meson*; that central value, ξ, about which the fourth moment of a distribution is minimum. Clearly, ξ is also that point about which the third moment is zero. Setting $\xi = \mu + k\sigma$,

$$k = \left(\sqrt{1+\tfrac{1}{4}\alpha_3^2} + \tfrac{1}{2}\alpha_3\right)^{1/3} - \left(\sqrt{1+\tfrac{1}{4}\alpha_3^2} - \tfrac{1}{2}\alpha_3\right)^{1/3}, \qquad (1)$$

and thus

$$\mu_4 = \mu_4^* + 6\sigma^4 k^2 + 3\sigma^4 k^4,$$

where $\mu_4^* = E[(X-\xi)^4]$ denotes the minimum fourth moment. Given this relation, Blest proposed

$$\alpha_4^* = \mu_4^*/\sigma^4 = \alpha_4 - 3k^2(2+k^2), \qquad (2)$$

as a measure of kurtosis adjusted for skewness, his clear intention being to try to eliminate the effects of skewness on α_4 noted earlier. Jones et al. (2011) note that k can be represented in terms of the sinh-arcsinh function as

$$k = 2\sinh(\tfrac{1}{3}\sinh^{-1}(\tfrac{1}{2}\alpha_3)) = 2S_{0,\frac{1}{3}}(\tfrac{1}{2}\alpha_3),$$

using the notation $S_{\varepsilon,\delta}(x) = \sinh(\delta\sinh^{-1}(x) - \varepsilon)$ of Jones and Pewsey (2009).

In Section 2, we reconsider the definition of α_4^* and propose an adaptation of it, $\alpha_4^\dagger$. In the same section, we show that neither α_4^* nor $\alpha_4^\dagger$ are kurtosis measures that are completely unaffected by skewness. We also provide lower bounds for the two skewness adjusted kurtosis measures. In Section 3 we consider the problem of how α_4^* and $\alpha_4^\dagger$ might be estimated, and present results of an extensive simulation study designed to explore the performance of various estimators based on popular estimators of the skewness measure α_3 and the kurtosis measure α_4. The paper ends with Section 4 where concluding remarks are drawn.

2. An alternative measure: comparative performance and bounds

2.1. An alternative skewness adjusted measure

It is easy to show that $\mu_2^* = E[(X-\xi)^2] = \sigma^2(1+k^2)$. This result raises the question as to why, in the definition of α_4^* in Equation (2), μ_4^* is divided by σ^4 and not $\sigma^4(1+k^2)^2$. We therefore propose the alternative skewness adjusted coefficient of kurtosis

$$\alpha_4^\dagger = \frac{\mu_4^*}{(\mu_2^*)^2} = \frac{\alpha_4^*}{(1+k^2)^2} = \frac{\alpha_4}{(1+k^2)^2} - \frac{3k^2(2+k^2)}{(1+k^2)^2}. \qquad (3)$$

Like α_4 and α_4^*, $\alpha_4^\dagger$ does not exist if the fourth moment of X does not exist. As is the case for α_4^*, the new measure $\alpha_4^\dagger$ is a function of α_4 and a sinh-arcsinh transformation of the coefficient of skewness, α_3.

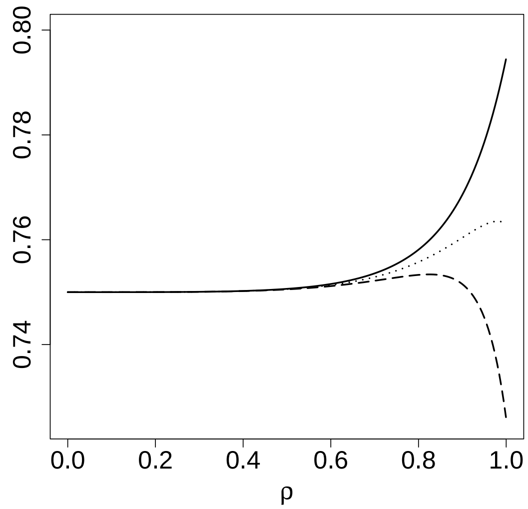

Figure 1: Kurtosis measures $\alpha_4/(1+\alpha_4)$ (solid), $\alpha_4^*/(1+\alpha_4^*)$ (dotted) and $\alpha_4^\dagger/(1+\alpha_4^\dagger)$ (dashed), as functions of $\rho = \alpha/(1+\alpha), \alpha > 0$, for the skew-normal distribution with density (4).

2.2. Performance of the skewness adjusted kurtosis measures

Although α_4^* and $\alpha_4^\dagger$ are generally less affected by skewness than α_4 is, they are not, however, skewness invariant measures. This fact is illustrated in Figures 1 and 2. Figure 1 represents all three measures for the popular skew-normal class of distributions of Azzalini (1985) with density

$$f_\alpha(x) = 2\phi(x)\Phi(\alpha x), \qquad -\infty < x, \alpha < \infty, \tag{4}$$

where ϕ and Φ are the density and distribution function, respectively, of the standard normal distribution. The parameter α is a shape parameter which affects both the skewness and kurtosis. The skew-normal distribution has shapes ranging from that of the normal distribution ($\alpha = 0$) to those of half-normal distributions ($\alpha = \pm\infty$). In Figure 1, both the measures and the shape parameter, constrained without loss of generality to be positive, have been transformed to put them on to $(0, 1)$. (When $\rho = \alpha/(1 + \alpha) = 0$, each kurtosis measure is $3/(1 + 3) = 0.75$, the kurtosis value of the normal distribution.) If the effects of asymmetry were eliminated completely for all members of the class, we would expect to see lines that were parallel with the horizontal axis in such a plot. Clearly they are not, but $\alpha_4^\dagger$ appears to do a better job than α_4^* at removing the effects of skewness for all but the most asymmetric of cases, in the neighbourhood of the half-normal ($\alpha = \infty,\ \rho = 1$) distribution.

Panels (a)–(c) of Figure 2 present contour plots of $\alpha_4/(1+\alpha_4)$, $\alpha_4^*/(1+\alpha_4^*)$ and $\alpha_4^\dagger/(1+\alpha_4^\dagger)$, as functions of $\rho_1 = \varepsilon/(1 + \varepsilon)$, $\varepsilon \geq 0$, and $\lambda_1 = \delta/(1 + \delta)$, for the sinh-arcsinhed normal (or SAS-normal, for short) family of distributions of Jones and Pewsey (2009) with density

$$f_{\varepsilon,\delta}(x) = \{2\pi(1 + x^2)\}^{-1/2}\delta C_{\varepsilon,\delta}(x) \exp\{-\tfrac{1}{2}S^2_{\varepsilon,\delta}(x)\}, \qquad -\infty < x, \varepsilon < \infty,\ \delta > 0, \tag{5}$$

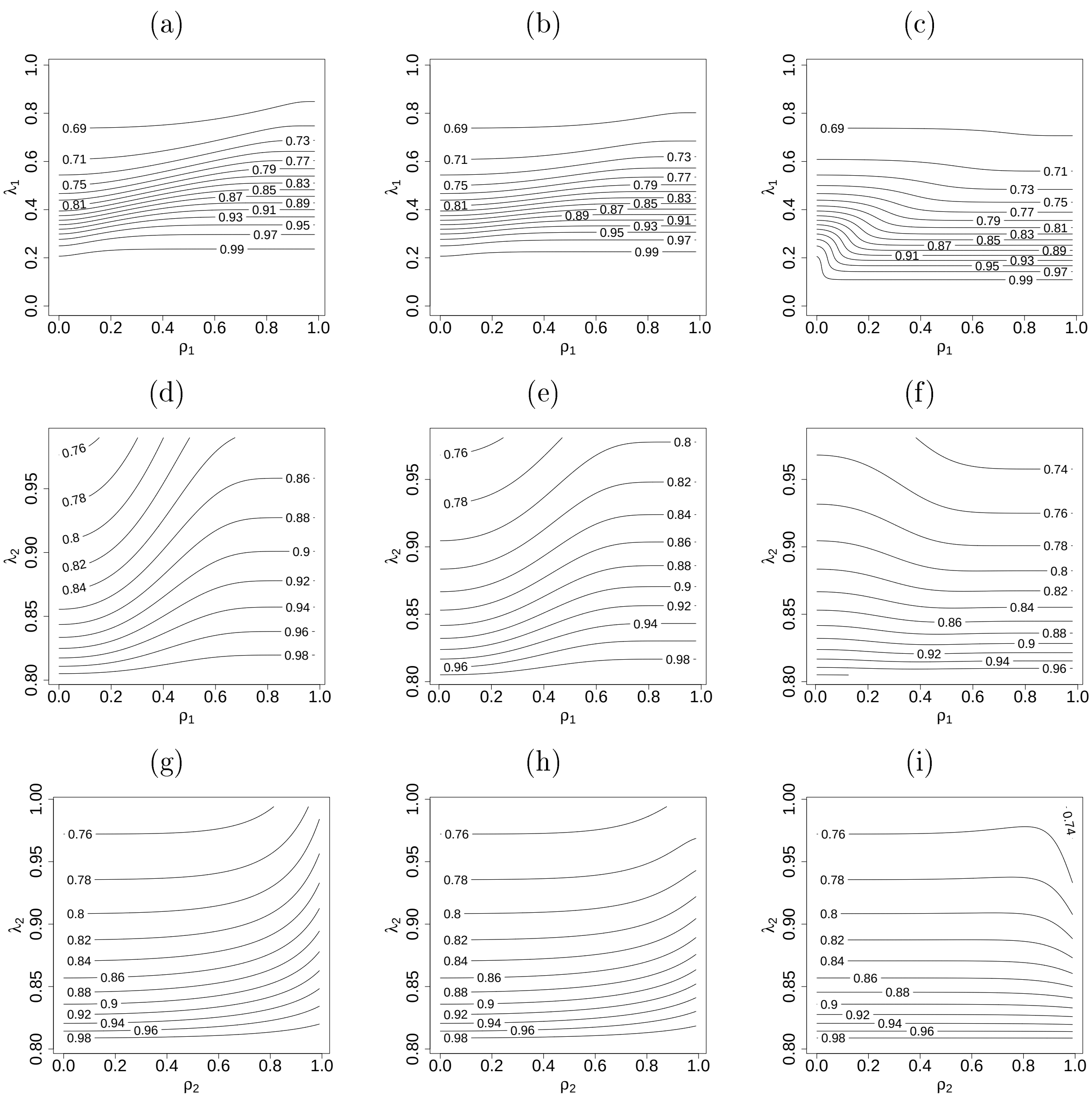

Figure 2: Contour plots of the three kurtosis measures $\alpha_4/(1+\alpha_4)$ (first column), $\alpha_4^*/(1+\alpha_4^*)$ (second column) and $\alpha_4^\dagger/(1+\alpha_4^\dagger)$ (third column). Panels (a)–(c) correspond to the SAS-normal distribution with density (5), as functions of $\rho_1 = \varepsilon/(1+\varepsilon), \varepsilon \geq 0$, and $\lambda_1 = \delta/(1+\delta)$. Panels (d)–(f) are their analogues for the SAS-t distribution with density (6), as functions of $\rho_1 \geq 0$ and $\lambda_2 = \nu/(1+\nu)$. Panels (g)–(i) correspond to the skew-t distribution with density (7), as functions of $\rho_2 = \alpha/(1+\alpha), \alpha \geq 0$, and λ_2.

where $C_{\varepsilon,\delta}(x) = \cosh\{\delta \sinh^{-1}(x) - \varepsilon\} = \{1 + S^2_{\varepsilon,\delta}(x)\}^{1/2}$. Here, δ is a tailweight parameter, while ε regulates the skewness of the distribution. The SAS-normal distribution has tails ranging from the extremely heavy ($\delta \simeq 0$), through those of the normal distribution ($\delta = 1$) to the extremely light ($\delta \to \infty$). Its densities are symmetric if $\varepsilon = 0$, and increasingly positively (negatively) skewed as $\varepsilon \to \infty$ ($\varepsilon \to -\infty$). In the contour plots of panels (a)–(c), we would expect to see contour lines that were parallel with the horizontal axis if the effects of asymmetry were eliminated completely. Here it is debatable which of the two forms of correction does best at removing the effects of asymmetry, effects that are not especially strong to start with in this case. Certainly for moderate levels of asymmetry and perhaps for high levels of asymmetry, $\alpha_4^\dagger$ performs best. However, for distributions with heavy tails ($\delta < 1, \lambda_1 < 1/2$) and low levels of asymmetry ($\varepsilon \simeq 0, \rho_1 \simeq 0$), α_4^* performs better.

Panels (d)–(f) of Figure 2 portray contour plots analogous to those in panels (a)–(c), now as functions of $\rho_1 = \varepsilon/(1+\varepsilon)$, $\varepsilon \geq 0$, and $\lambda_2 = \nu/(1+\nu)$, for the sinh-arcsinhed t distribution of Rosco et al. (2011) with density

$$f_{\varepsilon,\nu}(x) = \frac{K_\nu C_{\varepsilon,1}(x)}{\sqrt{1+x^2}(1+\nu^{-1}S^2_{\varepsilon,1}(x))^{(\nu+1)/2}}, \qquad -\infty < x, \varepsilon < \infty, \quad \nu > 0, \tag{6}$$

where $K_\nu = \Gamma((\nu+1)/2)/(\sqrt{\nu\pi}\Gamma(\nu/2))$. As for the SAS-normal distribution, ε is the skewness regulating parameter. However, ν replaces δ as the tailweight parameter. The SAS-t distribution has tails ranging from the extremely heavy ($\nu \simeq 0$), through those of the Cauchy distribution ($\nu = 1$), all the way to those of the normal distribution ($\nu \to \infty$). However, the moment-based kurtosis measures are only defined for $\nu > 4$, or $\lambda_2 > 0.8$. For this family of distributions, $\alpha_4^\dagger$ can probably be judged to generally perform best.

Finally, panels (g)–(i) of Figure 2 provide analogous contour plots, now as functions of $\rho_2 = \alpha/(1+\alpha)$, $\alpha \geq 0$, and $\lambda_2 = \nu/(1+\nu)$, for the skew-t distribution of Azzalini and Capitanio (2003) with density

$$f_{\alpha,\nu}(x) = 2t_\nu(x)T_{\nu+1}\left\{\alpha x \left(\frac{\nu+1}{x^2+\nu}\right)^{1/2}\right\}, \qquad -\infty < x, \alpha < \infty, \quad \nu > 0, \tag{7}$$

where t_ν and T_ν denote the density and distribution function, respectively, of the t-distribution with ν degrees of freedom. Here, α is a skewness parameter (as for the skew-normal class) and ν is a tailweight parameter (as for the SAS-t family). Again, the moment-based kurtosis measures are only defined for $\nu > 4$, or $\lambda_2 > 0.8$. For this family, $\alpha_4^\dagger$ generally performs best, particularly for distributions with low to moderate levels of skewness.

Thus, although the ability of the kurtosis measures α_4^* and $\alpha_4^\dagger$ to remove the influence of skewness clearly depends on the family of distributions under consideration and the level of skewness, our findings for the four flexible families of unimodal distributions considered here indicate that $\alpha_4^\dagger$ generally outperforms α_4^*, if not by a huge amount. It is noteworthy that, in the examples of Figure 2, α_4^* actually makes little difference compared with α_4; on the other hand, $\alpha_4^\dagger$ makes more difference, although sometimes it seems to adjust α_4 a little bit too much.

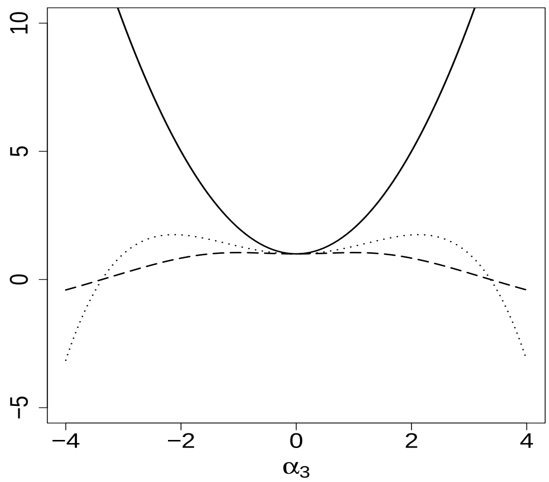

Figure 3: Lower bounds for α_4 (solid), α_4^* (dotted) and $\alpha_4^\dagger$ (dashed) as functions of the skewness measure α_3.

2.3. Lower bounds

As stated in the Introduction, the standard kurtosis measure α_4 is bounded below by $\alpha_3^2 + 1$. Here we consider lower bounds for the two skewness adjusted measures, α_4^* and $\alpha_4^\dagger$.

The key to obtaining a lower bound for α_4^* is the following simple bound for the 'symmetric' (actually, odd) sinh-arcsinh function $S_{0,\delta}(x) = \sinh(\delta \sinh^{-1}(x))$ when $0 \leq \delta \leq 1$ and $x \geq 0$: $S_{0,\delta}(x) \leq \delta x$. This follows because $S_{0,\delta}(0) = 0$, $S'_{0,\delta}(0) = \delta$ and, with just a little effort, $S_{0,\delta}(x)$ with $0 \leq \delta \leq 1$ can be shown to be concave on $x \geq 0$. It follows that $k = 2S_{0,\frac{1}{3}}(\alpha_3/2) \leq \alpha_3/3$ for $\alpha_3 \geq 0$ and hence, since k is an odd function of α_3,

$$k^2 \leq \tfrac{1}{9}\alpha_3^2.$$

(Blest (2003) notes essentially that $k \approx \alpha_3/3$ which is indeed a good approximation for small α_3.) Finally,

$$\alpha_4^* = \alpha_4 - 3k^2(2 + k^2) \geq \alpha_3^2 + 1 - \tfrac{1}{3}\alpha_3^2\left(2 + \tfrac{1}{9}\alpha_3^2\right) = 1 + \tfrac{1}{3}\alpha_3^2 - \tfrac{1}{27}\alpha_3^4.$$

The same bound divided by $(1 + \frac{1}{9}\alpha_3^2)^2$ clearly holds for $\alpha_4^\dagger = \alpha_4^*/(1 + k^2)^2$. That is,

$$\alpha_4^\dagger \geq \frac{1 + \frac{1}{3}\alpha_3^2 - \frac{1}{27}\alpha_3^4}{(1 + \frac{1}{9}\alpha_3^2)^2} = 3\left(\frac{27 + 9\alpha_3^2 - \alpha_3^4}{81 + 18\alpha_3^2 + \alpha_3^4}\right).$$

Figure 3 portrays the lower bounds for α_4, α_4^* and $\alpha^\dagger$ as functions of α_3. All three lower bounds are clearly identical, and equal to one, if the underlying distribution is symmetric ($\alpha_3 = 0$). The bounds for α_4^* and $\alpha_4^\dagger$ are not dissimilar for α_3 values within the plotted range. However, $\alpha_4^* \to -\infty$ as $|\alpha_3| \to \infty$, while $\alpha_4^\dagger \to -3$ as $|\alpha_3| \to \infty$. The lower bounds on the skewness-adjusted kurtosis measures are much less stringent than the classical lower bound on the value of α_4.

3. Estimation

When working with data, it will of course be of interest to estimate the values of α_4^* and $\alpha_4^\dagger$, and this is the problem we consider here. Specifically, we focus on estimators of them based on popular estimators of α_3 and α_4. We introduce the underlying estimators of α_3 and α_4 in Section 3.1, and present the results from a simulation study designed to explore the performance of twelve estimators of each of α_4^* and $\alpha_4^\dagger$ in Section 3.2.

3.1. Estimators of α_3 and α_4

Let $X_1, ..., X_n$ denote a random sample from some unspecified distribution, and $\bar{X} = n^{-1}\sum_{i=1}^n X_i$, $M_k = n^{-1}\sum_{i=1}^n (X_i - \bar{X})^k$, $\tilde{\alpha}_3 = M_3/M_2^{3/2}$ and $\tilde{\alpha}_4 = g_2 + 3 = M_4/M_2^2$ denote the sample mean, the kth moment about the mean, and the classical sample moment estimators of α_3 and α_4, respectively. For data from a normal distribution, $\tilde{\alpha}_3$ is unbiased for α_3, whereas $\tilde{\alpha}_4$ is only asymptotically unbiased for α_4. For data from other distributions, the two estimators are asymptotically unbiased (see, for example, Đorić et al. 2009). $\tilde{\alpha}_3$ and $\tilde{\alpha}_4$ are the estimators of α_3 and α_4 implemented in the statistical software package STATA and the `moments` package of R.

Fisher (1930) proposed

$$\tilde{\alpha}_3' = \tilde{\alpha}_3 \frac{\sqrt{n(n-1)}}{n-2} \quad \text{and} \quad G_2 = \frac{n-1}{(n-2)(n-3)}\{(n+1)(\tilde{\alpha}_4 - 3) + 6\}$$

as estimators of α_3 and $\alpha_4 - 3$. We will denote the corresponding estimator of α_4 by $\tilde{\alpha}_4' = G_2 + 3$. For samples drawn from the normal distribution, $\tilde{\alpha}_3'$ and $\tilde{\alpha}_4'$ are unbiased. These are the estimators of α_3 and α_4 implemented within the packages SAS, SPSS and STATISTICA.

Making use of the unbiased estimators $M_2' = nM_2/(n-1)$, $M_3' = n^2M_3/\{(n-1)(n-2)\}$ and

$$M_4' = \frac{n(n^2-2n+3)}{(n-1)(n-2)(n-3)}M_4 - \frac{3n(2n-3)}{(n-1)(n-2)(n-3)}M_2^2,$$

of their population central moment counterparts, Cramér (1946) considered the estimators

$$\frac{M_3'}{(M_2')^{3/2}} = \tilde{\alpha}_3' \quad \text{and} \quad \tilde{\alpha}_4'' = \frac{M_4'}{(M_2')^2}. \tag{8}$$

As Đorić et al. (2009) explain, $\tilde{\alpha}_4''$ is biased with the same bias as $\tilde{\alpha}_4$ when the data are normal. More generally, $\tilde{\alpha}_3'$ and $\tilde{\alpha}_4''$ are biased but with smaller biases than $\tilde{\alpha}_3$ and $\tilde{\alpha}_4$.

The estimators of α_3 and $\alpha_4 - 3$ implemented in MINITAB, BMDP and the `timeDate` package of R are

$$\tilde{\alpha}_3'' = \frac{M_3}{(M_2')^{3/2}} = \tilde{\alpha}_3\left(\frac{n-1}{n}\right)^{3/2} \quad \text{and} \quad b_2 = \frac{M_4}{(M_2')^2} - 3 = \tilde{\alpha}_4\left(\frac{n-1}{n}\right)^2 - 3.$$

We will use $\tilde{\alpha}_4''' = b_2 + 3$ to denote the corresponding estimator of α_4. Like $\tilde{\alpha}_3'$, $\tilde{\alpha}_3''$ is a multiple of $\tilde{\alpha}_3$ and thus is also an unbiased estimator of $\alpha_3 = 0$ when the data are normal.

Joanes and Gill (1998) present results for the variances of the estimators $\tilde{\alpha}_3$, $\tilde{\alpha}'_3$ and $\tilde{\alpha}''_3$ and for the biases and variances of the estimators g_2, G_2 and b_2 for samples drawn from the normal distribution. They also summarise Monte Carlo based results for the bias and mean squared error (MSE) of the same estimators for data drawn from chi-squared distributions with varying levels of asymmetry, specifically, with 1, 10 and 50 degrees of freedom. They found all six estimators to be negatively biased for samples drawn from these positively skewed distributions, the bias decreasing with increasing sample size, n, and number of degrees of freedom. Based on their results, it can be concluded that $\tilde{\alpha}''_3$ and $\tilde{\alpha}'''_4$ have the smallest variances for samples drawn from the normal distribution, while $\tilde{\alpha}''_3$ and $\tilde{\alpha}_4$ have the smallest MSEs in the normal case. On the other hand, $\tilde{\alpha}'_3$ and $\tilde{\alpha}'_4$, for $n < 100$, and $\tilde{\alpha}_4$, for $100 \leq n \leq 200$, have the smallest MSEs for samples from a very skewed distribution like the chi-squared distribution with 1 degree of freedom.

3.2. Simulation study

There are twelve possible combinations of the three estimators $\tilde{\alpha}_3$, $\tilde{\alpha}'_3$ and $\tilde{\alpha}''_3$ of α_3 and the four estimators $\tilde{\alpha}_4$, $\tilde{\alpha}'_4$, $\tilde{\alpha}''_4$ and $\tilde{\alpha}'''_4$ of α_4 which one might contemplate substituting for α_3 and α_4 in (1)–(3) so as to obtain estimators of k, α_4^* and $\alpha_4^\dagger$. We identify these twelve combinations using the numbers: 1 for $(\tilde{\alpha}_3, \tilde{\alpha}_4)$, 2 for $(\tilde{\alpha}'_3, \tilde{\alpha}_4)$, 3 for $(\tilde{\alpha}''_3, \tilde{\alpha}_4)$, 4 for $(\tilde{\alpha}_3, \tilde{\alpha}'_4)$, 5 for $(\tilde{\alpha}'_3, \tilde{\alpha}'_4)$, 6 for $(\tilde{\alpha}''_3, \tilde{\alpha}'_4)$, 7 for $(\tilde{\alpha}_3, \tilde{\alpha}''_4)$, 8 for $(\tilde{\alpha}'_3, \tilde{\alpha}''_4)$, 9 for $(\tilde{\alpha}''_3, \tilde{\alpha}''_4)$, 10 for $(\tilde{\alpha}_3, \tilde{\alpha}'''_4)$, 11 for $(\tilde{\alpha}'_3, \tilde{\alpha}'''_4)$, 12 for $(\tilde{\alpha}''_3, \tilde{\alpha}'''_4)$. In order to study the small-sample bias and MSE properties of the twelve resulting estimators of α_4^* and of $\alpha_4^\dagger$, we carried out a simulation study.

In our study we generated samples of size $n = 10, 20, 50, 100$ and 200 from the SAS-normal distribution with density (5), the SAS-t distribution with density (6), and Azzalini and Capitanio's skew-t distribution with density (7). We chose these three models because of their unimodal flexibility. For each of the three families of distributions we considered values of their skewness parameters (ε for the first two, and α for the last) of $0, 0.5, 1$ and 10. For the two asymmetric t distributions we explored values of their tailweight parameter, ν, of $4.1, 10$ and ∞. (The $\nu = \infty$ cases correspond to the SAS-normal distribution with $\delta = 1$ and the skew-normal distribution, respectively.) And for the SAS-normal we investigated values for its tailweight parameter, δ, of $0.2, 0.5, 2, 5$ and 20. These parameter combinations correspond to ranges of α_4 of: $(2.14, 1154.60)$ for the SAS-normal; $(3, 266.18)$ for the SAS-t; $(3, 230.70)$ for the skew-t. For each distribution, sample size, asymmetry parameter value and tailweight parameter value combination we simulated $10,000$ samples, and from these samples we calculated the sample bias and MSE of each of the twelve estimators of α_4^* and each of the twelve estimators of $\alpha_4^\dagger$.

Consistent with the results quoted above from Joanes and Gill (1998) and there being relatively little difference between α_4 and α_4^*, the biases of all the estimates of α_4^* were found to be negative, the bias decreasing (in absolute value) with increasing sample size and as the tailweight tends to that of the normal distribution and, generally, as the skewness tends to 0 (i.e. to symmetry). With regard to the MSE of the twelve estimators of α_4^*, for distributions with normal or heavier tails we observed patterns which are well represented by panels (a) and (c) of Figure 4. For distributions with lighter than normal tails, patterns like those displayed in panel (e) of the same figure were obtained. As panels (a), (c) and (e) of Figure

4 illustrate, there is little or no difference between the MSEs of the twelve estimators of α_4^* for sample sizes of 100 or more.

The results obtained for the estimators of $\alpha_4^\dagger$ were very similar to those for the estimators of α_4^*, except that their biases and MSEs were generally found to be somewhat larger. The difference between their MSEs can be appreciated by comparing the panels corresponding to the estimates of α_4^* in the first column of Figure 4 with their counterparts for the estimates of $\alpha_4^\dagger$ in its second column.

For samples drawn from distributions with heavier than normal tails, for example, Figure 4(a),(b), the estimators of α_4^* and $\alpha_4^\dagger$ which generally had lowest MSEs were those based on the combinations 9 $(\tilde{\alpha}_3'', \tilde{\alpha}_4'')$, 7 $(\tilde{\alpha}_3, \tilde{\alpha}_4'')$ and 8 $(\tilde{\alpha}_3', \tilde{\alpha}_4'')$ (ordered according to increasing MSE). The estimator $\tilde{\alpha}_4''$ which appears in all three of these combinations was not considered by Joanes and Gill (1998) as a potential estimator of α_4. For samples from distributions with normal-like tails, for example, Figure 4(c),(d), the estimators with lowest MSEs generally corresponded to the combinations 3 $(\tilde{\alpha}_3'', \tilde{\alpha}_4)$ and 9 $(\tilde{\alpha}_3'', \tilde{\alpha}_4'')$. Both of these combinations involve $\tilde{\alpha}_3''$ which was found by Joanes and Gill (1998) to be the estimator of α_3 with smallest MSE for data drawn from the normal distribution. Finally, for samples drawn from distributions with lighter than normal tails, for example, Figure 4(e),(f), the estimators based on the combinations 1 $(\tilde{\alpha}_3, \tilde{\alpha}_4)$, 2 $(\tilde{\alpha}_3', \tilde{\alpha}_4)$ and 3 $(\tilde{\alpha}_3'', \tilde{\alpha}_4)$ were found generally to be those with lowest MSEs. All three of these combinations contain the raw moment estimator $\tilde{\alpha}_4$ of α_4. Here a comparison with the results reported in Joanes and Gill (1998) is impossible because they did not investigate the performance of the different estimators for data simulated from light tailed distributions. The estimators corresponding to the combinations 11 $(\tilde{\alpha}_3', \tilde{\alpha}_4''')$ and 10 $(\tilde{\alpha}_3, \tilde{\alpha}_4''')$ were found consistently to be the ones with the largest MSEs, and this is the reason why the results for them have been omitted from Figure 4. Both combinations involve the estimator $\tilde{\alpha}_4'''$ of α_4.

The lessons gleaned from our simulation study are pulled together in Section 4 below.

4. Concluding remarks

In this paper we have proposed $\alpha_4^\dagger$, an adaptation of Blest's (2003) coefficient of kurtosis adjusted for skewness, α_4^*. For four flexible unimodal models considered in Section 2.2, $\alpha_4^\dagger$ was found generally to outperform α_4^* in terms of its ability to remove the effects of asymmetry. Also, the lower bound for $\alpha_4^\dagger$ is closer to being constant than that for α_4^*.

Our Monte Carlo investigation of the MSEs of various moment-based estimators of α_4^* and $\alpha_4^\dagger$, reported in Section 3.2, identified the estimators corresponding to the combinations of any of the estimators of α_3 with $\tilde{\alpha}_4''$ as being the ones which generally performed best when working with samples drawn from distributions with heavier than normal tails. On the other hand, for samples drawn from distributions with lighter than normal tails, the estimators based on the combinations of any of the estimators of α_3 with $\tilde{\alpha}_4$ were found generally to perform best. In the intermediate case, for samples from distributions with close to normal tails, the estimators which generally performed best were those corresponding to the combinations 3 $(\tilde{\alpha}_3'', \tilde{\alpha}_4)$ and 9 $(\tilde{\alpha}_3'', \tilde{\alpha}_4'')$. It seems appropriate, therefore, to recommend use of $\tilde{\alpha}_3''$ throughout. The most appropriate estimator of α_4 depends on tailweight; $\tilde{\alpha}_4''$

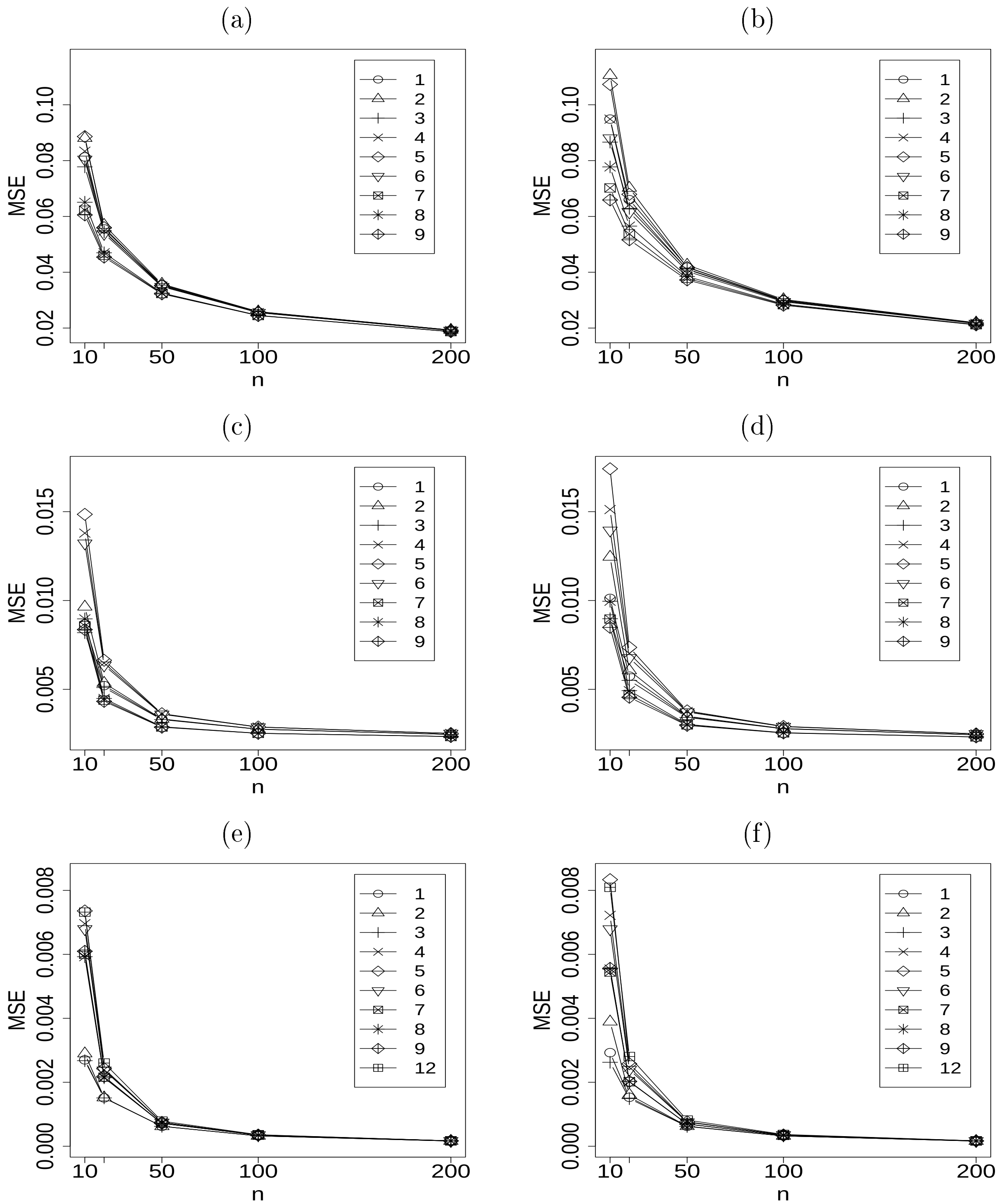

Figure 4: Empirical MSE, as a function of sample size, n, of estimators of $\alpha_4^*/(1+\alpha_4^*)$ (first column) and $\alpha_4^\dagger/(1+\alpha_4^\dagger)$ (second column) based on the combinations of the estimators of α_3 and α_4 identified in the keys. The rows correspond to data simulated from the: t distribution with $\nu = 4.1$ (first); skew-normal distribution with $\alpha = 1$ (second); SAS-normal distribution with $\delta = 20$ and $\varepsilon = 10$ (third). The results for those combinations producing the highest MSEs have been omitted so as to aid the identification of the combinations corresponding to the estimators with the lowest MSEs.

would seem to be the more usual choice, as it is good for heavier and normal tails, but users should be aware that its performance is not so good for light tails. That said, of all the different combinations considered, only combination 1 involves estimators of both α_3 and α_4 — the classical moment estimators — which are readily available within any of the major statistical packages. These conclusions all apply to estimators of both α_4^* and $\alpha_4^\dagger$, but it has to be admitted that the performance of estimators of $\alpha_4^\dagger$ is generally a little inferior to those of α_4^*.

Like α_4 and α_4^*, $\alpha_4^\dagger$ is a moment-based measure which will not exist if the fourth moment does not exist. As stated in the Introduction, the potential non-existence of moment-based kurtosis measures has rightly led researchers to propose numerous alternative measures of kurtosis. In Jones et al. (2011), we identify two wide classes of quantile-based kurtosis measures which are skewness-invariant for certain families of distributions. Development of the ideas explored there we consider to warrant future investigation.

Acknowledgements

Financial support for the research which led to the production of this paper was received by Rosco in the form of grant PRE08028 from the Junta de Extremadura, and by Pewsey in the form of grant MTM2010-16845 from the Spanish Ministry of Science and Education and grant PRI08A094 from the Junta de Extremadura.

References

Azzalini, A., 1985. A class of distributions which includes the normal ones. Scand. J. Statist. 12, 171–178.

Azzalini, A., Capitanio, A., 2003. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution. J. Roy. Statist. Soc. B 65, 367–389.

Balanda, K.P., MacGillivray, H.L., 1988. Kurtosis a critical review. Amer. Statist. 42, 111–119.

Balanda, K.P., MacGillivray, H.L., 1990. Kurtosis and spread. Canad. J. Statist. 18, 17–30.

Blest, D.C., 2003. A new measure of kurtosis adjusted for skewness. Aust. N. Z. J. Statist. 45, 175–179.

Cramér, H., 1946. Mathematical Methods of Statistics. Princeton University Press, Princeton.

Đorić, D., Nikolić-Đorić, E., Jevremović, V., Mališić, J., 2009. On measuring skewness and kurtosis. Qual. Quant. 43, 481–493.

Fisher, R.A., 1930. The moments of the distribution for normal samples of measures of departures from normality. Proc. Roy. Soc. Lond. A 130, 16–28.

Joanes, D.N., Gill, C.A., 1998. Comparing measures of sample skewness and kurtosis. Statistician 47, 183–189.

Jones, M.C., Pewsey, A., 2009. Sinh-arcsinh distributions. Biometrika 96, 761–780.

Jones, M.C., Rosco, J.F., Pewsey, A., 2011. Skewness-invariant measures of kurtosis. Amer. Statist. 65, 89–95.

Pearson, K., 1905. Skew variation, a rejoinder. Biometrika 4, 169–212.

Pearson, K., 1916. IX. Mathematical contributions to the theory of evolution.—XIX. Second supplement to a memoir on skew variation. Philos. Trans. Roy. Soc., A 216, 429–457.

Rosco, J.F., Jones, M.C., Pewsey, A., 2011. Skew-t distributions via the sinh-arcsinh transformation. Test, doi: 10.1007/s11749-010-0222-2 (in press).

Thiele, T.N., 1889. Forlæsinger over Almindelig Iagttagelseslære: Sandsynlighedsregning og Mindste Kvadraters Methode. C.A. Reitzel, Copenhagen.