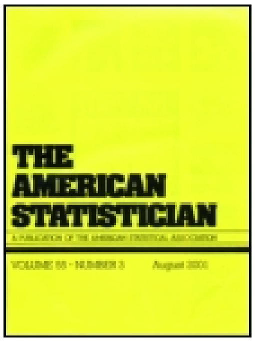

The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: https://www.tandfonline.com/loi/utas20

Kurtosis as Peakedness, 1905–2014. *R.I.P.*

Peter H. Westfall

Kurtosis as Peakedness, 1905–2014. *R.I.P.*

Peter H. WESTFALL

The incorrect notion that kurtosis somehow measures "peakedness" (flatness, pointiness, or modality) of a distribution is remarkably persistent, despite attempts by statisticians to set the record straight. This article puts the notion to rest once and for all. Kurtosis tells you virtually nothing about the shape of the peak—its only unambiguous interpretation is in terms of tail extremity, that is, either existing outliers (for the sample kurtosis) or propensity to produce outliers (for the kurtosis of a probability distribution). To clarify this point, relevant literature is reviewed, counterexample distributions are given, and it is shown that the proportion of the kurtosis that is determined by the central $\mu \pm \sigma$ range is usually quite small.

KEY WORDS: Fourth moment; Inequality; Leptokurtic; Mesokurtic; Platykurtic.

1. INTRODUCTION

By anyone's standard, a lifespan of 109 years is a good run. But it is time we put the term "peakedness," as a descriptor of kurtosis, to rest for good.

We have Karl Pearson to thank for this connection. In Pearson (1905), he defined kurtosis as

$$\text{kurtosis} = \mathrm{E}\{(\mathrm{X} - \mu)^4\}/\sigma^4 - 3$$

to measure departure from normality, and coined the terms "leptokurtic," "mesokurtic," and "platykurtic" to indicate cases where kurtosis is >0, $=0$, and <0, respectively, stating,

> "[departure from normality involves a] degree of flat-toppedness which is greater or less than that of the normal curve. Given two frequency distributions which have the same variability as measured by the standard deviation, they may be relatively more or less flat-topped than the normal curve. If more flat-topped I term them platykurtic, if less flat-topped leptokurtic, and if equally flat-topped mesokurtic." (p. 173)

Since then, numerous articles in statistics journals have appeared concerning the precise interpretation of kurtosis. While many have questioned the notion that kurtosis measures "peakedness" (specific citations given below), most state that the peak is also relevant, perhaps in deference to Pearson. And this is where the problem lies: because even articles in journals such as *The American Statistician* (*TAS*) have given a nod to "peakedness" as a descriptor of kurtosis, the incorrect interpretation of kurtosis in terms of peakedness persists.

The Internet shapes peoples' perceptions. No less than Wikipedia (*en.wikipedia.org/wiki/Kurtosis*, accessed 19 March 2014) reports "... kurtosis ... is any measure of the 'peakedness' of the probability distribution of a real-valued random variable." YourDictionary (*www.yourdictionary.com/kurtosis*, accessed 19 March 2014) reports "[kurtosis is] the degree of peakedness of the graph of a statistical distribution, indicative of the concentration around the mean ... The general form or a quantity indicative of the general form of a statistical frequency curve near the mean of the distribution." PsychWiki (*www.psychwiki.com/wiki/How_do_I_determine_whether_my_data_are_normal%3F*, accessed 3/19/2014) reports "Kurtosis involves the peakedness of the distribution. Kurtosis that is normal involves a distribution that is bell-shaped and not too peaked or flat. Positive kurtosis is indicated by a peak. Negative kurtosis is indicated by a flat distribution."

Similar gaffes are found in modern textbooks. Lee, Lee, and Lee (2013, p. 401) reported "The fourth moment around the mean—kurtosis—which characterizes peakedness, is defined by ...." Coolican (2013, p. 402) stated "... kurtosis ... refers to the overall shape of the curve in terms of its peak ...." Katz et al. (2013, p. 115) stated, "Kurtosis is characterized by a vertical stretching or flattening of the frequency distribution..... A kurtotic distribution could appear more peaked or more flattened than the normal bell-shaped distribution." Sapp (2006, p. 130) wrote "when [kurtosis is] positive the distributions are leptokurtic or peaked." Taylor (2008, p. 580) stated, "... positive and negative [kurtosis values] refer to whether the peak of the distribution is 'sharper' or higher than a normal distribution or if the peak is 'flatter' or lower ...." In McDonald (2007, p. 3), one finds a misinterpretation of both kurtosis and standard deviation: "Leptokurtic—a high-pointed, narrow-based distribution. These tend to have small standard deviations ... Kurtosis refers to 'peakedness' ...." Reinard (2006, p. 75) stated, in a section called "peakedness," "A measure of kurtosis may be computed to identify the degree to which the distribution is peaked or flat." Cohen (2008, p. 86) wrote, "... a distribution can be leptokurtic due to extreme tails or extreme peakedness ...."

Peter H. Westfall is Horn Professor in the Area of Information Systems and Quantitative Sciences, Texas Tech University, Lubbock, TX 79409 (E-mail: *peter.westfall@ttu.edu*). The author thanks the National Institutes of Health (NIH RO1 DK089167) for partial support of this work, and Frits H. Ruymgaart for helpful discussions. R code to perform the analyses in this article is available from the author.

 DOI: 10.1080/00031305.2014.917055

But the erroneous interpretations are not confined to textbooks. Even academic journals, where one might assume peer review to catch mistakes, promote the fiction. DeCarlo (1997, p. 292) stated "positive kurtosis indicates heavy tails and peakedness relative to the normal distribution, whereas negative kurtosis indicates light tails and flatness," while An and Ahmed (2008, p. 2669) stated "kurtosis describes the peakedness and tail behavior." *TAS* shares blame for promoting the confusion, publishing an article by Darlington (1970), who claimed that kurtosis is a measure of bimodality, a spectacular misdirection: while it is true that kurtosis values at or very near the minimum are indicative of bimodality, Figures 2 and 3 demonstrate that bimodal distributions are possible for all possible values of the kurtosis. A *TAS* article by Ruppert (1987) attempts to strike a balance between tail and peak definitions, concluding that kurtosis measures both peak and tails. Balanda and MacGillivray (1988, p. 111) defined kurtosis in *TAS* as "…the location- and scale-free movement of probability mass from the shoulders of a distribution into its centre and tails." The definition is vague, as the authors admit, but the real problem with this definition is that leaves the door too easily opened by those who cling to the interpretation of kurtosis as a measure of the center. A similar statement can be made about the *TAS* article by Kotz and Seiera (2008), "Visualizing Peak and Tails to Introduce Kurtosis": the title alone suggests that kurtosis concerns the peak. On the other hand, a *TAS* article by Chissom (1970, p. 22) nicely separated kurtosis from peakedness: "…It is difficult to determine the shape of a distribution from the kurtosis value alone, since almost any distribution may have a negative kurtosis value."

Correct interpretations of kurtosis as regards peakedness were also given by Ali (1974, p. 543), "…[kurtosis] measures only tailedness…" and by Johnson, Tietjen, and Beckman (1980, p. 277): "Several densities for which [Pearson's kurtosis = 0] are plotted in Figure B. Each of these distributions has zero mean, unit variance, zero skewness, and kurtosis equal to three. A considerable degree of shape differences is observed. The notion that kurtosis measures 'peakedness' is clearly not true." Unfortunately, these articles appeared in *Journal of the American Statistical Association*, a journal not widely read by nonstatisticians; thus the erroneous misinterpretations of kurtosis in terms of peakedness persist.

A possible reason for the conflation of kurtosis with peakedness is the observation that heavy-tailed distributions sometimes have higher peaks than light-tailed distributions; this makes the word "peak" seem relevant when discussing kurtosis. However, heavy-tailed distributions do not always have higher peaks: Kaplansky (1945) provided examples of leptokurtic distributions with higher and lower peak than the standard normal distribution, as well similar examples of platykurtic distributions. More extreme counterexamples are given in Figures 2 and 3, where peaks are infinite for both small and large kurtosis. In any event, the height of the peak is not relevant to my discussion of peakedness; rather, I am concerned with misinterpretations involving shape of the peak. Additionally, comparative height of peaks has little relevance for interpreting frequency histograms, particularly for datasets with different sample sizes.

To eradicate the persistent, erroneous interpretations of kurtosis in terms of peakedness, the terms "kurtosis" and "peakedness" must be disassociated. To be clear, I am not suggesting that discussions of peakedness itself (independent of kurtosis) should be eliminated; see Proschan (1965) for a thoughtful discussion. Nor am I considering variants of kurtosis such as multivariate measures (e.g., Mardia 1970) or quantile-based measures (e.g., Moors 1988). In this article, I give simple, easily understood arguments to demonstrate why the classical kurtosis measure and peakedness are unrelated. I show that both extremely large and small kurtosis values can easily be associated with distributions that have pointy, flat, and infinite bimodal peaks. I also show that the portion of the kurtosis determined by the $\mu \pm h\sigma$ range is vanishingly small, for all h, when kurtosis is large. Finally, I show that the proportion of the kurtosis determined by the $\mu \pm \sigma$ range is small, even for distributions with moderately small kurtosis.

2. TERMINOLOGY

This article concerns the expected standardized fourth moment (assumed to exist) of a random variable X, defined as $\kappa = \mathrm{E}Z^4$, where $Z = (X - \mu_X)/\sigma_X$. (Pearson used "β_2" instead of κ.) The term $\kappa - 3$ ($= \beta_2 - 3$ in Pearson's terminology) is also termed either kurtosis or "excess kurtosis," depending on the source. For the purposes of this article, κ will be called kurtosis (departing from Pearson), having the well-known properties (a) $\kappa \geq 1$ and (b) $\kappa = 3$ when X is normally distributed. Property (a) follows from Jensens' inequality: $\mathrm{E}(Z^4) \geq \{\mathrm{E}(Z^2)\}^2 = 1^2 = 1$; the minimum is attained for the equiprobable two-point distribution. Property (b) follows from direct integration.

3. HOW DATA LOOK

Perhaps the persistence of the term "peakedness" comes from people who look at histograms of heavy-tailed data and see a strongly pronounced "peak." For example, a realization of $n = 1000$ Cauchy random variables typically produces a histogram as shown in Figure 1. (Data were generated using the "rcauchy"

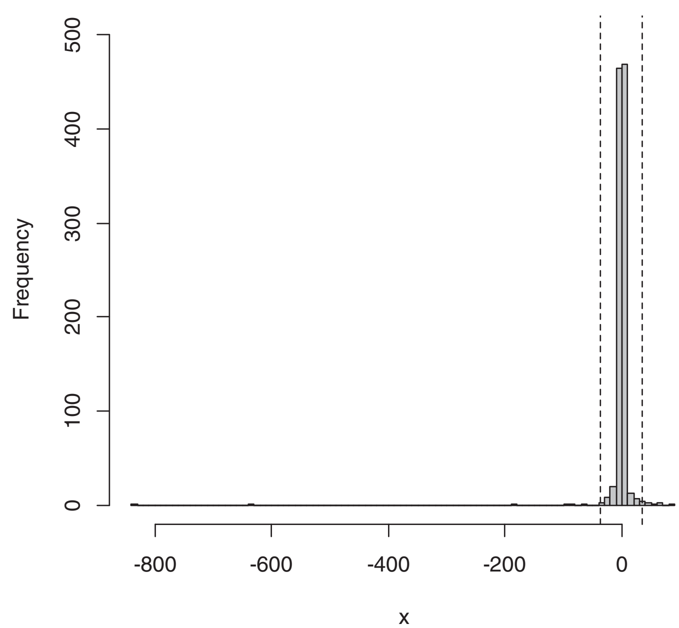

Figure 1. Histogram of a random sample of 1000 Cauchy random numbers. Dotted lines show mean ± one standard deviation of the empirical distribution.

function of R with seed 12344.) The graph seems to show a distinct narrow "peak" in the center, but this is just an artifact of the scaling of the x-axis. It is the outliers, that is, the tails, which determine this appearance.

The data graphed in Figure 1 have sample mean $m = -1.55$ and standard deviation $s = 34.85$ (using n in the denominator to correspond to the empirical distribution). Letting $z_i = (x_i - m)/s$, the kurtosis of the empirical distribution is

$$\kappa = (1/1000) \sum_i z_i^4 = 437.3871.$$

Let us separate the calculation of kurtosis by data within one standard deviation or outside one standard deviation of the mean:

$$\kappa = (1/1000) \sum_{|z_i| \le 1} z_i^4 + (1/1000) \sum_{|z_i| > 1} z_i^4$$
$$= 0.0073 + 437.3798 = 437.3871.$$

The proportion of the kurtosis statistic that is determined by the data within one standard deviation of the mean is thus $0.0073/437.3871 = 0.000017$. As this calculation shows, the notion that the kurtosis statistic has anything to do with the data near the peak is nothing short of silly with these data. On the other hand kurtosis clearly measures primarily the outliers in this example. On a related note, Livesay (2007) showed that kurtosis statistic can be used as a test for outliers, while Cramér (1946, p. 256) and Rohatgi (1976, p. 102) displayed inequalities relating kurtosis and the propensity of a distribution to produce outliers.

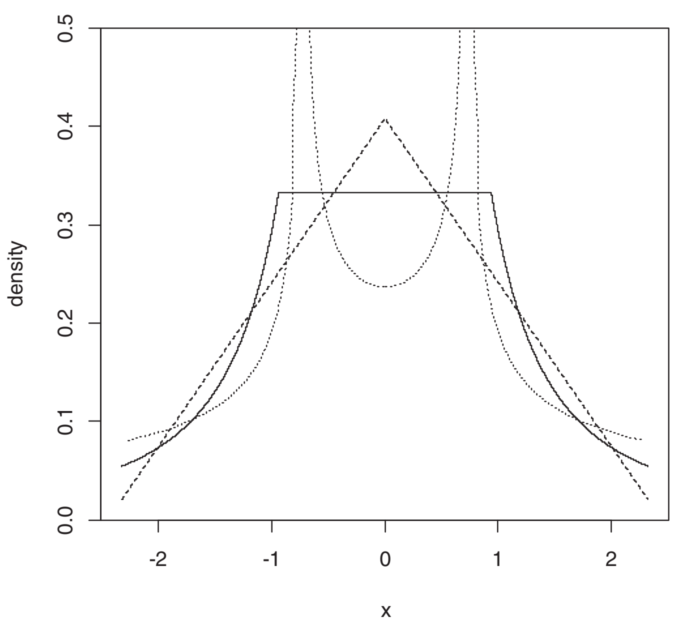

Figure 2. Distributions with identical kurtosis = 2.4: solid = devil's tower, dashed = triangular, dotted = slip-dress.

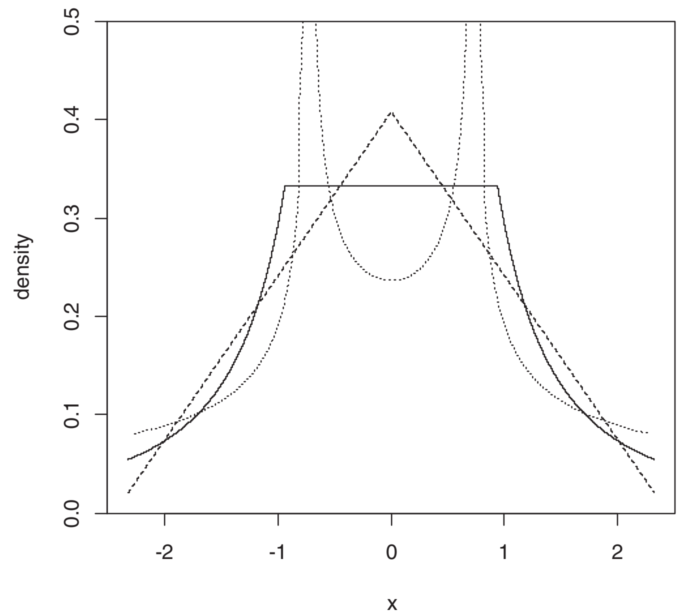

Figure 3. Central 0.99999 probability range of distributions with identical kurtosis ≅ 6000.

4. NEITHER SMALL NOR LARGE KURTOSIS CONVEY ANY USEFUL INFORMATION ABOUT THE SHAPE OF THE PEAK

Johnson, Tietjen, and Beckman (1980) provided examples of distributions with identical kurtosis as the normal, with widely different peaks. I provide six additional distributions, three with kurtosis less than normal, three with near infinite kurtosis; these examples provide even more compelling evidence for the point stated in the title of this section. For the first example, I consider symmetric distributions whose means, variances, and kurtosis values are all 0, 1, and 2.4, respectively. In Pearson's terminology, these should be "platykurtic," or flat-topped, and are graphed in Figure 2. Apparently, the peaks differ: the triangular distribution has a pointy peak, the devil's tower distribution has a flat peak, and the slip-dress distribution has no finite peak at all, but rather two infinite peaks.

Details of the distributions are given as follows.

Devil's tower: $f(x) = 0.3334$, for $|x| < 0.9399$; $f(x) = 0.2945/x^2$, for $0.9399 \le |x| < 2.3242$; and $f(x) = 0$, for $2.3242 \le |x|$.

Triangular: $f(x) = 0.4082 - 0.1667\,|x|$, for $|x| < 2.4495$; and $f(x) = 0$, for $2.4495 \le |x|$.

The slip-dress distribution is easiest to define as a mixture of beta random variables.

Slip-dress: Let $Y \sim \text{beta}(0.5, 1)$, let $m = 0.7241$ and $c = 1.5423$. Let X be either $m + cY$, $m - cY$, $-m + cY$, or $-m - cY$, each with probability 0.25; such an X has the slip-dress distribution.

So, a small kurtosis such as $\kappa = 2.4$ obviously tells you nothing about the peak, whether it is flat, pointy, or bimodal.

What about large kurtosis? Figure 3 displays three distributions with the same kurtosis $\kappa \cong 6000$; these graphs appear virtually identical to those in Figure 2. In Pearson's terminology, these should all be "leptokurtic," or less flat-topped than the normal distribution.

The distributions in Figure 3 are the ones in Figure 2 mixed with the $T(4 + 0.0000001)$ density, with mixing probability 0.0001. Again, the obvious conclusion is that kurtosis tells you nothing about the peak, whether it is flat, pointy, or bimodal. While there is more to the distributions than shown in Figure 3 (obviously the tails extend to infinity), the central 0.99999 portions that are shown cover their peaks adequately.

While the counterexample distributions shown in Figures 2 and 3 are symmetric, it would be a simple matter to introduce asymmetry without detracting from the point the kurtosis is unrelated to peakedness.

5. WHY TAILS MOSTLY DETERMINE KURTOSIS

On the face of it, the notion that κ is informative about the peak of a distribution is suspect: expected values are averages, and averages are highly influenced by extremes; see the calculations in Section 3, for example. In this section, I provide upper bounds on the proportion of kurtosis determined by the center, for general distributions, and calculate these proportions for specific distributions. The material contained in this section is an elaboration upon similar material presented by Westfall and Henning (2013, pp. 252–253).

Define $Z = (X - \mu)/\sigma$ for general random variable X, and define a central portion of the distribution F_Z of Z as $A_h = \{z\colon |z| \le h\}$. Define the quantities $\text{Center}_h = \int_{A_h} z^4 dF_Z$ and $\text{Tail}_h = \int_{A_h^c} z^4 dF_Z$. Then $\kappa = \text{Center}_h + \text{Tail}_h$. But clearly,

$$0 \le \int_{A_h} z^4 dF_Z \equiv \text{Center}_h \le h^4. \qquad (1)$$

This implies that the kurtosis is largely determined by the tail, in general: $\text{Tail}_h \le \kappa \le \text{Tail}_h + h^4$. It also follows that for large kurtosis, the portion determined by the center is vanishingly small, no matter how many standard deviations from the mean define the center: for any sequence of distributions where the kurtosis values tend to infinity, we have $\text{Center}_h/\kappa \le h^4/\kappa \to 0$, for all $h > 0$.

The Appendix shows that inequality (1) can be sharpened for $h = 1$ when the density of Z^2 is continuous and decreasing over the range [0,1]; call this class of distributions C. Defining $\text{Center} = \text{Center}_1 = \int_{A_1} z^4 dF_Z$, the Appendix proves that $\text{Center} \le 0.5$ in the case where the distribution lies in C. Inequality (1) and its extension to the class C show that kurtosis κ is determined to within ± 0.5 (within ± 0.25 for class C) by data outside the range $\mu \pm \sigma$. Specifically, defining $\text{Tail} = \text{Tail}_1 = \int_{A_1^c} z^4 dF_Z$, inequality (1) shows in general that $\text{Tail} \le \kappa \le \text{Tail} + 1$, and when the distribution lies in the class C, the range of possible kurtosis values can be sharpened to $\text{Tail} \le \kappa \le \text{Tail} + 0.5$.

These inequalities are somewhat loose. For common distributions, the proportion of kurtosis determined by the center is usually much smaller. Table 1 shows various distributions and the proportion of the kurtosis that is determined by $\text{Center} = \int_{A_1} z^4 dF_Z$. Entries are computed using simulations with 10,000,000 random numbers, generated using R with the Mersenne twister random number generator (Matsumoto and Nishimura 1998), except when calculated analytically with ease. R code to perform all simulation-based calculations is available online.

Table 1. Proportion of kurtosis determined by the range $[\mu - \sigma, \mu + \sigma]$ for various distributions

Distribution	Kurtosis	Center/Kurtosis
Two point (0.5, 0.5)	1.00	1.000
Two point$(0.5+\varepsilon, 0.5-\varepsilon)$	1.00	0.500
$0.5\text{N}(-9,1) + 0.5\text{N}(9,1)$	1.05	0.352
$0.5\text{N}(-2,1) + 0.5\text{N}(2,1)$	1.72	0.099
Uniform	1.80	0.064
Slip-dress	2.40	0.063
Devil's tower	2.40	0.055
$0.5\text{N}(-1,1) + 0.5\text{N}(1,1)$	2.50	0.048
Triangular (symmetric)	2.40	0.045
Normal	3.00	0.037
T_{10}	4.00	0.028
Logistic	4.20	0.026
Exponential	9.00	0.023
Laplace	6.00	0.015
T_5	9.00	0.012
Empirical, Figure 1	437.4	1.7×10^{-5}

Apparently, for all but the most extreme case of the two-point distribution and for distributions close to it, very little of the kurtosis is determined by the portion of the distribution that is within one standard deviation of its mean. And of course, for distributions with infinite kurtosis, the proportion that is determined by any finite central portion is zero, no matter how the "finite central portion" is defined.

6. SUMMARY

As I have shown, kurtosis tells you very little about the peak or center of a distribution. Thus, kurtosis should never be defined in terms of peakedness. To do so is counterproductive to the aim of fostering statistical literacy. The relationship of peakedness with kurtosis is now officially over.

APPENDIX: BOUNDS ON PORTION OF KURTOSIS DETERMINED BY THE CENTER

The inequality $\int_{A_1} z^4 dF_Z \le 0.5$, which states that the portion of the kurtosis that is determined the range within one standard deviation of the mean is less than 0.5, is true in general when the distribution of Z^2 satisfies the monotonicity condition given in the following theorem.

Theorem A.1. Let $Z \sim p_Z(z)$, and let $W = Z^2$. Suppose that $W \sim p_W(w)$, where $p_W(w)$ is continuous and decreasing on the interval [0,1]. Then $\text{E}\{Z^4\, \text{I}(|Z| < 1)\} \le 0.5$.

Proof. Let $Y = W^2$. Then $p_Y(y) = p_W(y^{1/2})/(2y^{1/2})$ and $\text{E}\{W^2\, \text{I}(W < 1)\} = \int_{[0,1]} y\, p_W(y^{1/2})/(2y^{1/2}) dy = 0.5 \int_{[0,1]} y^{1/2}\, p_W(y^{1/2})\, dy \le 0.5 \int_{[0,1]} y^{1/2}\, p_W(y)\, dy$ since $p_W(w)$ is decreasing on the interval [0,1].

Hence,

$E\{W^2\ I(W < 1)\} \leq 0.5\ P(|Z| \leq 1)\ E\{|Z| \mid |Z| \leq 1\} \leq 0.5$. The result $E\{Z^4\ I(|Z| < 1)\} \leq 0.5$ follows. □

As a corollary, if the distribution of Z is symmetric, then the result $E\{Z^4\ I(|Z| < 1)\} \leq 0.5$ follows even when the distribution of Z is increasing away for 0 in the [0,1] range, provided that it increases slower than z.

Corollary A.1. Suppose $p_Z(z)$ is symmetric. Then result $E\{Z^4\ I(|Z| < 1)\} \leq 0.5$ follows if $p_Z(z)/z$ is decreasing for $0 < z < 1$.

Proof. Let $W = Z^2$. Then $p_W(w) = p_Z(w^{1/2})/w^{1/2}$ and the result follows. □

As suggested by the inequality $E\{W^2\ I(W < 1)\} \leq 0.5\ P(|Z| \leq 1)\ E\{|Z| \mid |Z| \leq 1\}$ shown in the proof of the Theorem, and by the results in Table 1, the "≤ 0.5" inequality result of the Theorem and Corollary is somewhat loose. To investigate the looseness, consider a class of densities for which $f(z) = c|z|^p$ when $|z| < 1$. Since $\int f(z)dz = 1$ it follows that $p > -1$, but the densities are otherwise completely unspecified for $|z| \geq 1$. Also because $\int f(z)dz = 1$, it follows that $c \leq (p+1)/2$. Then

$$E\{Z^4 I(|Z| \leq 1)\} = 2c/(p+5) \leq (p+1)/(p+5).$$

The corollary condition that $p_Z(z)/z$ is decreasing for $0 < z < 1$ implies $p < 1$ in this class, in which case $E\{Z^4\ I(|Z| \leq 1)\} < 1/3$, rather than ≤ 0.5 as shown in the Theorem and Corollary. Additionally, $E\{Z^4\ I(|Z| \leq 1)\} \to 0$ as $p \to -1$ in this class.

[Received January 2014. Revised April 2014.]

REFERENCES

Ali, M. M. (1974), "Stochastic Ordering and Kurtosis Measure," *Journal of the American Statistical Association*, 69, 543–545. [192]

An, L., and Ahmed, S. E. (2008), "Improving the Performance of Kurtosis Estimator," *Computational Statistics and Data Analysis*, 52, 2669–2681. [192]

Balanda, K. P., and MacGillivray, H. L. (1988), "Kurtosis: A Critical Review," *The American Statistician*, 42, 111–119. [192]

Chissom, B. S. (1970), "Interpretation of the Kurtosis Statistic," *The American Statistician*, 24, 19–22. [192]

Cohen, B. H. (2013). *Explaining Psychological Statistics* (4th ed.), Hoboken, NJ: Wiley. [191]

Coolican, H. (2013), *Research Methods and Statistics in Psychology* (5th ed.), New York: Routledge. [191]

Cramér, H. (1946), *Mathematical Methods of Statistics*, Princeton, NJ: Princeton University Press. [193]

Darlington, R. B. (1970), "Is Kurtosis Really 'Peakedness'?" *The American Statistician*, 24, 19–22. [192]

DeCarlo, L. T. (1997), "On the Meaning and Use of Kurtosis," *Psychological Methods*, 2, 292–307. [192]

Johnson, M. E., Tietjen, G. L., and Beckman, R. J. (1980), "A New Family of Probability Distributions With Applications to Monte Carlo Studies," *Journal of the American Statistical Association*, 75, 276–279. [192,193]

Kaplansky, I. (1945), "A Common Error Concerning Kurtosis," *Journal of the American Statistical Association*, 40, 259. [192]

Katz, D. L., Elmore, J. G., Wild, D. M. G., and Lucan, S. C. (2013), *Jekel's Epidemiology, Biostatistics and Preventive Medicine* (4th ed.), Philadelphia, PA: Elsevier. [191]

Kotz, S., and Seiera, E. (2008), "Visualizing Peak and Tails to Introduce Kurtosis," *The American Statistician*, 62, 348–354. [192]

Lee, C.-F., Lee, J. C., and Lee, A. C. (2013), *Statistics for Business and Financial Economics* (3rd ed.), New York: Springer. [191]

Livesay, J. H. (2007), "Kurtosis Provides a Good Omnibus Test for Outliers in Small Samples," *Clinical Biochemistry*, 40, 1032–1036. [193]

Mardia, K. V. (1970), "Measures of Multivariate Skewness and Kurtosis With Applications," *Biometrika*, 57, 519–530. [192]

Matsumoto, M., and Nishimura, T. (1998), "Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator," *ACM Transactions on Modeling and Computer Simulation*, 8, 3–30. [194]

McDonald, T. H. (2007), *Basic Concepts in Statistics and Epidemiology*, Oxon, UK: Radcliffe. [191]

Moors, J. J. A. (1998), "A Quantile Alternative for Kurtosis," *Journal of the Royal Statistical Society*, Series D, 37, 25–32. [192]

Pearson, K. (1905), "Das Fehlergesetz und seine Verallgemeinerungen durch Fechner und Pearson. A Rejoinder," *Biometrika*, 4, 169–212. [191]

Proschan, F. (1965), "Peakedness of Distributions of Convex Combinations," *Annals of Mathematical Statistics*, 36, 1703–1706. [192]

Reinard, J. C. (2006), *Communication Research Statistics*, Thousand Oaks, CA: Sage. [191]

Rohatgi, V. K. (1976), *An Introduction to Probability Theory and Mathematical Statistics*, New York: Wiley. [193]

Ruppert, D. (1987), "What is Kurtosis? An Influence Function Approach," *The American Statistician*, 41, 1–5. [192]

Sapp, M. (2006), *Basic Psychological Measurement, Research Designs, and Statistics Without Math*, Springfield, IL: Charles C. Thomas. [191]

Taylor, S. E. (2008), "Kurtosis," in *Encyclopedia of Epidemiology*, eds. Sarah Boslaugh and Louise-Anne McNutt, CA: Sage Publications, pp. 580–581. [191]

Westfall, P. H., and Henning, K. S. S. (2013), *Understanding Advanced Statistical Methods*, Boca Raton, FL: Chapman & Hall/CRC Texts in Statistical Science Series. [194]