The Statistician (1998)
47, *Part* 1, *pp.* 183–189

Comparing measures of sample skewness and kurtosis

D. N. Joanes† and C. A. Gill

University of Leeds, UK

[Received January 1997. Revised September 1997]

Summary. Over the years, various measures of sample skewness and kurtosis have been proposed. Comparisons are made between those measures adopted by well-known statistical computing packages, focusing on bias and mean-squared error for normal samples, and presenting some comparisons from simulation results for non-normal samples.

Keywords: Bias; Kurtosis; Mean-squared error; Skewness

1. Introduction

Skewness and kurtosis measures are often used to describe shape characteristics of a distribution. They have also been used in tests of normality and in studies of robustness to normal theory procedures, as, for example, in Wilcox (1990). The Pearson family of distributions is characterized by the first four moments, and skewness and kurtosis may be used to help to select an appropriate member of this family.

Balanda and MacGillivray (1988) suggested that, like location and scale, skewness and kurtosis should be viewed as 'vague concepts' which can be formalized in many ways. Accordingly, many different definitions have been suggested. In this paper, the traditional measures of sample skewness and kurtosis proposed by, for example, Cramér (1946) are compared with those measures defined and adopted by various statistical computing packages such as SAS and MINITAB. In large samples, the differences in definition are unimportant but for small samples very different values of sample skewness and kurtosis can be obtained by using the various definitions.

We consider the traditional measures of skewness and kurtosis, g_1 and g_2 respectively, with measures G_1 and G_2 adopted by SAS, and b_1 and b_2 adopted by MINITAB. We establish a general relationship between the variances of the three measures and some further results for bias and mean-squared error for samples from a normal distribution.

For large samples, there is very little to choose between the various measures. For small samples from a normal distribution, b_1 and b_2 have smaller mean-squared error than G_1 and G_2 have. However, for small samples from non-normal distributions, a simulation study shows that G_1 and G_2 have the smallest mean-squared error.

2. Definitions and background

Over the years, many measures of sample skewness and kurtosis have been proposed. Cramér

†*Address for correspondence*: Department of Statistics, University of Leeds, Leeds, LS2 9JT, UK.
E-mail: D.N.Joanes@leeds.ac.uk

 0039–0526/98/47183

(1946) is typical of many older texts in defining such coefficients as $g_1 = m_3/m_2^{3/2}$ and $g_2 = m_4/m_2^2 - 3$, where the sample moments for samples of size n are given by

$$m_r = \frac{1}{n} \sum (x_i - \bar{x})^r.$$

As is well known, the sample moments are not unbiased estimates of the population moments μ_r. Thus, for example

$$E(m_2) = \frac{n-1}{n} \mu_2,$$

$$E(m_3) = \frac{(n-1)(n-2)}{n^2} \mu_3,$$

$$E(m_4) = \frac{(n-1)(n^2-3n+3)}{n^3} \mu_4 + \frac{3(n-1)(2n-3)}{n^3} \mu_2^2.$$

Unbiased estimates of the μ_r are readily available by making the appropriate simple corrections.

In a similar way, unbiased cumulant estimates K_j may be defined such that $E(K_j) = \kappa_j$ where, for example, $\kappa_2 = \mu_2$, $\kappa_3 = \mu_3$ and $\kappa_4 = \mu_4 - 3\mu_2^2$.

It may be seen that, in terms of the population cumulants, skewness and kurtosis coefficients γ_1 and γ_2 are given by $\gamma_1 = \kappa_3/\kappa_2^{3/2}$ and $\gamma_2 = \kappa_4/\kappa_2^2$.

As reported in Cramér (1946),

$$K_2 = \frac{n}{n-1} m_2,$$

$$K_3 = \frac{n^2}{(n-1)(n-2)} m_3,$$

$$K_4 = \frac{n^2}{(n-1)(n-2)(n-3)} \{(n+1)m_4 - 3(n-1)m_2^2\}.$$

In the particular case of sampling from a normal distribution, it may be shown (Fisher, 1930) that $E(g_1) = 0$ and $E(g_2) = -6/(n+1)$, or equivalently $E(m_4/m_2^2) = 3(n-1)/(n+1)$.

To remove the bias in g_2 it would be possible to apply a simple correction. However, to achieve consistency, quantities analogous to g_1 and g_2, but involving ratios of the unbiased cumulant estimates K_j, are often preferred.

Let

$$G_1 = \frac{K_3}{K_2^{3/2}} = \frac{\surd\{n(n-1)\}}{n-2} g_1$$

and

$$G_2 = \frac{K_4}{K_2^2} = \frac{n-1}{(n-2)(n-3)} \{(n+1)g_2 + 6\}.$$

The quantities G_1 and G_2 are the definitions of sample skewness and kurtosis adopted by the computing packages SAS and SPSS, and also by the EXCEL spreadsheet program. In contrast, MINITAB and BMDP define skewness and kurtosis by

$$b_1 = \frac{m_3}{s^3} = \left(\frac{n-1}{n}\right)^{3/2} \frac{m_3}{m_2^{3/2}}$$

and

$$b_2 = \frac{m_4}{s^4} - 3 = \left(\frac{n-1}{n}\right)^2 \frac{m_4}{m_2^2} - 3,$$

where

$$s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2.$$

In large samples, the differences in definition will be unimportant. However, in small or moderate samples the differences can be quite startling. For example, consider the data set with the frequency distribution

value	10	11	12	13	14	15	16	17
frequency	1	1	3	6	4	3	1	1

The sample kurtosis G_2 (as given by SAS and SPSS) is 0.2743, whereas b_2 (as used by MINITAB) is -0.3605. Thus the kurtosis excess, compared with a normal distribution, may be positive, or it may be negative!

3. Comparison of skewness and kurtosis measures for normal samples

We now focus attention on the comparison between the skewness measures g_1, G_1 and b_1, and the kurtosis measures g_2, G_2 and b_2 for samples of size n from a normal distribution.

Since both G_1 and b_1 are simple multiples of g_1, it is clear that all three skewness measures are unbiased. Of the three kurtosis measures, however, only G_2 is unbiased. Since $E(g_2) = -6/(n+1)$ and

$$E(b_2) = 3\,\frac{(n-1)^3}{n^2(n+1)} - 3 \simeq \frac{-12}{n+1},$$

the bias in b_2 is approximately twice the bias in g_2 for normal samples.

Further insight may be gained by comparing variances and mean-squared errors. In general let $\mathrm{var}(g_1) = W$, say. Then

$$\mathrm{var}(b_1) = \mathrm{var}\left\{\left(\frac{n-1}{n}\right)^{3/2} g_1\right\} = \left(\frac{n-1}{n}\right)^3 W \simeq \left(1 - \frac{3}{n}\right) W$$

and

$$\mathrm{var}(G_1) = \mathrm{var}\left[\frac{\surd\{n(n-1)\}}{n-2}\, g_1\right] = \frac{\surd\{n(n-1)\}}{(n-2)^2}\, W \simeq \left(1 + \frac{3}{n}\right) W.$$

Similarly let $\mathrm{var}(g_2) = V$. Then

$$\mathrm{var}(b_2) = \mathrm{var}\left[\left(\frac{n-1}{n}\right)^2 g_2 + 3\left\{\left(\frac{n-1}{n}\right)^2 - 1\right\}\right] = \left(\frac{n-1}{n}\right)^4 V \simeq \left(1 - \frac{4}{n}\right) V$$

and

$$\mathrm{var}(G_2) = \mathrm{var}\left[\frac{n-1}{(n-2)(n-3)}\{(n+1)g_2 + 6\}\right] = \left\{\frac{(n-1)(n+1)}{(n-2)(n-3)}\right\}^2 V \simeq \left(1 + \frac{10}{n}\right) V.$$

Thus,

$$\mathrm{var}(b_1) < \mathrm{var}(g_1) < \mathrm{var}(G_1)$$

and

$$\mathrm{var}(b_2) < \mathrm{var}(g_2) < \mathrm{var}(G_2),$$

irrespective of the distribution being sampled.

As given by Cramér (1946), in normal samples

$$\mathrm{var}\left(\frac{m_3}{m_2^{3/2}}\right) = \frac{6(n-2)}{(n+1)(n+3)}$$

and

$$\mathrm{var}\left(\frac{m_4}{m_2^2}\right) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}.$$

Expressions for the variances of all our skewness and kurtosis measures follow simply for normal samples from these equations.

Since for normal samples all three skewness measures are unbiased, the mean-squared error mse will be the same as the variance. Thus

$$\mathrm{mse}(b_1) < \mathrm{mse}(g_1) < \mathrm{mse}(G_1).$$

As far as kurtosis is concerned, the relationship involving mean-squared error turns out to be

$$\mathrm{mse}(g_2) < \mathrm{mse}(b_2) < \mathrm{mse}(G_2), \qquad \text{for all } n \geqslant 2.$$

G_2 has the largest mean-squared error because it has the largest variance, even though it is unbiased. The mean-squared error for b_2 is larger than for g_2 because of the larger bias term in b_2 which provides the dominant contribution to the mean-squared error.

To give some idea of the magnitude of variation between the mean-squared errors in the normal samples case, in Table 1 we compare results for sample sizes $n = 10$, 20, 50, 100. Clearly, for large samples, there is very little to choose between the various measures. For small samples, b_1 and b_2 have smaller mean-squared errors than G_1 and G_2 have.

4. Comparison of skewness and kurtosis measures for non-normal samples

For non-normal distributions, theoretical expressions for the bias and mean-squared error of the sample skewness and kurtosis measures are not generally available, but they can be estimated. As an example, we give average estimates of bias and mean-squared error based on 100 000 samples of various sample sizes n, generated in turn from χ^2_v-distributions with $v = 1$, 10 and 50 degrees of freedom.

When g_1 and g_2 are used as measures of sample skewness and kurtosis, and sampling is from a

Table 1. Relative mean-squared errors of sample skewness and kurtosis in normal samples

n	*Skewness*			*Kurtosis*		
	mse(b_1)	*mse*(g_1)	*mse*(G_1)	*mse*(g_2)	*mse*(b_2)	*mse*(G_2)
10	0.73	1	1.41	1	1.61	2.05
20	0.86	1	1.17	1	1.17	1.49
50	0.94	1	1.06	1	1.03	1.18
100	0.97	1	1.03	1	1.01	1.09

χ^2_v-distribution, the bias in g_1 is $E(g_1) - \sqrt{(8/v)}$ and is $E(g_2) - 12/v$ for g_2. The expectations and variances of g_1 and g_2 are estimated by the mean and variance of the sampling distributions of g_1 and g_2 based on 100 000 random samples of size n from the appropriate χ^2-distribution. The same method is used for the other measures of sample skewness and kurtosis, b_1 and G_1, and b_2 and G_2. The results are presented in Tables 2 and 3.

The consistently negative bias for all three measures of sample skewness (Table 2) decreases as the degrees of freedom v and the sample size n increase, and follows the relationship

$$\text{bias}(G_1) < \text{bias}(g_1) < \text{bias}(b_1).$$

The mean-squared error also decreases with increasing n and v. For χ^2_{10} and χ^2_{50} and any value of n, we see that

$$\text{mse}(b_1) < \text{mse}(g_1) < \text{mse}(G_1),$$

like the normal case. For χ^2_1, owing to the large bias terms in b_1 and g_1, we observe that

$$\text{mse}(G_1) < \text{mse}(g_1) < \text{mse}(b_1).$$

Table 2. Bias and mean-squared error for skewness measures g_1, b_1 and G_1

v	n	*Bias*			*Mean-squared error*		
		g_1	b_1	G_1	g_1	b_1	G_1
1	10	−1.55	−1.73	−1.31	2.78	3.29	2.26
	20	−1.09	−1.22	−0.95	1.67	1.90	1.46
	50	−0.63	−0.70	−0.56	0.94	1.00	0.89
	100	−0.39	−0.43	−0.35	0.66	0.68	0.65
10	10	−0.44	−0.50	−0.35	0.54	0.51	0.61
	20	−0.26	−0.31	−0.21	0.34	0.33	0.37
	50	−0.13	−0.15	−0.10	0.19	0.18	0.19
	100	−0.07	−0.08	−0.05	0.11	0.11	0.11
50	10	−0.19	−0.22	−0.16	0.37	0.29	0.49
	20	−0.11	−0.13	−0.09	0.25	0.22	0.28
	50	−0.05	−0.06	−0.04	0.12	0.12	0.13
	100	−0.03	−0.03	−0.02	0.07	0.07	0.07

Table 3. Bias and mean-squared error for kurtosis measures g_2, b_2 and G_2

v	n	*Bias*			*Mean-squared error*		
		g_2	b_2	G_2	g_2	b_2	G_2
1	10	−11.25	−11.96	−9.71	129.62	145.12	103.75
	20	−9.26	−9.82	−8.05	95.04	103.99	80.70
	50	−6.37	−6.71	−5.63	64.65	67.25	61.28
	100	−4.37	−4.58	−3.91	54.82	55.31	54.84
10	10	−1.60	−2.09	−0.94	3.49	4.99	3.82
	20	−1.09	−1.39	−0.69	2.95	3.37	3.45
	50	−0.59	−0.73	−0.39	2.54	2.55	2.84
	100	−0.33	−0.41	−0.22	2.01	1.99	2.15
50	10	−0.76	−1.23	−0.20	1.22	1.94	2.04
	20	−0.45	−0.72	−0.14	1.01	1.18	1.40
	50	−0.21	−0.33	−0.07	0.74	0.75	0.85
	100	−0.11	−0.17	−0.04	0.49	0.49	0.53

The estimates of bias in the kurtosis measures (Table 3) appear somewhat larger than in the skewness measures but behave in a similar way, decreasing as n and v increase, and following the relationship

$$\text{bias}(G_2) < \text{bias}(g_2) < \text{bias}(b_2).$$

The variances of both b_2 and g_2 increase sharply and then slowly decay, reaching a peak around $n = 200$ for χ^2_1, $n = 45$ for χ^2_{10} and $n = 20$ for χ^2_{50}. The variance of G_2 decreases as n increases and is rather unstable when n and v are very small. As expected, the variance satisfies the relationship

$$\text{var}(b_2) < \text{var}(g_2) < \text{var}(G_2).$$

The relationship between the mean-squared errors of the three measures depends on whether the mean-squared error is being dominated by a large variance term or by a large bias term. Thus, for χ^2_1 and $n < 100$, we have

$$\text{mse}(G_2) < \text{mse}(g_2) < \text{mse}(b_2)$$

but, as n increases further, the bias becomes small relative to the variance and we have

$$\text{mse}(g_2) < \text{mse}(b_2) < \text{mse}(G_2).$$

For very large n (> 200) we have

$$\text{mse}(b_2) < \text{mse}(g_2) < \text{mse}(G_2).$$

As v increases, G_2 generally has the largest mean-squared error whatever the value of n. For small n, g_2 has the smallest mean-squared error whereas b_2 achieves the smallest mean-squared error for n sufficiently large, e.g. $n = 100$ for χ^2_{50}.

The measures of kurtosis were also compared in samples from a t-distribution with 5 degrees of freedom as a representative example of a symmetric heavy-tailed distribution. The bias behaved in the same way as for the χ^2-distributions with

$$\text{bias}(G_2) < \text{bias}(g_2) < \text{bias}(b_2).$$

Again, comparisons of mean-squared errors are dependent on the sample size, since the effect of the variance increasingly dominates the mean-squared error as n increases.

5. Conclusions

For samples from a normal distribution, all three measures of skewness are unbiased, but in small samples the mean-squared error is less for MINITAB's b_1 than for g_1 and greater for SAS's G_1 than for g_1. The kurtosis measure G_2 also has the largest mean-squared error whereas b_2 has a mean-squared error that is only slightly larger than that of g_2. For large samples, there is very little difference between the three measures.

The variances of G_1 and G_2 are greatest whereas b_1 and b_2 have the smallest variances whatever distribution is being sampled.

In samples from the χ^2-distribution, all the measures show negative bias with G_1 and G_2 less biased than g_1 and g_2, and b_1 and b_2 more biased than g_1 and g_2 respectively. The bias decreases for larger samples from the more symmetric χ^2-distributions, i.e. for larger v, and follows the same behaviour seen in normal samples.

The measures G_1 and G_2 have the smallest mean-squared error, and b_1 and b_2 have the largest, for the asymmetric χ^2_1-distribution, because the bias terms involved are larger than the variance terms. For larger n and v, the mean-squared error behaves as for normal samples, except that b_2 has the least mean-squared error when n is very large.

In conclusion, MINITAB's b_1 and b_2 have smaller variance and mean-squared error in normal samples, but SAS's G_1 and G_2 have smaller mean-squared error in samples from a very skewed distribution such as the χ_1^2-distribution.

References

Balanda, K. P. and MacGillivray, H. L. (1988) Kurtosis: a critical review. *Am. Statistn*, **42**, 111–119.
Cramér, H. (1946) *Mathematical Methods of Statistics*. Princeton: Princeton University Press.
Fisher, R. A. (1930) The moments of the distribution for normal samples of measures of departure from normality. *Proc. R. Soc.* A, **130**, 16–28.
Wilcox, R. R. (1990) Comparing the means of two independent groups. *Biometr. J.*, **32**, 771–780.