

On the maximum of the standardized fourth moment

Robert H. Byers, Jr.
Centers for Disease Control
& Prevention

Abstract

For sample size n , it is well known that the kurtosis statistic b_2 (the standardized fourth moment) is less than n . We show that $b_2 \leq n - 2 + \frac{1}{n-1}$, and this maximum value will be reached only when $n - 1$ of the data points are equal to each other and unequal to the n -th point. The skewness statistic also has a maximum with this configuration, and it is $n - 3 + \frac{1}{n-1}$.

Keywords: *kurtosis, skewness, moments.*

Introduction.

The kurtosis statistic is often used to measure how flat or peaked a frequency distribution is. It is also used to judge how close the underlying distribution is to normal. It is well known and easy to prove that kurtosis is less than the sample size. Johnson and Lowe(1979) report the inequalities $b_2 \leq N$ and $|b_1|^{\frac{1}{2}} \leq (N - 1)^{\frac{1}{2}}$. Johnson and Lowe point out that "these bounds have serious consequences for situations in which $b_1^{\frac{1}{2}}$ and b_2 are used to estimate the population skewness and kurtosis." They give an example for the log-normal distribution with kurtosis 113.9. Clearly a small sample size would seriously underestimate the population moments.

This paper establishes upper bounds for b_2 and b_1 that are smaller than those previously known. As a side issue we will also discover that for a sample of size n there is an upper bound on the standardized data which depends only on n .

Notation.

Let $x_i, i = 1, \dots, n$ represent n real numbers. Let m_i be the moment statistics:

$$m_i = \sum_{j=1}^n (x_j - \bar{x})^i / n.$$

The skewness is measured by $b_1^{\frac{1}{2}} = m_3/m_2^{3/2}$, and kurtosis is $b_2 = m_4/m_2^2$. We will refer to the standardized data values as $z_i = (x_i - \bar{x})/\hat{s}$. The mean is $\bar{x} = \sum_{i=1}^n x_i/n$, and the standard deviation is $\hat{s} = (\sum_{i=1}^n (x_i - \bar{x})^2/(n-1))^{\frac{1}{2}}$.

Derivation.

Suppose we have a sample of $n+1$ real numbers x_i , $i = 1, \dots, n+1$. Let x_{n+1} be an arbitrarily chosen x . Write b_2 as sums of powers of the x_i .

$$\begin{aligned} b_2 = (n+1) & \left[\sum_{i=1}^n (x_i^4 + x_{n+1}^4) - 4 \left(\sum_{i=1}^n x_i^3 + x_{n+1}^3 \right) \left(\sum_{i=1}^n x_i + x_{n+1} \right) / (n+1) \right. \\ & + 6 \left(\sum_{i=1}^n x_i^2 + x_{n+1}^2 \right) \left(\sum_{i=1}^n x_i + x_{n+1} \right)^2 / (n+1)^2 \\ & \left. - 3 \left(\sum_{i=1}^n x_i + x_{n+1} \right)^4 / (n+1)^3 \right] \\ & \div \left[\left(\sum_{i=1}^n x_i^2 + x_{n+1}^2 \right) - \left(\sum_{i=1}^n x_i + x_{n+1} \right)^2 / (n+1) \right]^2 \end{aligned} \quad (1)$$

Expanding the numerator and denominator as polynomials in x_{n+1} and performing polynomial division by means of *Mathematica* (Wolfram, 1996), we find that this can be written as $n-1 + 1/n + R$ for a sample of size $n+1$. So for a sample of size n :

$$b_2 = n - 2 + \frac{1}{n-1} + R \quad (2)$$

The remainder R has the form:

$$R = \frac{(c_0 + c_1 x_{n+1} + c_2 x_{n+1}^2)}{\left[\left(\sum_{i=1}^n x_i^2 + x_{n+1}^2 \right) - \left(\sum_{i=1}^n x_i + x_{n+1} \right)^2 / (n+1) \right]^2}. \quad (3)$$

The coefficients of the numerator are:

$$\begin{aligned}
c_0 = \frac{1}{n} & \left[n(n+1) \sum_{i=1}^n x_i^4 - 4n \sum_{i=1}^n x_i \sum_{i=1}^n x_i^3 \right. \\
& - (n^2 - n + 1) \left(\sum_{i=1}^n x_i^2 \right)^2 + 2(n+1) \sum_{i=1}^n x_i^2 \sum_{i=1}^n x_i \\
& \left. - \left(\sum_{i=1}^n x_i \right)^4 \right] \tag{4}
\end{aligned}$$

$$c_1 = -\frac{4}{n} \left[n \sum_{i=1}^n x_i^3 - (n+1) \sum_{i=1}^n x_i \sum_{i=1}^n x_i^2 + \left(\sum_{i=1}^n x_i \right)^3 \right] \tag{5}$$

$$c_2 = \frac{1}{n} \left[\left(\sum_{i=1}^n x_i \right)^2 - n \sum_{i=1}^n x_i^2 \right] \tag{6}$$

A new upper bound on b_2 will be established if we can show that R is nonpositive. We have the identity

$$b_2 = \frac{m_4}{m_2^2} = n - 2 + \frac{1}{n-1} + R. \tag{7}$$

Solving for R gives

$$\begin{aligned}
R &= \frac{m_4 - (n-2 + \frac{1}{n-1})m_2^2}{m_2^2} \\
&= \frac{n \Sigma(x-\bar{x})^4 - (n-2 + \frac{1}{n-1})(\Sigma(x-\bar{x})^2)^2}{(\Sigma(x-\bar{x})^2)^2} \tag{8}
\end{aligned}$$

We write

$$(\Sigma(x-\bar{x})^2)^2 = \Sigma(x-\bar{x})^4 + \sum_{i=1}^n \sum_{j \neq i}^n (x_i - \bar{x})^2 (x_j - \bar{x})^2$$

which can be written more compactly as:

$$S_2^2 = S_4 + S_{22}$$

Now R can be written:

$$R = \frac{n S_4 - (n - 2 + \frac{1}{n-1})(S_4 + S_{22})}{S_4 + S_{22}} \quad (9)$$

Mitrinović(1970) gives this inequality, which he attributes to G. Kalajdžić.

$$\sum_{\nu_1, \dots, \nu_n=0}^k \frac{x_1^{\nu_1} \cdots x_n^{\nu_n}}{\nu_1! \cdots \nu_n!} \leq \frac{n(n^k - 1)}{(k+1)!} \left(\frac{\sum_{i=1}^n x_i}{n} \right)^2 \quad (10)$$

where x_1, \dots, x_n are all nonnegative, $k \geq 1$ is an integer, and $\sum_{i=1}^n \nu_i = k + 1$. Setting

$x_i = y_i^2$ with $k = 1$ gives

$$\frac{S_{22}}{2} \leq \frac{n(n-1)}{2} \left(\frac{\sum_{i=1}^n y_i^2}{n} \right)^2 = \frac{n-1}{2n} (S_4 + S_{22}), \quad (11)$$

which implies that

$$S_4 \geq \frac{S_{22}}{n-1}. \quad (12)$$

Now by substituting $(n-1)S_4$ for S_{22} in the numerator of R we get:

$$\frac{n S_4 - (n - 2 + \frac{1}{n-1})(S_4 + S_{22})}{S_4 + S_{22}} \leq - \frac{n(n-2)^2 S_4}{(n-1)(S_{22} + S_4)} \leq 0. \quad (13)$$

This proves that R is negative, and consequently

$$b_2 \leq n - 2 + \frac{1}{n - 1}. \quad (14)$$

The fact that b_2 can reach this upper limit can be shown by applying l'Hospital's rule twice to R , revealing that $\lim_{x_k \rightarrow \infty} R = 0$.

This result seems rather abstract, since we don't expect to find samples with one extremely large outlier to force b_2 near its limit. Are there reasonable configurations of data points that have kurtosis at or near the limit? There is, in fact, just one.

To find a set of $n + 1$ numbers with maximum kurtosis, observe what happens to the standardized value of an arbitrary element x_k as x_k increases.

$$\lim_{x_k \rightarrow \infty} z_k = \lim_{x_k \rightarrow \infty} \frac{x_k - \left(\sum_{i=1}^n x_i + x_k\right)/(n+1)}{\left[\frac{\sum_{i=1}^n x_i^2 + x_k^2 - \left(\sum_{i=1}^n x_i + x_k\right)^2/(n+1)}{n}\right]^{\frac{1}{2}}} = \frac{n}{(n+1)^{\frac{1}{2}}}. \quad (15)$$

Similarly

$$\lim_{x_k \rightarrow \infty} z_i = -\frac{1}{(n+1)^{\frac{1}{2}}}, \quad i \neq k. \quad (16)$$

Substituting these values into b_2 yields $n - 1 + \frac{1}{n}$. In fact, any set of n numbers with $n - 1$ equal and one different will work. Say there are $n - 1$ numbers x and one y . Then

$$\begin{aligned}
b_2 &= n \frac{(n-1)(x - \frac{(n-1)x+y}{n})^4 + (y - \frac{(n-1)x+y}{n})^4}{((n-1)(x - \frac{(n-1)x+y}{n})^2 + (y - \frac{(n-1)x+y}{n})^2)^2} \quad (17) \\
&= n \frac{(n-1)(\frac{x-y}{n})^4 + ((\frac{n-1}{n})(x-y))^4}{((n-1)(\frac{x-y}{n})^2 + ((\frac{n-1}{n})(x-y))^2)^2} \\
&= n \frac{(x-y)^4(\frac{n-1}{n^4} + (\frac{n-1}{n})^4)}{(x-y)^4(\frac{n-1}{n^2} + (\frac{n-1}{n})^2)^2} \\
&= n \frac{(n-1)(n^3 - 3n^2 + 3n)}{n^2(n-1)^2} \\
&= n - 2 + \frac{1}{n-1}.
\end{aligned}$$

Since b_2 measures how flat or peaked a distribution is, this result makes intuitive sense.

The most peaked distribution possible would have all points equal, but that is degenerate and b_2 is undefined. The next most peaked would place $n - 1$ of the values at one point and one anywhere else.

When there are n values at x and m values at y , $b_2 = \frac{m}{n} + \frac{n}{m} - 1$. When $n = m$, we have a minimum, since $b_2 = 1$.

By a calculation similar to (12), we find that b_1 is $n - 3 + \frac{1}{n-1}$ when b_2 is maximum.

According to Johnson & Kotz (1970, p. 14), $b_2 - b_1 \geq 1$. Therefore $n - 3 + \frac{1}{n-1}$ is a maximum for b_1 .

Discussion.

Dr. L. R. Shenton posed the question "What is the maximum value of the fourth moment in a sample of size n ?" This has led to three surprising results:

$$b_2 \leq n - 2 + \frac{1}{n - 1},$$

$$b_1 \leq n - 3 + \frac{1}{n - 1},$$

$$\max_{i \leq n}(z_i) \leq \frac{n - 1}{n^{\frac{1}{2}}}.$$

Teuscher and Guiard(1995) have investigated inequalities involving skewness and kurtosis. Their work assumes the existence of standardized unimodal frequency distributions, which this paper does not. The inequalities above are only dependent on the existence of n real numbers for their validity. Teuscher and Guiard also point out that the inequality $b_2 - b_1 - 1 \geq 0$ attains equality for 2-point distributions. That is the case when b_2 and b_1 are at their maxima.

References.

1. Johnson, M. E. and V. W. Lowe, Jr.(1979) Bounds on the Sample Skewness and Kurtosis, *Technometrics* 21(3) 377-378.
3. Johnson, Norman L. & Samuel Kotz(1970) *Continuous Univariate Distributions-1* Houghton-Mifflin Co., Boston, MA.
2. Mitrinović, D. S.(1970) *Analytic Inequalities*, Springer-Verlag, Berlin.
4. Teuscher, F. & V. Guiard(1995) Sharp inequalities between skewness and kurtosis for unimodal distributions, *Statistics and Probability Letters* 22, 257-260.
5. Wolfram, Stephen(1996) *The Mathematica Book*, 3rd ed., Wolfram Media/Cambridge University Press.