MULTIVARIATE SKEWNESS

Steven J. Schwager
Biometrics Unit, Cornell University, Ithaca, NY 14853

BU-803-M* January, 1983

Abstract

Multivariate skewness is defined and its properties are discussed. Its use in testing for multivariate normality and in robustness studies is described. Both Mardia's definition and that of Malkovich and Afifi are treated.

* In the Biometrics Unit series, Cornell University, Ithaca, NY 14853.

Multivariate skewness is a generalization of (univariate) skewness*, the standardized third moment, to multivariate distributions and samples.

Mardia [3] defined the multivariate skewness of a distribution with $p \times 1$ mean vector μ and $p \times p$ covariance matrix Σ as

$$\beta_{1,p} = E\{[(X-\mu)'\Sigma^{-1}(Y-\mu)]^3\} \quad ,$$

where X and Y are independent $p \times 1$ random vectors with this distribution. He defined [3] the multivariate sample skewness of the set of $p \times 1$ observations $X_1, \cdots, X_n$ as

$$b_{1,p} = (1/n^2) \sum_{i,j=1}^{n} [(X_i - \bar{X})'S^{-1}(X_j - \bar{X})]^3 \quad ,$$

where the sample mean vector and covariance matrix are

$$\bar{X} = (1/n)\sum_{i=1}^{n} X_i, \qquad S = (1/n)\sum_{i=1}^{n} (X_i - X)(X_i - \bar{X})' \quad .$$

When the dimension p is 1, $\beta_{1,p}$ and $b_{1,p}$, reduce to the squares of the usual univariate population skewness $\sqrt{\beta_1}$ and sample skewness $\sqrt{b_1}$. When $p = 2$ and $\Sigma = I$,

$$\beta_{1,p} = \mu_{30}^2 + \mu_{03}^2 + 3\mu_{12}^2 + 3\mu_{21}^2 \quad .$$

Similar but more complex expressions are available for $p > 2$, $\Sigma \neq I$ [see 3]. For any nonsingular $p \times p$ matrix A and any $p \times 1$ vector D, $b_{1,p}$ is invariant* under the affine transformation $AX + D$ of the sample; $\beta_{1,p}$ is also invariant under this transformation.

The skewness of any distribution symmetric about its mean is $\beta_{1,p} = 0$. Thus the p-dimensional multivariate normal

$N(\mu,\Sigma)$ distribution has skewness $\beta_{1,p} = 0$. For a random sample from this distribution, the statistic $nb_{1,p}/6$ has an asymptotic chi-square distribution* with $p(p+1)(p+2)/6$ degrees of freedom. An improved version of this result follows from the mean of $b_{1,p}$ under normal sampling [4].

Mardia [3,4,5] suggested using his multivariate sample skewness and multivariate sample kurtosis* to test for normality. A test based on skewness is given by rejecting the hypothesis of multivariate normality if $b_{1,p}$ is very large. To perform such a test, tables of critical points of the distribution of $b_{1,p}$ under normal sampling are necessary for small to moderately large values of n. Tables for $p = 2$ and selected values of n from 10 to 5000, produced by Monte Carlo simulations* and smoothing, appear in Mardia [4]. For extremely large n, critical points of $b_{1,p}$ can be approximated from its asymptotic behavior [4].

The magnitude of the multivariate skewness $\beta_{1,p}$ is a measure of the nonnormality of a distribution, which can be useful in robustness studies. Nonnormality reflected by $\beta_{1,p}$ affects the size of Hotelling's T^2 test*, but does not appear to have much effect on the normal theory likelihood ratio test* for equal covariance matrices in several populations [3,4,5].

Malkovich and Afifi [2] introduced another definition of multivariate skewness, based on Roy's union-intersection principle#. If the distribution of X has $p \times 1$ mean vector μ and $p \times p$ covariance matrix Σ, then, for any nonzero $p \times 1$ vector C, the scalar variable C'X has squared skewness

$$\beta_1(C) = \{E[(C'X - C'\mu)^3]\}^2/(C'\Sigma C)^3 \ .$$

The multivariate skewness of the distribution of X is [2]

$$\beta_1^M = \max_C \beta_1(C) \ ,$$

the largest squared skewness produced by any projection of the p-dimensional distribution onto a line.

For a sample $X_1, \cdots, X_n$ of $p \times 1$ observations and any nonzero $p \times 1$ vector C, the square of the sample skewness of the scalars $C'X_1, \cdots, C'X_n$ is

$$b_1(C) = n[\sum_{i=1}^{n}(C'X_i - C'\bar{X})^3]^2/[\sum_{i=1}^{n}(C'X_i - C'\bar{X})^2]^3 \ .$$

The multivariate sample skewness of $X_1, \cdots, X_n$ is [2]

$$b_1^M = \max_C b_1(C) \ .$$

When the dimension p is 1, β_1^M and b_1^M reduce to the squares β_1 and b_1 of the usual population skewness and sample skewness. Both β_1^M and b_1^M are invariant under nonsingular affine transformations $AX + D$. The multivariate normal $N(\mu,\Sigma)$ distribution has $\beta_1^M = 0$, since $\beta_1(C) = 0$ for every C. Malkovich and Afifi [2] advocated using their multivariate sample skewness and kurtosis to test for multivariate normality. A union-intersection test based on skewness is given by rejecting the hypothesis of normal random sampling whenever b_1^M is very large. The maximization needed to obtain b_1^M involves computations whose difficulty increases with p .

References

[1] Gnanadesikan, R. (1977). Methods for Statistical Data Analysis of Multivariate Observations. Wiley, New York. (Chap. 5 has examples of Mardia's sample skewness.)

[2] Malkovich, J. F. and Afifi, A. A. (1973). J. Amer. Statist. Ass., 68, 176-179.

[3] Mardia, K. V. (1970). Biometrika, 57, 519-530.

[4] Mardia, K. V. (1974). Sankhyā B, 36, 115-128.

[5] Mardia, K. V. (1975). Appl. Statist., 24, 163-171.

[6] Mardia, K. V. and Zemroch, P. J. (1975). Appl. Statist., 24, 262-265. (FORTRAN program for sample skewness.)

(DEPARTURES FROM NORMALITY, TESTS FOR

MULTIVARIATE KURTOSIS

SKEWNESS)