STATISTICS IN TRANSITION new series, Winter 2014
Vol. 15, No. 1, pp. 145–152

MODELLING OF SKEWNESS MEASURE DISTRIBUTION

Margus Pihlak[1]

ABSTRACT

In this paper the distribution of random variable skewness measure is modelled. Firstly, we present some results of matrix algebra useful in multivariate statistical analyses. Then, we apply the central limit theorem on modelling of skewness measure distribution. Finally, we give an idea for finding the confidence intervals of statistical model residuals' asymmetry measure.

Key words: central limit theorem, multivariate skewness measure, skewness measure distribution, statistical model residuals.

1. Introduction and basic notations

Firstly, we introduce some notations used in the paper. The zero vector is denoted as $\mathbf{0}$. The transposed matrix $\mathbf{A}$ is denoted as $\mathbf{A}^{\mathrm{T}}$.

Let us have random vectors $\mathbf{X}_i = (X_{i1}, X_{i2}, \ldots, X_{ik})^{\mathrm{T}}$ where index $i = 1,2,\ldots,n$ is for observations and k denotes the number of variables. These random vectors are independent and identically distributed copies (observations) of a random k-vector $\mathbf{X}$. Let

$$\overline{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i$$

and

$$\mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{X}_i - \overline{\mathbf{x}})(\mathbf{X}_i - \overline{\mathbf{x}})^{\mathrm{T}}$$

be the estimators of the sample mean $E(\mathbf{X}) = \boldsymbol{\mu}$ and the covariance matrix $D(\mathbf{X}) = \boldsymbol{\Sigma}$, respectively.

[1] Tallinn University of Technology (Estonia), Department of Mathematics, Ehitajate tee 5 19086 Tallinn. E-mail: margus.pihlak@ttu.ee.

Now, we present matrix operations used in this paper. One of the widely used matrix operation in multivariate statistics is Kronecker product (or tensor product) $\mathbf{A}\otimes\mathbf{B}$ of $\mathbf{A}: m\times n$ and $\mathbf{B}: p\times q$ which is defined as a partitioned matrix

$$\mathbf{A}\otimes\mathbf{B} = \lfloor a_{ij}\mathbf{B}\rfloor, \quad i = 1,2,\ldots,m;\ j = 1,2,\ldots,n.$$

By means of Kronecker product we can present the third and the fourth order moments of vector $\mathbf{X}$:

$$m_3(\mathbf{X}) = E(\mathbf{X}\otimes\mathbf{X}^{\mathrm{T}}\otimes\mathbf{X})$$

and

$$m_4(\mathbf{X}) = E(\mathbf{X}\otimes\mathbf{X}^{\mathrm{T}}\otimes\mathbf{X}\otimes\mathbf{X}^{\mathrm{T}}).$$

The corresponding central moments

$$\overline{m}_3(\mathbf{X}) = E\left\{(\mathbf{X}-\boldsymbol{\mu})\otimes(\mathbf{X}-\boldsymbol{\mu})^{\mathrm{T}}\otimes(\mathbf{X}-\boldsymbol{\mu})\right\}$$

and

$$\overline{m}_4(\mathbf{X}) = E\left\{((\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^{\mathrm{T}})\otimes((\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^{\mathrm{T}})\right\}.$$

The operation $\mathrm{vec}(\mathbf{A})$ denotes a mn-vector obtained from $m\times n$-matrix by stacking its columns one under another in the natural order. For the properties of Kronecker product and vec-operator the interested reader is referred to Harville (1997), Kollo (1991) or Kollo and von Rosen (2005). In the next section skewness measure will be defined be means of the star-product of the matrices. The star-product was introduced in (MacRae, 1974) where some basic properties of the operation were presented and proved.

Definition 1. *Let us have a matrix* $\mathbf{A}: m\times n$ *and a partitioned matrix* $\mathbf{B}: mr\times ns$ *consisting of* $r\times s$*-blocks* B_{ij}, $i = 1,2,\ldots,m;\ j = 1,2,\ldots,n$. *Then, the star-product* $\mathbf{A}*\mathbf{B}$ *is a* $r\times s$*-matrix*

$$\mathbf{A}*\mathbf{B} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\mathbf{B}_{ij}.$$

The star-product is an inverse operation of Kronecker product in a sense of increasing and decreasing the matrix dimensions. One of the star-product applications is presented in the paper (Pihlak, 2004).

We also use the matrix derivative defined following Neudecker (1969).

Definition 2. *Let the elements of a matrix* $\mathbf{Y}: r\times s$ *be functions of a matrix* $\mathbf{X}: p\times q$. *Assume that for all* $i = 1,2,\ldots,p$, $j = 1,2,\ldots,q$, $k = 1,2,\ldots,r$ *and* $l = 1,2,\ldots,s$ *partial derivatives* $\dfrac{\partial Y_{kl}}{\partial X_{ij}}$ *exist and are continuous in an open set A. Then, the matrix* $\dfrac{d\mathbf{Y}}{d\mathbf{X}}$ *is called matrix derivative of the matrix* $\mathbf{Y}: r\times s$ *by the matrix* $\mathbf{X}: p\times q$ *in a set A, if*

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d}{d\mathrm{vec}^T(\mathbf{X})} \otimes \mathrm{vec}(\mathbf{Y})$$

where

$$\frac{d}{d\mathrm{vec}^{\mathrm{T}}(\mathbf{X})} = \left(\frac{\partial}{\partial x_{11}}, \ldots, \frac{\partial}{\partial x_{p1}}, \ldots, \frac{\partial}{\partial x_{1q}}, \ldots, \frac{\partial}{\partial x_{pq}} \right).$$

The matrix derivative defined by Definition 2 is called Neudecker matrix derivative. This matrix derivative has been in the last 40 years a useful tool in multivariate statistics.

2. Multivariate measures of skewness

In this section we present a multivariate skewness measure by means of the matrix operation described above. A skewness measure in multivariate case was introduced in Mardia (1970). Mori et al. (1993) introduced a skewness measure as a vector. B. Klar (2002) gave a thorough overview of the skewness problem. In this paper asymptotic distribution of different skewness characteristics is also examined. In Kollo (2008) a skewness measure vector is introduced and applied in Independent Component Analyses (ICA). In this paper we give an idea for the application of a skewness measure to residuals of statistical models. Our aim is to estimate the distribution of skewness measure and to find confidence intervals of the asymmetry characteristics.

The skewness measure in the multivariate case is presented through the third order moments:

$$\mathbf{s}(\mathbf{X}) = E(\mathbf{Y} \otimes \mathbf{Y}' \otimes \mathbf{Y}) \tag{2.1}$$

where

$$\mathbf{Y} = \mathbf{\Sigma}^{-\frac{1}{2}}(\mathbf{X} - \boldsymbol{\mu}).$$

In Kollo (2008) the skewness measure based on (2.1) is introduced by means of the star product:

$$\mathbf{b}(\mathbf{X}) = \mathbf{1}_{k\times k} * \mathbf{s}(\mathbf{X}) \tag{2.2}$$

where $k \times k$ -matrix

$$\mathbf{1}_{k\times k} = \begin{pmatrix} 1 & \cdots & 1 \\ \cdots & \ddots & \cdots \\ 1 & \cdots & 1 \end{pmatrix}.$$

In Kollo and Srivastava (2004) the Mardia's skewness measure is presented through the third order moment:

$$\beta = \mathrm{tr}(m_3^{\mathrm{T}}(\mathbf{Y}) m_3(\mathbf{Y}))$$

where operation tr denotes the trace of matrix.

The equality (2.2) generalizes the univarite ($k=1$) skewness measure

$$b(X)=\frac{E(X-\mu)^3}{\sigma^3}$$

where σ denotes standard deviation of the random variable *X*. Thus, we can express the estimator of the univariate skewness measure as

$$\hat{b}(X)=\frac{E(X_i-\bar{x})^3}{s^3} \tag{2.3}$$

where *s* denotes unbiased estimator of standard deviation σ and $\bar{x}$ is the sample mean estimator.

3. Modelling distribution of the univariate skewness measure

In this section we model the distribution of the random variable $\hat{b}(X)$ defined by the equation (2.3). Let us have independent and identically distributed random variables $X_1, X_2, \ldots, X_n$.

Let $E(X)=\mu$ and $D(X)=\sigma^2$. Then, according to the central limit theorem the distribution of the random variable $\frac{\sqrt{n}(\bar{x}-\mu)}{\sigma}$ converges to the normal distribution $N(0,1)$. In the multivariate case the distribution of the random vector $\sqrt{n}(\bar{\mathbf{x}}-\boldsymbol{\mu})$ converges to normal distribution $N(\mathbf{0},\boldsymbol{\Sigma})$.

Let us have $k+k^2$-vector

$$\mathbf{Z}_n=\begin{pmatrix}\bar{\mathbf{x}}\\ \mathrm{vec}(\mathbf{S})\end{pmatrix}.$$

Then

$$\sqrt{n}(\mathbf{Z}_n-E(\mathbf{Z}_n))\mapsto N(\mathbf{0},\boldsymbol{\Pi})$$

in distribution. Here, $(k^2+k)\times(k^2+k)$ -dimensional partitioned matrix

$$\boldsymbol{\Pi}=\begin{pmatrix}\boldsymbol{\Sigma} & \overline{m}_3^{\mathrm{T}}(\mathbf{X})\\ \overline{m}_3(\mathbf{X}) & \boldsymbol{\Pi}_4\end{pmatrix}$$

where $k^2\times k^2$ -matrix $\boldsymbol{\Pi}_4=\overline{m}_4(\mathbf{X})-\mathrm{vec}(\boldsymbol{\Sigma})\mathrm{vec}^{\mathrm{T}}(\boldsymbol{\Sigma})$ (Parring, 1979). This convergence can be generalized by means of the following theorem.

Theorem 1. Let $\{\mathbf{Z}_n\}$ be a sequence of $k+k^2$-component random vectors and $\mathbf{v}$ be a fixed vector such that $\sqrt{n}(\mathbf{Z}_n - \mathbf{v})$ has the limiting distribution $N(\mathbf{0},\mathbf{\Pi})$ as $n \to \infty$. Let the function $g: R^{k^2+k} \to R^k$ have continuous partial derivatives at $\mathbf{z}_n = \mathbf{v}$. Then, the distribution of random variable $\sqrt{n}\{g(\mathbf{Z}_n) - g(\mathbf{v})\}$ converges to the normal distribution $N(\mathbf{0}, g_{\mathbf{z}_n}^{\mathrm{T}} \mathbf{\Pi} g_{\mathbf{z}_n})$ where $(k^2+k)\times k$ -matrix

$$g_{\mathbf{z}_n} = \frac{dg(\mathbf{z}_n)}{d\mathbf{z}_n}\Big|_{\mathbf{z}_n=\mathbf{v}}$$

is Neudecker matrix derivative at $\mathbf{z}_n = \mathbf{v}$.

The proof of the theorem can be found in the book of T. W. Anderson (2003, page 132). In Theorem 1 vector $\mathbf{v} = E(\mathbf{Z}_n)$.

In the univariate case

$$\mathbf{Z}_n = \begin{pmatrix} \overline{X} \\ s^2 \end{pmatrix},$$

$E(\mathbf{Z}_n) = \begin{pmatrix}\mu & \sigma^2\end{pmatrix}^{\mathrm{T}}$ and the function $g(\mathbf{z}_n)$ is defined by equality (2.3).

In this case 2×2 -matrix

$$\mathbf{\Pi} = \begin{pmatrix} \sigma^2 & \overline{m}_3(X) \\ \overline{m}_3(X) & \Pi_4 \end{pmatrix}$$

where

$$\Pi_4 = \overline{m}_4(X) - \sigma^4.$$

Let us take

$$g(\mathbf{z}_n) = g(\overline{x}, s^2) = \hat{b}(X) = \frac{E(X-\overline{x})^3}{(s^2)^{\frac{3}{2}}}.$$

We get

$$g_{\mathbf{z}_n} = \left(\frac{\partial}{\partial \overline{x}} \frac{E(X-\overline{x})^3}{(s^2)^{\frac{3}{2}}} \quad \frac{\partial}{\partial s^2} \frac{E(X-\overline{x})^3}{(s^2)^{\frac{3}{2}}} \right)^{\mathrm{T}} \Bigg|_{\mathbf{z}_n = \binom{\mu}{\sigma^2}} =$$

$$= \left(\frac{-3E(X-\overline{x})^2}{s^3} \quad \frac{-3E(X-\overline{x})^3}{2s^5} \right)^{\mathrm{T}} \Bigg|_{\mathbf{z}_n = \binom{\mu}{\sigma^2}} =$$

$$= -\left(\frac{3}{\sigma} \quad \frac{3\overline{m}_3(X)}{2\sigma^5} \right)^{\mathrm{T}}.$$

According to the Theorem 1 we can say that the random variable $\sqrt{n}\left\{\hat{b}(X)-b(X)\right\}$ has approximately normal distribution $N(0,\sigma_b^2)$. Variance σ_b^2 can be found in the following way:

$$\sigma_b^2 = g_{\mathbf{Z}_n}^{\mathrm{T}} \mathbf{\Pi} g_{\mathbf{Z}_n} =$$

$$= \begin{pmatrix} \dfrac{-3}{\sigma} & \dfrac{-3\overline{m}_3(X)}{2\sigma^5} \end{pmatrix} \begin{pmatrix} \sigma^2 & \overline{m}_3(X) \\ \overline{m}_3(X) & \Pi_4 \end{pmatrix} \begin{pmatrix} \dfrac{-3}{\sigma} \\ \dfrac{-3\overline{m}_3(X)}{2\sigma^5} \end{pmatrix} =$$

$$= 9\frac{\overline{m}_3^2(X)}{\sigma^6}\left(\frac{\Pi_4}{4\sigma^4}+1\right)+9.$$

Thus, we have

$$\sigma_b^2 = 9\frac{\overline{m}_3^2(X)}{\sigma^6}\left(\frac{\Pi_4}{4\sigma^4}+1\right)+9. \tag{3.1}$$

Example. Let us generate *m* times random variable *X* with a sample size *n*. Let random variable *X* have exponential distribution with parameter $\lambda > 0$. Then, the *i*-th order moment $E(X^i) = \dfrac{i!}{\lambda^i}$. Using these moments we get that $\overline{m}_3(X) = \dfrac{2}{\lambda^3}$ and $\Pi_4 = \dfrac{8}{\lambda^4}$. According to the formula (3.1) variance $\sigma_b^2 = 117$. Thus, we get the following approximate 0.95-confidence interval for the skewness measure $b(X)$:

$$\hat{b}(X) \pm \frac{1.96\sqrt{117}}{\sqrt{nm}}.$$

4. Summary: skewness confidence intervals for statistical models

The problem concerns the estimation of statistical models. This is the problem of skewness or lack of symmetry, which means the distribution of statistical model residuals is frequently non-gaussian, as Kolmogorov-Smirnov test shows. In this case the skewness has to be estimated for testing the goodness of models. The confidence intervals of that parameter have to be found. This enables us to improve the diagnosis of statistical models. By means of skewness confidence intervals we can estimate the influence of outliers. These outliers are typical in forestry. The main question is: does the zero value belong to the estimated confidence interval? To answer this question we can estimate the variance of residuals by means of equality (3.1). This variance depends on standard deviation, skewness and kurtosis of residuals.

Acknowledgements

This paper is supported by Estonian Ministry of Education and Science target financed theme No. SF0140011s09.

REFERENCES

ANDERSON, T. W., (2003). An Introduction to Multivariate Statistical Analysis. Wiley Interscience.

HARVILLE, A., (1997). Matrix Algebra from a Statistician's Perspective. Springer, New York.

KLAR, B., (2002). A Treatment of Multivariate Skewness, Kurtosis, and Related Statistics. Journal of Multivariate Analysis, 83, 141–165.

KOLLO, T., (1991). Matrix Derivative in Multivariate Statistics. Tartu University Press, Tartu (in Russian).

KOLLO, T., SRIVASTAVA, M. S., (2004). Estimation and testing of parameters in multivariate Laplace distribution. Comm. Statist., 33, 2363–2687.

KOLLO, T., VON ROSEN, D., (2005). Advanced Multivariate Statistics with Matrices. Springer, Dordrecht.

KOLLO, T., (2008). Multivariate skewness and kurtosis measures with an application in ICA. Journal of Multivariate Analyses, 99, 2328–2338.

MacRAE, E. C., (1974). Matrix derivatives with an application to an adaptive linear decision problem. Ann. Statist., 2, 337–346.

MARDIA, K. V., (1970). Measures of multivariate skewness and kurtosis measures with applications. Biometrika, 57, 519–530.

MORI, T. F., ROHATGI, V. K., SZÉKELY., (1993). On multivariate skewness and kurtosis. Theory Probab. Appl, 38, 547–551.

NEUDECKER, H., (1969). Some theorems on matrix differentiations with special reference to Kronecker matrix products. Journal of the American Statistical Association, 64, 953–963.

PARRING, A-M., (1979). Estimation asymptotic characteristic function of sample (in Russian). Acta et Commetationes Universitatis Tartuensis de Mathematica, 492, 86–90.

PIHLAK, M., (2004). Matrix integral. Linear Algebra and its Applications, 388, 315–325.