Listy Biometryczne - Biometrical Letters
Vol. 33 (1996), No. 1, 25-31

Asymptotically normal distributions in the multivariate Gauss-Markoff model

Wiktor Oktaba

Agricultural University of Lublin
Akademicka 13, 20–934 Lublin, Poland

SUMMARY

Explicit formulas for asymptotically normal distributions of three random variables involving determinants of sums of squares and products matrices for error, hypothesis and "total" and the determinant of the matrix $\sigma^2\mathbf{\Sigma}$ in the multinormal Gauss-Markoff model with the covariance matrix $\sigma^2\mathbf{\Sigma}\otimes\mathbf{V}$ are given. By applying these results the asymptotically normal confidence intervals for $|\sigma^2\mathbf{\Sigma}|$ on the basis of three sums of squares and products matrices are presented.

KEY WORDS: digamma and trigamma function, multivariate gamma function, Euler's constant, standard multivariate distribution, asymptotical normality, determinant, confidence interval, Wishart distribution

1. Multivariate Gauss-Markoff model and notation

Let us consider a multivariate Gauss-Markoff (MGM) model of the form

$$(\mathbf{U},\mathbf{XB},\sigma^2\mathbf{\Sigma}\otimes\mathbf{V}), \tag{1.1}$$

where

$$E(\mathbf{U})=\mathbf{XB} \tag{1.2}$$

is the expected value of the $n\times p$ random matrix $\mathbf{U}$ and

$$Cov(\mathbf{U})=\sigma^2\mathbf{\Sigma}\otimes\mathbf{V}$$

is the covariance matrix of $\mathbf{U}$. Let us assume that

$$\mathbf{U}\sim N_{n,p}(\mathbf{XB},\sigma^2\mathbf{\Sigma}\otimes\mathbf{V}),$$

where $N_{n,p}$ determines multivariate distribution with a $p\times p$ unknown matrix $\mathbf{\Sigma}>\mathbf{0}$, an $n\times n$ matrix $\mathbf{V}\geq 0$, and a scalar $\sigma^2>0$.

Following Rao (1973) we introduce an $n \times n$ matrix

$$\mathbf{T} = \mathbf{V} + \mathbf{XMX}'$$

where $\mathbf{M} = \mathbf{M}'$ is such that the range of $\mathbf{X}$, $R(\mathbf{X})$, is contained in $R(\mathbf{T})$, i.e. $R(\mathbf{X}) \subset R(\mathbf{T})$.

It is known (Rao, 1973) that

$$\begin{bmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X}' & \mathbf{0} \end{bmatrix}^{-} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \\ \mathbf{C}_3 & -\mathbf{C}_4 \end{bmatrix},$$

where

$$\mathbf{C}_1 = \mathbf{T}^{-} - \mathbf{T}^{-}\mathbf{X}(\mathbf{X}'\mathbf{T}^{-}\mathbf{X})^{-}\mathbf{X}'\mathbf{T}, \tag{1.3}$$

$$\mathbf{C}_3 = \mathbf{C}_2' = (\mathbf{X}'\mathbf{T}^{-}\mathbf{X})^{-}\mathbf{X}'\mathbf{T}^{-}, \tag{1.4}$$

$$\mathbf{C}_4 = (\mathbf{X}'\mathbf{T}^{-}\mathbf{X})^{-} - \mathbf{M}. \tag{1.5}$$

The symbol $\mathbf{A}^{-}$ denotes here a g-inverse of the matrix $\mathbf{A}$, i.e. such a matrix $\mathbf{A}^{-}$ that $\mathbf{AA}^{-}\mathbf{A} = \mathbf{A}$.

For (1.1) the error sums of squares and products (SSP) matrix $\mathbf{S}_e$ takes the form

$$\mathbf{S}_e = \mathbf{U}'\mathbf{C}_1\mathbf{U} \tag{1.6}$$

with

$$\nu_e = r(\mathbf{V} \vdots \mathbf{X}) - r(\mathbf{X}) \tag{1.7}$$

degrees of freedom; the symbol $(\mathbf{V} \vdots \mathbf{X})$ denotes here the matrix involving two submatrices $\mathbf{V}$ and $\mathbf{X}$.

If we put in (1.1) $\mathbf{V} = \mathbf{I}_n$ and $\sigma^2 = 1$, where $\mathbf{I}_n$ is a $n \times n$ unit matrix, we get the standard multivariate model of the form

$$(\mathbf{U}, \mathbf{XB}, \mathbf{\Sigma} \otimes \mathbf{I}_n). \tag{1.8}$$

The function (Johnson and Kotz, 1992, p.7)

$$\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)} \tag{1.9}$$

is called the digamma function or the psi function and the function

$$\psi'(x) = \frac{d}{dx}[\psi(x)] \tag{1.10}$$

is called the trigamma function. The tables of the digamma and trigamma functions are contained in Davis (1933) and Janke *et al.* (1968, p.58).

2. Expected value and variance of logarithm of the chi-square variate

The expected value and variance of logarithm of the chi-square variate are equal to

$$E(\ln \chi^2) = \ln 2 + \psi\left(\frac{\nu}{2}\right) \tag{2.1}$$

and

$$\text{var}(\ln \chi^2) = \psi'\left(\frac{\nu}{2}\right) \tag{2.2}$$

respectively, where ν is the number of degrees of freedom while ψ and ψ' are given in (1.9) and (1.10).

3. Asymptotically normal distribution in a multivariate model

The determinant of a matrix $\mathbf{A}$ is denoted as $|\mathbf{A}|$. Let us consider the multinormal Gauss-Markoff model. Then it is known (Oktaba and Kieloch, 1993) that,

$$\mathbf{S}_e = \mathbf{U}'\mathbf{C}_1\mathbf{U} \sim W_p(\nu_e, \sigma^2\mathbf{\Sigma}), \tag{3.1}$$

where ν_e is as in (1.7), while W_p denotes the Wishart distribution. Moreover, if the hypothesis $H_0 : \mathbf{KB} = \boldsymbol{\varphi}$ is true, then the corresponding SSP matrix is

$$\mathbf{S}_H = (\mathbf{K}\hat{\mathbf{B}} - \boldsymbol{\varphi})'\mathbf{L}^-(\mathbf{K}\hat{\mathbf{B}} - \boldsymbol{\varphi}) \sim W_p(\nu_H, \sigma^2\mathbf{\Sigma}) \tag{3.2}$$

and the total SSP matrix is

$$\mathbf{S}_y = \mathbf{S}_H + \mathbf{S}_e \sim W_p(\nu_y, \sigma^2\mathbf{\Sigma}), \tag{3.3}$$

where $\mathbf{L} = \mathbf{KC}_4\mathbf{K}'$, $\hat{\mathbf{B}} = \mathbf{C}_3\mathbf{U}$, $\nu_H = r(\mathbf{L})$ and $\nu_y = \nu_e + \nu_H$.

It is known (Oktaba, 1995a) that if $\mathbf{S} \sim W_p(\nu, \mathbf{\Sigma})$, then

$$\frac{|\mathbf{S}|}{|\mathbf{\Sigma}|} \sim z_1 \cdot z_2 \cdot \ldots \cdot z_p \tag{3.4}$$

is distributed as the product of p mutually independent random variables $z_1, z_2, \ldots, z_p$ each distributed as χ^2 with $\nu, \nu - 1, \ldots, \nu - p + 1$ degrees of freedom, respectively.

From (3.4) we get

$$\ln \frac{|\mathbf{S}|}{|\mathbf{\Sigma}|} \sim \sum_{j=1}^{p} \ln z_j \text{ where } z_j \sim \chi^2_{\nu-j+1}. \tag{3.5}$$

Olsen (1938) considered the distribution of variate $\ln X$, when X has standard gamma distribution, and in particular has chi-square distribution. Approximating of the distribution of the variate $\ln z_j$ by the normal distribution (Johnson and Kotz, 1970, p.181,196) with the expected value (cf.(2.1))

$$\mu_j = E \ln z_j = \psi\left(\frac{\nu - j + 1}{2}\right) + \ln 2, \; j = 1, 2, ..., p, \tag{3.6}$$

and the variance (cf.(2.2))

$$\sigma_j^2 = \psi'\left(\frac{\nu - j + 1}{2}\right), \tag{3.7}$$

gives

$$\ln \frac{|\mathbf{S}|}{|\mathbf{\Sigma}|} \sim AN(\mu, \sigma^2), \tag{3.8}$$

where

$$\mu = \sum_{j=1}^{p} \mu_j = p \ln 2 + \sum_{j=1}^{p} \psi\left(\frac{\nu - j + 1}{2}\right), \tag{3.9}$$

$$\sigma^2 = \sum_{j=1}^{p} \sigma_j^2 = \sum_{j=1}^{p} \psi'\left(\frac{\nu - j + 1}{2}\right) \tag{3.10}$$

and $AN(\mu, \sigma^2)$ denotes the asymptotically normal distribution with mean μ and variance σ^2.

4. Asymptotically normal distributions in MGM model with singular covariance matrix

We are interested in asymptotically normal distribution of sum of logarithms of p mutually independent variates distributed as chi-square with $\nu, \nu - 1, ..., \nu - p + 1$ degrees of freedom (cf. (3.4)).

THEOREM 4.1. *In the multivariate model of the form (1.1) under condition $\nu_e \to \infty, (\nu_e \geq p)$ we have*

$$\ln \frac{|\mathbf{S}_e|}{|\sigma^2 \mathbf{\Sigma}|} = \sum_{j=1}^{p} \ln \chi^2_{\nu_e - j+1} \sim AN(\mu_e, \sigma_e^2), \tag{4.1}$$

where

$$\mu_e = p \ln 2 + \sum_{j=1}^{p} \psi\left(\frac{\nu_e - j + 1}{2}\right), \tag{4.2}$$

$$\sigma_e^2 = \sum_{j=1}^{p} \psi'\left(\frac{\nu_e - j + 1}{2}\right). \tag{4.3}$$

Proof. It is known (Oktaba, 1995b) that

$$\frac{|\mathbf{S}_e|}{|\sigma^2 \mathbf{\Sigma}|} = \prod_{j=1}^{p} \zeta_j, \tag{4.4}$$

where ζ_j are mutually independent variates distributed as chi-square with $\nu, \nu - 1, ..., \nu - p + 1$ degrees of freedom, respectively. Moreover (cf. (3.1)) $\mathbf{S}_e \sim W_p(\nu_e, \sigma^2\mathbf{\Sigma})$ in the MGM model of the form (1.1), where $\mathbf{S}_e$ and ν_e are as in (1.6) and (1.7), respectively. In spite of (4.4), (3.5) and the procedure giving (3.5)-(3.10) and the definition of Wishart distribution we get (4.2). It is sufficient to substitute ν_e and $\sigma^2\mathbf{\Sigma}$ instead of ν and $\mathbf{\Sigma}$ respectively in relations (3.8), (3.9) and (3.10). □

THEOREM 4.2. *In the multivariate model of the form (1.1) under condition $\nu_H \to \infty, (\nu_H \geq p)$ we have*

$$\ln \frac{|\mathbf{S}_H|}{|\sigma^2\mathbf{\Sigma}|} = \sum_{j=1}^{p} \ln \chi^2_{\nu_H - j + 1} \sim AN(\mu_H, \sigma_H^2), \tag{4.5}$$

where

$$\mu_H = p \ln 2 + \sum_{j=1}^{p} \psi\left(\frac{\nu_H - j + 1}{2}\right), \tag{4.6}$$

$$\sigma_H^2 = \sum_{j=1}^{p} \psi'\left(\frac{\nu_H - j + 1}{2}\right), \tag{4.7}$$

and $\mathbf{S}_H$ is as in (3.2).

Proof is analogous to the proof of the Theorem 4.1.

THEOREM 4.3. *In the multivariate model of the form (1.1) under condition $\nu_y \to \infty, (\nu_y \geq p)$ we have*

$$\ln \frac{|\mathbf{S}_y|}{|\sigma^2\mathbf{\Sigma}|} = \sum_{j=1}^{p} \ln \chi^2_{\nu_y - j + 1} \sim AN(\mu_y, \sigma_y^2), \tag{4.8}$$

where

$$\mu_y = p \ln 2 + \sum_{j=1}^{p} \psi\left(\frac{\nu_y - j + 1}{2}\right), \tag{4.9}$$

$$\sigma_y^2 = \sum_{j=1}^{p} \psi'\left(\frac{\nu_y - j + 1}{2}\right), \tag{4.10}$$

and $\mathbf{S}_H$ is as in (3.3) and $\nu_y = \nu_e + \nu_H$.

Proof is analogous to the proof of the Theorem 4.1.

5. Approximate confidence intervals for generalized variance

THEOREM 5.1. *Asymptotically normal* $(1-\alpha)$ *confidence interval for the determinant* $|\sigma^2\mathbf{\Sigma}|$ *in the MGM model (1.1) under* $\nu_e \to \infty, (\nu_e \geq p)$ *is of the form*

$$P\left\{\frac{|\mathbf{S}_e|}{\exp(\mu_e+\sigma_e u_\alpha)} \leq |\sigma^2\mathbf{\Sigma}| \leq \frac{|\mathbf{S}_e|}{\exp(\mu_e+\sigma_e u_\alpha)}\right\} \approx 1-\alpha, \tag{5.1}$$

where $\mathbf{S}_e, \mu_e$ *and* σ_e^2 *are as in (1.6), (4.2) and (4.3), respectively,* $u\left(1-\frac{\alpha}{2}\right)$ *is the value of the variate* $\mathcal{U}$ *with standard normal distribution* $N(0,1)$ *such that*

$$P\left\{-u\left(1-\frac{\alpha}{2}\right) < \mathcal{U} < u\left(1-\frac{\alpha}{2}\right)\right\} = 1-\alpha. \tag{5.2}$$

Proof. In virtue of (4.1) we have directly

$$P\left\{-u_\alpha < \frac{u_e-\mu_e}{\sigma_e} < u_\alpha\right\} \approx 1-\alpha, \tag{5.3}$$

where

$$u_e = \ln\frac{|\mathbf{S}_e|}{|\sigma^2\mathbf{\Sigma}|}. \tag{5.4}$$

From (5.3) we get immediately (5.1). □

The remaining two asymptotically normal confidence intervals for $|\sigma^2\mathbf{\Sigma}|$ we obtain by using 1) $\mathbf{S}_H$, μ_H, σ_H and 2) $\mathbf{S}_y, \mu_y, \sigma_y$ (cf. (4.6), (4.7),(4.9),(4.10)). It is enough to substitute symbols $\mathbf{S}_e, \mu_e, \sigma_e$ as in (5.1) by 1° $\mathbf{S}_H, \mu_H, \sigma_H$ and 2° $\mathbf{S}_y, \mu_y, \sigma_y$, respectively.

REFERENCES

Davis, H.T. (1933). *Tables of the higher mathematical functions.* Vol.1,2, Bloomington, In: Principia Press, 1935.

Janke, E., Emde, F. and Lösch, F. (1968). (Tafeln Höheres Funktionen). Spetsyalnye funktsii, formuly, grafiki, tablitsy, Moskva, 56.

Johnson, N.L. and Kotz, S. (1970). Continuous univariate distributions. Vol.2, J.Wiley, New York.

Johnson, N.L. and Kotz, S. (1992). Discrete distributions. Vol.1, J.Wiley, New York.

Oktaba, W. (1995a). Densities of determinant ratios, their moments and some simultaneous confidence intervals in the multivariate Gauss-Markoff model. *Applications of Mathematics*, Czechoslovak Academy of Sciences, Praha, **40,** No.1, 47-54.

Oktaba, W. (1995b). Asymptotically normal confidence intervals for a determinant in a generalized multivariate Gauss-Markoff model. *Applications of Mathematics*, Czechoslovak Academy of Sciences, Praha, **40**, No.1, 55-59.

Oktaba, W. and Kieloch, A. (1993). Wishart distributions in the Multivariate Gauss-Markoff model with singular covariance matrix. *Applications of Mathematics*, Czechoslovak Academy of Sciences, Praha, **38**, No.1, 61-66.

Olsen, A.C. (1938). Transformations of the Pearson Type III distributions. *Ann.Math.Statist.* **9**, 176-200.

Rao, C.R. (1973). Linear statistical inference and its applications. 2nd ed., J.Wiley, New York.

Received 10 August 1995; Revised 12 December 1995

Rozkłady asymptotyczne normalne w wielowymiarowym modelu Gaussa-Markowa

STRESZCZENIE

W wielowymiarowym modelu Gaussa-Markowa z macierzą kowariancji $\sigma^2\mathbf{\Sigma} \otimes \mathbf{V}$ podano jawne wzory na asymptotycznie normalne rozkłady trzech zmiennych losowych zawierających wyznaczniki macierzy sum kwadratów i iloczynów dla błędu, hipotezy i zmienności całkowitej oraz wyznacznik macierzy $\sigma^2\mathbf{\Sigma}$. Korzystając z tych rezultatów przedstawiono asymptotycznie normalne przedziały ufności dla $|\sigma^2\mathbf{\Sigma}|$ bazujące na trzech macierzach sum kwadratów i iloczynów.

SŁOWA KLUCZOWE: funkcje digamma i trigamma, wielowymiarowa funkcja gamma, stała Eulera, standardowy wielowymiarowy rozkład normalny, asymptotyczna normalność, wyznacznik, przedział ufności, rozkład Wisharta