

Two Piece Multinormal Distribution

Dr. Parag Shah, Dr. C. D. Bhavsar

ABSTRACT:

We have extended the univariate binormal distribution to two-piece multinormal distribution. The density function of two-piece multinormal distribution is given and the maximum likelihood estimators of the parameters when the variance - covariance matrix of two components of multivariate normal distributions are proportional to each other, are obtained. Its truncated distribution and estimation of parameters are also considered in this paper.

KEYWORDS: Two piece multinormal distribution, maximum likelihood estimators, log likelihood, variance-covariance matrix, truncated, ordered statistics and concomitants.

1. INTRODUCTION:

Multivariate analysis deals with observations on more than one variable where there is some inherent interdependence between the variables. The basic central distribution & building block in classical multivariate analysis is the multivariate normal distribution because (i) it is often the case that multivariate observations are at least approximately, normally distributed, and (ii) the multivariate normal distribution of the sampling distributions give rise to the tractable distribution. This is not generally the case with other multivariate distributions, even for ones which appear to be close to the normal.

The Statistical theory based on the normal distribution has the advantage that the multivariate methods based on it are extensively developed and can be studied in an organized and systematic way. This is due not only to the need for such methods because they are of practical use, but also to the fact that normal theory is amenable to exact mathematical treatment. The suitable methods of analysis are mainly based on standard operations of matrix algebra; the distributions of many statistics involved can be obtained exactly or characterized, and in many cases optimum properties of procedures can be deduced.

Most of the Statistical methods developed and evaluated in the context of the multivariate normal distribution, though many of the procedure are useful and effective when the distribution sampled is not normal. A major reason for basing Statistical analysis on the normal distribution is that this probabilistic model approximates well the distribution of continuous measurement in many sampled populations.

Dr. Parag Shah, Dept. of Statistics, H L College of Commerce, India, pbs1374@yahoo.co.in

Dr. C. D. Bhavsar, Dept. of Statistics, Gujarat University, India, chetna_bhavsar@yahoo.co.in

In fact, most of the methods and theory have been developed to serve statistical analysis of data. Although the bivariate normal distribution has been studied at the beginning of the nineteenth century, interest in multivariate distributions remained at a low level until it was stimulated by the work of Galton (1877) in the last quarter of the century. He did not, himself, introduce new forms of joint distribution, but he developed the idea of correlation and regression and focused attention on the need for greater knowledge of possible forms of multivariate distribution. Galton (1889) enunciated the theory of the multivariate normal distribution as a generalization of observed properties of samples.

For studying problems in genetics, biology, agricultural experiments, botany, anthropology, engineering, economic and in other fields, the multivariate normal distributions have been found to be sufficiently close approximations to the populations so that statistical analysis based on these models are justified.

The moment generating function of the Truncated Multinormal distribution was given by Tallis (1961). Sharples, J.J and Pezzey, J.C.V. (2007) have obtained the expectations of linear functions with respect to truncated multinormal distribution. He has also shown how it can be applied to uncertainty analysis in environmental modeling. The truncated multivariate normal distribution is a reasonable distribution for modeling many natural occurring random outcomes. The distribution of rainfalls in adjacent geographical areas follows truncated multivariate normal distribution. The usefulness of modeling rainfall in this way and the way in which it contributes to wheat yield uncertainty is illustrated in Griffiths et al (2001).

Similar to two-piece normal distribution, there are many advantages of the two-piece multivariate normal (TPMN) distribution over the other distributions those are used to handle asymmetry in data. One such advantage is that it

can handle a wide range of skewness both positive and negative for more than one correlated variables. John (1982), Kimber (1985), Nabar and Deshmukh (2001) have studied some important properties of two-piece normal distribution.

In section 2, we have defined two piece multinormal (TPMN) distribution. In section 3 we have given m.l.e's of the parameters. Truncated two piece multivariate normal (TTPMN) distribution is defined in section 4. It's m.l.e's are given in section 5 of this paper.

2. THE DENSITY FUNCTION

The density function of two piece multinormal (TPMN) distribution is given by

$$f(x) = \begin{cases} \left(\frac{2}{\pi}\right)^p (|V_1|^{1/2} + |V_2|^{1/2})^{-1} \exp\left\{-\frac{1}{2}(x-\underline{\mu})' V_1^{-1}(x-\underline{\mu})\right\}, x \leq \underline{\mu} \\ \left(\frac{2}{\pi}\right)^p (|V_1|^{1/2} + |V_2|^{1/2})^{-1} \exp\left\{-\frac{1}{2}(x-\underline{\mu})' V_2^{-1}(x-\underline{\mu})\right\}, x > \underline{\mu} \end{cases} \dots\dots (2.1)$$

If V_2 is proportional to V_1 i.e. $V_1 = V$ and $V_2 = kV_1, (k \neq 0)$ then the density function (2.1) reduces to

$$f(x) = \begin{cases} \left(\frac{2}{\pi}\right)^p (1+k^p)^{-1} |V|^{-1/2} \exp\left\{-\frac{1}{2}(x-\underline{\mu})' V^{-1}(x-\underline{\mu})\right\}, x \leq \underline{\mu} \\ \left(\frac{2}{\pi}\right)^p (1+k^p)^{-1} |V|^{-1/2} \exp\left\{-\frac{1}{2}(x-\underline{\mu})' (kV)^{-1}(x-\underline{\mu})\right\}, x > \underline{\mu} \end{cases} \dots\dots (2.2)$$

3 ESTIMATION OF PARAMETERS

According to the suggestion of Xinlei Wang et al (2006) we define concomitants of multivariate order statistics as follows

Let $(x_{h1}, x_{h2}, \dots, x_{hp-1}, x_{hp})_{h=1}^N$ be an i i d random sample from a p-variate distribution, where the random variables x_1, x_2, \dots, x_{p-1} are absolutely continuous. Denote the order of x_{h1} among $x_{11}, x_{21}, \dots, x_{N1}$ by $R_{h:N}^1$, denote the order of x_{h2} among $x_{12}, x_{22}, \dots, x_{N2}$ by $R_{h:N}^2$ and so on, denote the order of x_{hp-1} among $x_{1p-1}, x_{2p-1}, \dots, x_{Np-1}$ by $R_{h:N}^{p-1}$ and so on, denote the

order of x_{hp-1} among $x_{1p-1}, x_{2p-1}, \dots, x_{Np-1}$ by $R_{h:N}^{p-1}$.

We consider the random variable x_p given the ranks $R_{h:N}^1 = r_1, R_{h:N}^2 = r_2, \dots, R_{h:N}^{p-1} = r_{p-1}$ as the concomitant of the r_1^{th} order statistics of x_1, r_2^{th} order statistics of x_2 and so on r_{p-1}^{th} order statistics of x_{p-1} and is

denoted by $x_{p:h}[r_1, r_2, \dots, r_{p-1}, N]$. For simplicity, we ignore the subscripts N and h and denote concomitant as $x_{p(r_1, r_2, \dots, r_{p-1})}$. Let

$\underline{x}_{(1)} = \underline{x}_{p1}(r_1^1, r_2^1, \dots, r_{p-1}^1), \underline{x}_{(2)} = \underline{x}_{p2}(r_1^2, r_2^2, \dots, r_{p-1}^2), \dots, \underline{x}_{(n)} = \underline{x}_{pn}(r_1^n, r_2^n, \dots, r_{p-1}^n)$ be n concomitants of a multivariate order statistics from TPMN distribution. On the assumption that $\underline{x}_{(t)} < \underline{\sim} < \underline{x}_{(t+1)}$

$$= \underline{x}_{(pt)}(r_1^t, r_2^t, \dots, r_{p-1}^t) < \underline{\sim} < \underline{x}_{(p,t+1)}(r_1^{t+1}, r_2^{t+1}, \dots, r_{p-1}^{t+1})$$

For some $t (t = 1, 2, \dots, n-1)$, the likelihood function can be written as

$$\log L = \frac{np}{2} \log\left(\frac{2}{\pi}\right) - n \log(1+k^p) - \frac{n}{2} \log |V| - \frac{1}{2} \sum_1 (x-\underline{\mu})' V^{-1}(x-\underline{\mu}) - \frac{1}{2} \sum_2 (x-\underline{\mu})' (kV)^{-1}(x-\underline{\mu}) \dots\dots (3.1)$$

where $\sum_1 (\sum_2)$ denotes summation over all observations less than or equal to (greater than) $\underline{\mu}$.

Differentiating $\log L$ with respect to $\underline{\mu}$ and equating it to zero, we get

$$V^{-1} \sum_1 (x-\underline{\mu}) + k^{-1} V^{-1} \sum_2 (x-\underline{\mu}) = 0 \dots\dots (3.2)$$

$$\therefore k \sum_1 (x-\underline{\mu}) + \sum_2 (x-\underline{\mu}) = 0 \dots\dots (3.3)$$

Let $\hat{\underline{\mu}}$ denote the m.l.e of $\underline{\sim}$ and

$$G(\underline{\mu}) = k \sum_1 (x-\underline{\mu}) + \sum_2 (x-\underline{\mu}) \dots\dots (3.4)$$

We can observe that,

$$(1) \quad G(\underline{x}_{(1)}) > 0 \text{ and } G(\underline{x}_{(n)}) < 0.$$

Also $G(\underline{x}_{(t)}) > G(\underline{x}_{(t+1)})$ for $t = 1, 2, \dots, n-1$

$$(2) \quad \frac{\partial G(\underline{\mu})}{\partial \underline{\mu}} = -kt - (n-t) < 0 \text{ for } \underline{x}_{(t)} < \underline{\mu} < \underline{x}_{(t+1)}$$

As $G(\underline{\mu})$ is continuous, from (1) and (2) we conclude that

$G(\underline{\mu})$ is a decreasing function of $\underline{\mu}$ and there exists only one real vector $\hat{\underline{\mu}}$, between $\underline{x}_{(1)}$ and $\underline{x}_{(n)}$ at which $G(\underline{\mu}) = 0$. This implies that $\frac{\partial \log L}{\partial \underline{\mu}}$ is a decreasing

function of $\underline{\mu}$ and $\hat{\underline{\mu}}$ is the only solution of the likelihood equation. One can note that even though $\frac{\partial \log L}{\partial \underline{\mu}}$ is not

differentiable with respect to $\underline{\mu}$ at sample points, it is continuous.

As $\log L$ at $\underline{\mu} = \underline{x}_{(1)} - h$ is less than $\log L$ at $\underline{\mu} = \underline{x}_{(1)}$ and $\log L$ at $\underline{\mu} = \underline{x}_{(n)} + h$ is less than $\log L$ at $\underline{\mu} = \underline{x}_{(n)}$, the search for $\hat{\underline{\mu}}$ has to be restricted between $\underline{x}_{(1)}$ and $\underline{x}_{(n)}$.

We adopt the following procedure for obtaining $\hat{\underline{\mu}}$.

Sequentially, one can start with $t = 1$ and $\underline{\mu} = \underline{x}_{(1)}$, if $G(\underline{x}_{(1)}) > 0$, then take $t = 2$ with $\underline{\mu} = \underline{x}_{(2)}$ again if $G(\underline{x}_{(2)}) > 0$, then take $t = 3$ and so on. We can continue in this manner till $G(\underline{x}_{(t)}) < 0$ for some t say t_1 . Hence

$$\underline{x}_{(t_1-1)} < \hat{\underline{\mu}} < \underline{x}_{(t_1)}.$$

Using the value of $t = t_1 - 1$ in (3.3), $\hat{\underline{\mu}}$ the value of $\underline{\mu}$ can be obtained as

$$\hat{\underline{\mu}} = \frac{k \sum_1 \underline{x} + \sum_2 \underline{x}}{k t + (n-t)} \quad \dots\dots\dots (3.5)$$

Differentiating log likelihood function with respect to elements of V and equating it to zero, we get

$$\hat{V} = \frac{k \sum_1 (\underline{x} - \hat{\underline{\mu}})(\underline{x} - \hat{\underline{\mu}})' + \sum_2 (\underline{x} - \hat{\underline{\mu}})(\underline{x} - \hat{\underline{\mu}})'}{nk} \quad \dots\dots\dots (3.6)$$

4. THE DENSITY OF ITS TRUNCATED DISTRIBUTION

The density function of truncated two piece multinormal (TTPMN) distribution for $V_2 = kV_1$ is

$$f(\underline{x}) = \begin{cases} f_1(\underline{x}) = C_1^{-1} \left(\frac{\sqrt{2}}{\pi} \right)^p (1+k^p)^{-1} |V|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\underline{x} - \underline{\mu})' V^{-1} (\underline{x} - \underline{\mu}) \right\}, \underline{A} \leq \underline{x} \leq \underline{\mu} \\ f_2(\underline{x}) = C_2^{-1} \left(\frac{\sqrt{2}}{\pi} \right)^p (1+k^p)^{-1} |V|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\underline{x} - \underline{\mu})' (kV)^{-1} (\underline{x} - \underline{\mu}) \right\}, \underline{\mu} \leq \underline{x} < \underline{B} \end{cases} \quad \dots\dots\dots (4.1)$$

where $\underline{A}: p \times 1$ and $\underline{B}: p \times 1$ are vectors of truncation.

C_1, C_2 are constants due to truncation and are given as

$$C_1 = 2^{p+1} (1+k^p)^{-1} \prod_{i=1}^p \left(\frac{1}{2} - \Phi(A_i^*) \right)$$

$$C_2 = 2^{p+1} (1+k^p)^{-1} k^{p/2} \prod_{i=1}^p \left(\Phi \left(\frac{B_i^*}{\sqrt{k}} \right) - \frac{1}{2} \right) \quad \dots\dots\dots (4.2)$$

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2f}} e^{-\frac{1}{2}t^2 dt}$$

with

using (4.2), (4.1) can be rewritten as

$$f(\underline{x}) = \begin{cases} C_1^* |V|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\underline{x} - \underline{\mu})' V^{-1} (\underline{x} - \underline{\mu}) \right\}, \underline{A} \leq \underline{x} \leq \underline{\mu} \\ C_2^* |V|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2k} (\underline{x} - \underline{\mu})' V^{-1} (\underline{x} - \underline{\mu}) \right\}, \underline{\mu} \leq \underline{x} \leq \underline{B} \end{cases} \quad \dots\dots\dots (4.3)$$

where

$$C_1^* = \frac{1}{2(\sqrt{2f})^p \prod_{i=1}^p \left(\frac{1}{2} - \Phi(A_i^*) \right)} \quad \text{and}$$

$$C_2^* = \frac{1}{2(\sqrt{2f})^p \prod_{i=1}^p \left(\Phi \left(\frac{B_i^*}{\sqrt{k}} \right) - \frac{1}{2} \right)}$$

5. ESTIMATION OF PARAMETERS OF TTPMN DISTRIBUTION

Let $x_{(1)}, x_{(2)} \dots x_{(n)}$ be concomitants multivariate ordered sample from TTPMN distribution. On the assumption that $x_{(t)} \leq \tilde{x} \leq x_{(t+1)}$ for some $t, (t=1, \dots, n-1)$ the likelihood function can be written as

$$\begin{aligned} \text{Log}L &= t \log C_1^* + (n-t) \log C_2^* - \frac{n}{2} \log |V| \\ &\quad - \frac{1}{2} \sum_1 (x - \tilde{x})' V^{-1} (x - \tilde{x}) - \frac{1}{2k} \sum_2 (x - \tilde{x})' V^{-1} (x - \tilde{x}) \end{aligned} \dots\dots\dots (5.1)$$

where $\sum_1 (\sum_2)$ denotes summation over all observations lying between \underline{A} and \tilde{x} (lying between \tilde{x} and \underline{B}).

Differentiating logL with respect to \tilde{x} and equating it to zero, we get

$$\begin{aligned} V^{-1} \sum_1 (x - \tilde{x}) + k^{-1} V^{-1} \sum_2 (x - \tilde{x}) &= 0 \\ \therefore k \sum_1 (x - \tilde{x}) + \sum_2 (x - \tilde{x}) &= 0 \end{aligned} \dots\dots\dots (5.2)$$

Applying the same procedure as TPMN distribution, we can choose some $t = t_1$ such that $x_{(t_1-1)} < \hat{\tilde{x}} < x_{(t_1)}$ and similar to section 8.3, we get

$$\hat{\tilde{x}} = \frac{k \sum_1 x + \sum_2 x}{k t + (n - t)} \dots\dots\dots (5.3)$$

Differentiating loglikelihood function with respect to elements of V and equating it to zero, we get

$$\hat{V} = \frac{k \sum_1 (x - \hat{\tilde{x}})(x - \hat{\tilde{x}})' + \sum_2 (x - \hat{\tilde{x}})(x - \hat{\tilde{x}})'}{nk} \dots\dots\dots (5.4)$$

where $\sum_1 (\sum_2)$ denotes summation over all observations lying between \underline{A} and \tilde{x} (lying between \tilde{x} and \underline{B}) which are different from those in equation 3.5

REFERENCES :

[1]. Galton F. (1877) "Typical laws of Heredity in Man", Lecture at the Royal Institution of Great Britian.
 [2]. Galton, F. (1889) "Natural Inheritance", Mac Millan, London [1.2, 2.5]
 [3]. Griffiths, W.E., Newton, L.S. and O' Donell, C.J. (2001). "Predictive Densities for Shire level wheat yield in western Australia", Australian Agricultural and Resource Economic, Society Conference, Adelaide.
 [4]. John, S. (1982) "The parameter two piece normal family of distribution and its fitting", Communications in Statistic theory and Methods, 11, 8, 879-885.
 [5]. Kimber, A.C. (1985) "Methods for the two-piece normal distribution", Communications in Statistics: Theory and Methods, 14, 1, 235-245.
 [6]. Nabar, S.P. and Deshmukh, S.C. (2001). "On estimation of parameters of bilognormal distribution", JISA, 40, 1, 59-70.
 [7]. Sharples, J.J., Pezzey, J.C.V. (2007) "Expectations of linear functions with respect to truncated multinormal distribution-with applications for uncertainty analysis in environmental modeling", Environmental modeling and software, 22, 7, 915-923.
 [8]. Tallis, G.M. (1961) "The moment generating function of the truncated multi-normal distribution", Journal of Royal Statistical society, Series B (Methodological), 23, 1, 223-229.
 [9]. Xinlie Wang, Lynne Stokes, Johan Lim and Min Chew (2006) "Concomitants of Multivariate order statistics with applications to Judgment Post stratification", JASA, 101, 476, 1693-1704.