CRRAO Advanced Institute of Mathematics, Statistics and Computer Science (AIMSCS)

Research Report

Author (s): **B. L. S. Prakasa Rao**

Title of the Report: **Characterization of a Multivariate Normal Distribution from Samples of Random Size**

Research Report No.: **RR2013-13**

Date: **September 25, 2013**

Prof. C R Rao Road, University of Hyderabad Campus,
Gachibowli, Hyderabad-500046, INDIA.
www.crraoaimscs.org

Characterization of a Multivariate Normal Distribution from Samples of Random Size

B. L. S. PRAKASA RAO

CR Rao Advanced Institute of Mathematics, Statistics and Computer Science, Hyderabad 500046, India

Abstract: We obtain two characterizations of a multivariate normal distribution from samples of random size.

Key words : Characterization; Multivariate normal distribution; Samples of random size.

1 Introduction

Let $\mathbf{X}_i, 1 \leq i \leq N$ and $\mathbf{Y}_j, 1 \leq j \leq N$ be two independent samples of independent identically distributed k-dimensional random vectors with X_i distributed as F and Y_j distributed as G where N is a discrete integer valued random variable independent of $\mathbf{X}_i, 1 \leq i \leq N$ and $\mathbf{Y}_j, 1 \leq j \leq N$. Let

$$W = \sum_{j=1}^{N} [(\mathbf{a} - \mathbf{X}_j)'\Sigma^{-1}(\mathbf{a} - \mathbf{X}_j) + (\mathbf{b} - \mathbf{Y}_j)'\Sigma^{-1}(\mathbf{b} - \mathbf{Y}_j)]$$

where Σ is a known positive definite matrix for vectors $\mathbf{a}$ and $\mathbf{b}$ in R^k. Suppose that $E[e^{-\frac{1}{2}W}] = J(\mathbf{a}, \mathbf{b}) < \infty$. We prove that the function $J(\mathbf{a}, \mathbf{b}))$ is a measurable function of the function $\mathbf{a}'\Sigma^{-1}\mathbf{a} + \mathbf{b}'\Sigma^{-1}\mathbf{b}$ if and only if the distributions F and G are multivariate normal with mean zero vector and common covariance matrix $\sigma^2\Sigma$ for some constant $\sigma^2 > 1$. This result generalizes a similar result in the univariate case by Kotlarski and Cook (1977). Characterization problems of similar nature for identifiability in stochastic models are discussed in Prakasa Rao (1992).

2 Characterizations

We now state and prove the main results.

Theorem 2.1: Suppose that the function $J(\mathbf{a}, \mathbf{b})) = E[e^{-\frac{1}{2}W}] < \infty$ for all vectors $\mathbf{a}$ and $\mathbf{b}$

in R^k. Then the function $J(\mathbf{a},\mathbf{b}))$ is a measurable function of the function $\mathbf{a}'\Sigma^{-1}\mathbf{a}+\mathbf{b}'\Sigma^{-1}\mathbf{b}$ for $\mathbf{a},\mathbf{b}\in R^k$ if and only if the distributions F and G are multivariate normal with mean zero vector and common covariance matrix $\sigma^2\Sigma$ for some positive constant $\sigma^2>1$.

Proof : It is clear that

$$\begin{aligned}
E[e^{-\frac{1}{2}W}] &= \sum_{n=1}^{\infty} E[e^{-\frac{1}{2}W}|N=n]P(N=n)\\
&= \sum_{n=1}^{\infty}(E[\exp(-\frac{1}{2}(\mathbf{a}-\mathbf{X}_j)'\Sigma^{-1}(\mathbf{a}-\mathbf{X}_j))]E[\exp(-\frac{1}{2}(\mathbf{b}-\mathbf{Y}_j)'\Sigma^{-1}(\mathbf{b}-\mathbf{Y}_j))])^n P(N=n).
\end{aligned}$$

The last inequality follows from the assumption that $\mathbf{X}_i, 1\leq i\leq N$ and $\mathbf{Y}_j, 1\leq j\leq N$ are two independent samples of independent identically distributed k-dimensional random vectors independent of the random variable N. Let

$$\alpha(\mathbf{a}) = E[\exp(-\frac{1}{2}(\mathbf{a}-\mathbf{X}_j)'\Sigma^{-1}(\mathbf{a}-\mathbf{X}_j))]$$

and

$$\beta(\mathbf{b}) = E[\exp(-\frac{1}{2}(\mathbf{b}-\mathbf{Y}_j)'\Sigma^{-1}(\mathbf{b}-\mathbf{Y}_j))].$$

Then, it follows that,

$$E[e^{-\frac{1}{2}W}] = Q(\alpha(\mathbf{a})\beta(\mathbf{b}))$$

where

$$Q(x) = \sum_{n=1}^{\infty} x^n P(N=n) 0\leq x\leq 1.$$

Note that the function $Q(.)$ is a strictly increasing continuous function on the interval $[0,1]$. Hence its inverse is well defined. Suppose that the function $E[e^{-\frac{1}{2}W}]$ is a measurable function of the function $\mathbf{a}'\Sigma^{-1}\mathbf{a}+\mathbf{b}'\Sigma^{-1}\mathbf{b}$. Then there exists a measurable real-valued function $\psi(.)$ such that

$$\psi(\mathbf{a}'\Sigma^{-1}\mathbf{a}+\mathbf{b}'\Sigma^{-1}\mathbf{b}) = Q(\alpha(\mathbf{a})\beta(\mathbf{b})) \tag{2. 1}$$

or equivalently

$$\alpha(\mathbf{a})\beta(\mathbf{b}) = \gamma(\mathbf{a}'\Sigma^{-1}\mathbf{a}+\mathbf{b}'\Sigma^{-1}\mathbf{b}) \tag{2. 2}$$

where $\gamma = Q^{-1}o\psi$ for all $\mathbf{a},\mathbf{b}\in R^k$. It is easy to see that $\alpha(\mathbf{0})\neq 0$ and $\beta(\mathbf{0})\neq 0$ for $\mathbf{a}=\mathbf{0}$ and $\mathbf{b}=\mathbf{0}$. Substituting $\mathbf{a}=\mathbf{0}$ and $\mathbf{b}=\mathbf{0}$ alternately, we obtain that

$$\gamma(\mathbf{a}'\Sigma^{-1}\mathbf{a})\gamma(\mathbf{b}'\Sigma^{-1}\mathbf{b}) = \alpha(\mathbf{0})\beta(\mathbf{0})\gamma(\mathbf{a}'\Sigma^{-1}\mathbf{a}+\mathbf{b}'\Sigma^{-1}\mathbf{b}) \tag{2. 3}$$

for all $\mathbf{a}, \mathbf{b} \in R^k$. Let

$$\theta(t) = \frac{\gamma(t)}{\alpha(\mathbf{0})\beta(\mathbf{0})}, t \geq 0.$$

Note that the function $\theta(.)$ is measurable and the equation (2.3) implies that

$$\theta(\mathbf{a}'\Sigma^{-1}\mathbf{a})\theta(\mathbf{b}'\Sigma^{-1}\mathbf{b}) = \theta(\mathbf{a}'\Sigma^{-1}\mathbf{a} + \mathbf{b}'\Sigma^{-1}\mathbf{b}) \tag{2. 4}$$

for all $\mathbf{a}, \mathbf{b} \in R^k$. Hence the function $\theta(.)$ is a measurable function such that

$$\theta(t)\theta(s) = \theta(t+s) \tag{2. 5}$$

for all $t, s \geq 0$ since Σ^{-1} is a positive definite matrix. Therefore

$$\theta(t) = e^{c\,t}, t \geq 0 \tag{2. 6}$$

for some constant c. Hence

$$\gamma(t) = e^{c\,t}\alpha(\mathbf{0})\beta(\mathbf{0}), t \geq 0. \tag{2. 7}$$

Therefore, for any $\mathbf{a} \in R^k$,

$$\gamma(\mathbf{a}'\Sigma^{-1}\mathbf{a}) = e^{c\,\mathbf{a}'\Sigma^{-1}\mathbf{a}}\alpha(\mathbf{0})\beta(\mathbf{0}), \mathbf{a} \in R^k. \tag{2. 8}$$

Note that

$$\gamma(\mathbf{a}'\Sigma^{-1}\mathbf{a}) = \alpha(\mathbf{a})\beta(\mathbf{0}), \mathbf{a} \in R^k \tag{2. 9}$$

from (2.2). Combining the equations (2.8) and (2.9) and noting that $\beta(\mathbf{0}) \neq 0$, it follows that

$$\begin{aligned} e^{c\,\mathbf{a}'\Sigma^{-1}\mathbf{a}}\alpha(\mathbf{0}) &= \alpha(\mathbf{a}) \\ &= \int_{R^k} \exp[-\frac{1}{2}(\mathbf{a}-\mathbf{x})'\Sigma^{-1}(\mathbf{a}-\mathbf{x})]F(d\mathbf{x}). \end{aligned} \tag{2. 10}$$

This in turn gives the relation

$$\frac{\alpha(\mathbf{0})}{(2\pi)^{k/2}|\Sigma|^{1/2}}e^{c\,\mathbf{a}'\Sigma^{-1}\mathbf{a}} = \int_{R^k} \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp[-\frac{1}{2}(\mathbf{a}-\mathbf{x})'\Sigma^{-1}(\mathbf{a}-\mathbf{x})]F(d\mathbf{x}) \tag{2. 11}$$

for all $\mathbf{a} \in R^k$.The expression on the right side of the equation (2.11) is the convolution of a mutivariate normal density function with the distribution F. Hence the expression on the left side of the equation (2.11) also has to be a probability density function which implies that the constant $c = -\frac{1}{2\sigma^2}$ for some $\sigma^2 > 0$ and $\alpha(\mathbf{0})$ is a suitable normalizing constant. The characteristic functions of the probability densities on both sides of the equation (2.11), then, should satisfy the relation

$$\exp[-\frac{1}{2}(\mathbf{t}'\Sigma\mathbf{t})\sigma^2] = \exp[-\frac{1}{2}\mathbf{t}'\Sigma\mathbf{t}]\ \phi_X(\mathbf{t}), \mathbf{t} \in R^k \tag{2. 12}$$

where $\phi_{\mathbf{X}}$ is the characteristic function of the random vector $\mathbf{X}$. Hence

$$\phi_{\mathbf{X}}(\mathbf{t}) = \exp[-\frac{1}{2}(\mathbf{t}'(\sigma^2\Sigma - \Sigma)\mathbf{t})\sigma^2], \mathbf{t} \in R^k. \tag{2. 13}$$

Since $\phi_{\mathbf{X}}$ is the characteristic function of the random vector $\mathbf{X}$, it follows that $\sigma^2 > 1$ and the random vector $\mathbf{X}$ has the multivariate normal distribution with the mean vector zero and the covariance matrix $(\sigma^2 - 1)\Sigma$. Similar arguments prove that the random vector $\mathbf{Y}$ is also multivariate normal with mean vector zero and the covariance matrix $(\sigma^2 - 1)\Sigma$.

The converse part of the result stated in the theorem can be easily verified.

Suppose f and g are probability density functions on R^k. Let

$$Z = \Pi_{j=1}^{N} f(\mathbf{a} - \mathbf{X}_j) g(\mathbf{b} - \mathbf{Y}_j), \mathbf{a}, \mathbf{b} \in R^k.$$

Theorem 2.2: Suppose that the function $H(\mathbf{a}, \mathbf{b}) = E[Z] < \infty, \mathbf{a}, \mathbf{b} \in R^k$. Then the function $H(\mathbf{a}, \mathbf{b})$ is a measurable function of the function $\mathbf{a}'\Sigma^{-1}\mathbf{a} + \mathbf{b}'\Sigma^{-1}\mathbf{b}$ if and only if the distributions F and G are multivariate normal with mean vectors μ_F and μ_G and covariance matrices Σ_F and Σ_G respectively and the probability density functions f and g are multivariate normal probability density functions with mean vectors μ_f and μ_g and the covariance matrices Σ_f and Σ_g respectively with

$$\mu_F + \mu_f = \mu_G + \mu_g = \mathbf{0}$$

and

$$\Sigma_F + \Sigma_f = \Sigma_G + \Sigma_g = \sigma^2\Sigma$$

for some $\sigma^2 > 0$.

Proof : Let $\alpha(\mathbf{a}) = E[f(\mathbf{a} - \mathbf{X})]$ and $\beta(\mathbf{b}) = E[g(\mathbf{b} - \mathbf{Y})], \mathbf{a}, \mathbf{b} \in R^k$. It is easy to check that

$$\begin{aligned} E[Z] &= \sum_{n=1}^{\infty} [E(f(\mathbf{a} - \mathbf{X}))E(g(\mathbf{b} - \mathbf{Y}))]^n P(N = n) \\ &= Q(\alpha(\mathbf{a})\beta(\mathbf{b})) \quad \text{(say)}. \end{aligned} \tag{2. 14}$$

Suppose that $E(Z) = \psi(\mathbf{a}'\Sigma^{-1}\mathbf{a} + \mathbf{b}'\Sigma^{-1}\mathbf{b})$ for some function $\psi(.)$ Then

$$Q(\alpha(\mathbf{a})\beta(\mathbf{b})) = \psi(\mathbf{a}'\Sigma^{-1}\mathbf{a} + \mathbf{b}'\Sigma^{-1}\mathbf{b}), \mathbf{a}, \mathbf{b} \in R^k.$$

This relation is similar to that in equation (2.1). Arguments similar to those given earlier show that there exists a constant c such that

$$(2.\ 15) \qquad \alpha(\mathbf{0})\exp[c\,\mathbf{a}'\Sigma^{-1}\mathbf{a}] = \int_{R^k} f(\mathbf{a}-\mathbf{x})\ \ F(d\mathbf{x}), \mathbf{a}, \mathbf{b} \in R^k.$$

Note that the expression on the right side of the equation (2.15) is the convolution of the probability density function f with the distribution function F . Hence the function on the left side of the equation (2.15) has to be a probability density function which implies that $c = -\frac{1}{2\sigma^2}$ for some $\sigma^2 > 0$ and $\alpha(\mathbf{0})$ is a suitable normalizing constant for the multivariate normal density function with mean vector zero and the covariance matrix $\sigma^2\Sigma$. An application of the Cramer's theorem in R^k proves that f and F are multivariate normal probability density function and distribution function respectively with

$$\mu_f + \mu_F = \mathbf{0}$$

and

$$\Sigma_f + \Sigma_F = \sigma^2\Sigma.$$

Similar arguments show that g and G are also multivariate normal probability density function and distribution function respectively with

$$\mu_g + \mu_G = \mathbf{0}$$

and

$$\Sigma_f + \Sigma_G = \sigma^2\Sigma.$$

The converse part of the result in Theorem 2.2 can be established easily. We omit the details.

Acknowledgement : This work was supported under the scheme "Ramanujan Chair Professor" by grants from the Ministry of Statistics and Programme Implementation, Government of India (M 12012/15(170)/2008-SSD dated 8/9/09), the Government of Andhra Pradesh, India (6292/Plg.XVIII dated 17/1/08) and the Department of Science and Technology, Government of India (SR/S4/MS:516/07 dated 21/4/08) at the CR Rao Advanced Institute for Mathematics, Statistics and Computer Science, Hyderabad, India.

References

(1) Kotlarski, I.I. and Cook, W.L. (1977) Characterization of normality from samples of random size, Sankhya Series B, **39**, 196-200.

(2) Prakasa Rao, B.L.S. (1992) *Identifiability in Stochastic Models: Characterization of Probability Distributions*, Academic Press, Boston.