The Annals of Mathematical Statistics
1971, Vol. 42, No. 2, 824–827

A CHARACTERIZATION OF THE MULTIVARIATE NORMAL DISTRIBUTION[1]

BY MARY R. ANDERSON

Arizona State University

1. Summary. In this paper we characterize the multivariate normal distribution through the mean vector of the distribution. In Theorem 3, we find that the multivariate normal distribution can also be characterized through its variance, even though unknown, provided the variance–covariance matrix is free of the parameter involved. We are able to determine the mean vector to within two unknowns, the variance–covariance matrix and a constant vector.

2. Characterization through the mean vector. Let $\mathbf{X}' = (X_1, \cdots, X_n)$ be a random vector with mean $M'(\boldsymbol{\theta})$, $\boldsymbol{\theta}' = (\theta_1, \cdots, \theta_n)$, any real vector. We assume that $\mathbf{X}'$ has a positive definite variance–covariance matrix independent of $\boldsymbol{\theta}$.

In order to understand the nature of the main result more easily, we first state and prove a special case of the theorem for $n = 1$.

THEOREM 1. *If $E[X] = \theta$, then X is normal with positive variance free of θ if and only if X has a* pdf *of the form*

$$\exp\,[xb\theta + S(x) + Q(\theta)]$$

where b is a non-zero scalar.

PROOF. Suppose X has a pdf of the form $\exp\,[xb\theta + S(x) + Q(\theta)] = f_\theta(x)$. From the fact that $\int f_\theta(x)\,dx = 1$, by differentiating with respect to θ, we obtain

$$E(Xb) = \frac{-dQ(\theta)}{d\theta},$$

so that $\theta b = -dQ(\theta)/d\theta$, from which $Q(\theta) = -\frac{1}{2}b\theta^2 + c$, where c is independent of θ.

The moment-generating function of X exists and is easily verified to be

$$\begin{aligned} M_X(t) &= E[\exp(X(t)] \\ &= \exp(t\theta + \tfrac{1}{2}t^2/b) \int \exp\,[xb(\theta + t/b) + S(x) - \tfrac{1}{2}b(\theta + t/b)^2 + c]\,dx \\ &= \exp(t\theta + \tfrac{1}{2}t^2/b), \quad \text{for all} \quad t, \end{aligned}$$

which is normal with mean θ and variance $1/b$.

Now if X is normal with mean θ and variance v free of θ, it is obvious that the pdf of X is of the stated form.

We now state and prove the general theorem.

Received March 16, 1970; revised June 25, 1970.

[1] This paper is a portion of the author's doctoral thesis at the State University of Iowa.

THEOREM 2. *Let* $\mathbf{M}(\boldsymbol{\theta})$, $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_n)$ *be such that*

$$\mathbf{M}^* = \left[\frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_1}, \cdots, \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_n}\right]$$

is non-singular. If $E[\mathbf{X}] = \mathbf{M}(\boldsymbol{\theta})$, *then* $\mathbf{X}$ *is multivariate normal with positive definite variance–covariance matrix free of* $\boldsymbol{\theta}$, *if and only if* $\mathbf{X}$ *has a* pdf *of the form*

$$\exp[\mathbf{x}'\mathbf{B}\mathbf{M}(\boldsymbol{\theta}) + S(\mathbf{x}) + Q(\boldsymbol{\theta})],$$

with $\mathbf{B}$ *an* $n \times n$ *matrix free of* $\boldsymbol{\theta}$.

PROOF. Suppose $\mathbf{X}$ has the pdf $f_\theta(\mathbf{x})$ of the above form. From the fact that

$$\int_A \cdots \int f_\theta(\mathbf{x})\, d\mathbf{x}' = 1,$$

where A is the space of $\mathbf{X}'$, by taking the partial derivative with respect to θ_i, we obtain

$$E\left[\mathbf{X}'\mathbf{B}\frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_i}\right] = \frac{-\partial Q(\boldsymbol{\theta})}{\partial \theta_i},$$

so that

$$\mathbf{M}'(\boldsymbol{\theta})\mathbf{B}\frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_i} = \frac{-\partial Q(\boldsymbol{\theta})}{\partial \theta_i}, \qquad \text{for all } i. \tag{1}$$

Now,

$$\int_A \cdots \int \mathbf{x} f_\theta(\mathbf{x})\, d\mathbf{x}' = M(\theta),$$

so that by again taking the partial derivative with respect to θ_i, we obtain

$$E\left[\mathbf{X}\mathbf{X}'\mathbf{B}\frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_i}\right] = \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_i} + \mathbf{M}(\boldsymbol{\theta})\mathbf{M}'(\boldsymbol{\theta})\mathbf{B}\frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_i}. \tag{2}$$

Let $\mathbf{V}$ be such that $\mathbf{V} = E[\mathbf{X}\mathbf{X}'] - \mathbf{M}(\boldsymbol{\theta})\mathbf{M}'(\boldsymbol{\theta})$, the variance–covariance matrix of $\mathbf{X}'$, so that

$$\mathbf{V}\mathbf{B}\frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_i} = E\left[\mathbf{X}\mathbf{X}'\mathbf{B}\frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_i}\right] - \mathbf{M}(\boldsymbol{\theta})\mathbf{M}'(\boldsymbol{\theta})\mathbf{B}\frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_i}. \tag{3}$$

From (2) and (3), we have that

$$\frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_i} = \mathbf{V}\mathbf{B}\frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_i}, \qquad \text{for every } i.$$

Then

$$\left[\frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_1}, \cdots, \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_n}\right] = \mathbf{V}\mathbf{B}\left[\frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_1}, \cdots, \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_n}\right]$$

or

$$\mathbf{M}^* = \mathbf{V}\mathbf{B}\mathbf{M}^*.$$

Since $\mathbf{M}^*$ and $\mathbf{V}$ are non-singular, we have that $\mathbf{V} = \mathbf{B}^{-1}$ and hence $\mathbf{B}$ is non-singular and symmetric.

Since $\mathbf{B}$ is symmetric, we have from (1) that $Q(\boldsymbol{\theta}) = -\frac{1}{2}\mathbf{M}'(\boldsymbol{\theta})\mathbf{B}\mathbf{M}(\boldsymbol{\theta})+\mathbf{C}$, where $\mathbf{C}$ is independent of $\boldsymbol{\theta}$.

The moment-generating function of $\mathbf{X}$ exists and is

$$M_X(\mathbf{t}) = \int_{\mathbf{A}}\cdots\int \exp[\mathbf{x}'\mathbf{t}+\mathbf{x}'\mathbf{B}\mathbf{M}(\boldsymbol{\theta})+S(\mathbf{x})-\tfrac{1}{2}\mathbf{M}'(\boldsymbol{\theta})\mathbf{B}\mathbf{M}(\boldsymbol{\theta})+\mathbf{C}]\,d\mathbf{x}'$$

where $\mathbf{t}' = (t_1, \cdots, t_n)$. Since $\mathbf{B}$ is non-singular,

$$\begin{aligned} M_X(\mathbf{t}) &= \int_{\mathbf{A}}\cdots\int \exp\{\mathbf{x}'\mathbf{B}[\mathbf{M}(\boldsymbol{\theta})+\mathbf{B}^{-1}\mathbf{t}]+S(\mathbf{x})-\tfrac{1}{2}\mathbf{M}'(\boldsymbol{\theta})\mathbf{B}\mathbf{M}(\boldsymbol{\theta})\}\,d\mathbf{x}' \\ &= \exp\{-\tfrac{1}{2}\mathbf{M}'(\boldsymbol{\theta})\mathbf{B}\mathbf{M}(\boldsymbol{\theta})+\tfrac{1}{2}[\mathbf{M}(\boldsymbol{\theta})+\mathbf{B}^{-1}\mathbf{t}]'\mathbf{B}[\mathbf{M}(\boldsymbol{\theta})+\mathbf{B}^{-1}\mathbf{t}]\}\cdot \\ &\quad \int_{\mathbf{A}}\cdots\int \exp\{\mathbf{x}'\mathbf{B}[\mathbf{M}(\boldsymbol{\theta})+\mathbf{B}^{-1}\mathbf{t}]+S(\mathbf{x})-\tfrac{1}{2}[\mathbf{M}(\boldsymbol{\theta})+\mathbf{B}^{-1}\mathbf{t}]'\cdot \\ &\quad \mathbf{B}[\mathbf{M}(\boldsymbol{\theta})+\mathbf{B}^{-1}\mathbf{t}]+\mathbf{C}\}\,d\mathbf{x}'. \end{aligned}$$

But, the function under the integral sign has the form of a pdf and therefore its integral is equal to 1. Then

$$M_X(\mathbf{t}) = \exp[\mathbf{t}'\mathbf{M}(\boldsymbol{\theta})+\tfrac{1}{2}\mathbf{t}'\mathbf{B}^{-1}\mathbf{t}], \qquad \text{for all } \mathbf{t},$$

since $\mathbf{B}'^{-1} = \mathbf{B}^{-1}$. We have shown that $\mathbf{X}$ is multivariate normal with mean $\mathbf{M}(\boldsymbol{\theta})$ and variance–covariance matrix $\mathbf{B}^{-1}$.

Again, if $\mathbf{X}$ is multivariate normal with mean $\mathbf{M}(\boldsymbol{\theta})$ and variance–covariance matrix $\mathbf{V}$ free of $\boldsymbol{\theta}$, it is obvious that the pdf is of the stated form and the proof is complete.

3. Characterization through the variance.

THEOREM 3. *If the variance–covariance matrix of* $\mathbf{X}$ *is not a function of* $\boldsymbol{\theta}$, *then* $\mathbf{X}$ *is multivariate normal if and only if* $\mathbf{X}$ *has a* pdf *of the form*

$$\exp[\mathbf{x}'\mathbf{R}(\boldsymbol{\theta})+S(\mathbf{x})+Q(\boldsymbol{\theta})].$$

PROOF. Let $\mathbf{X}$ have the pdf $\exp[\mathbf{x}'\mathbf{R}(\boldsymbol{\theta})+S(\mathbf{x})+Q(\boldsymbol{\theta})] = \mathbf{g}_\theta(\mathbf{x})$. Since

$$\int_{\mathbf{A}}\cdots,\int \mathbf{x}g_\theta(\mathbf{x})\,d\mathbf{x}' = E[\mathbf{X}],$$

by taking the partial derivative with respect to θ_i, we obtain

$$E\left[\mathbf{X}\mathbf{X}'\frac{\partial\mathbf{R}(\boldsymbol{\theta})}{\partial\theta_i}\right] = \frac{\partial E[\mathbf{X}]}{\partial\theta_i} - E[\mathbf{X}]\frac{\partial Q(\boldsymbol{\theta})}{\partial\theta_i}. \tag{4}$$

However,

$$E\left[\mathbf{X}'\frac{\partial R(\boldsymbol{\theta})}{\partial\theta_i}\right] = \frac{-\partial Q(\boldsymbol{\theta})}{\partial\theta_i}$$

as established in the proof of Theorem 2, so that

$$E[\mathbf{X}]E[\mathbf{X}'] \frac{\partial R(\boldsymbol{\theta})}{\partial \theta_i} = -E[\mathbf{X}] \frac{\partial Q(\theta)}{\partial \theta_i}. \tag{5}$$

If we combine (4) and (5), we have that

$$\mathbf{V} \frac{\partial \mathbf{R}(\boldsymbol{\theta})}{\partial \theta_i} = \frac{\partial E[\mathbf{X}]}{\partial \theta_i}, \qquad \text{for all } i,$$

where $\mathbf{V}$ is the unknown variance–covariance matrix of $\mathbf{X}$. Since $\mathbf{V}$ is independent of $\boldsymbol{\theta}$, we have that $E[\mathbf{X}] = \mathbf{V}\mathbf{R}(\boldsymbol{\theta}) + \mathbf{C}$, $\mathbf{C}$ not a function of $\boldsymbol{\theta}$, or $R(\boldsymbol{\theta}) = \mathbf{V}^{-1}E[\mathbf{X}] - \mathbf{C}$.

We have now reduced the theorem to Theorem 2 and the result follows.

The converse is obvious.

COROLLARY. *If* $\mathbf{X}$ *has a variance–covariance matrix* $\mathbf{B}^{-1}$, *free of* $\boldsymbol{\theta}$, *and a* pdf *of the form* $\exp[\mathbf{x}'\mathbf{B}\mathbf{M}(\boldsymbol{\theta}) + S(\mathbf{x}) + Q(\boldsymbol{\theta})]$, *then* $\mathbf{X}$ *is multivariate normal with mean* $\mathbf{M}(\boldsymbol{\theta}) + \mathbf{C}$.

PROOF. Here $\mathbf{R}(\boldsymbol{\theta}) = \mathbf{B}\mathbf{M}(\boldsymbol{\theta})$ and thus $\mathbf{X}$ is multivariate normal with $E[\mathbf{X}] = \mathbf{V}\mathbf{B}\mathbf{M}(\boldsymbol{\theta}) + \mathbf{C}$ from Theorem 3. Since $\mathbf{V} = \mathbf{B}^{-1}$, the result follows.

Acknowledgment. I wish to express my appreciation to my thesis advisors Professors Allen T. Craig and Robert V. Hogg for their guidance and advice and for suggesting this problem and to the referees for their many helpful suggestions which made possible a considerable shortening of the main proof and a more general form of the last theorem.

REFERENCE

[1] HOGG, R. V. and CRAIG, A. T. (1959). *Introduction to Mathematical Statistics*. Macmillan, New York.