[3] S. N. Roy, "A Report on some Aspects of Multivariate Analysis," North Carolina Institute of Statistics Mimeograph Series No. 121.
[4] H. W. Turnbull and A. C. Aitken, *An Introduction to the Theory of Canonical Matrices*, Blackie and Son, Limited, London, 1952.
[5] K. D. Tocher, "The Design and Analysis of Block Experiments," JRSS Series B, Vol. 14 (1952), pp. 45–84.

A CHARACTERIZATION OF THE NORMAL DISTRIBUTION[1]

By J. N. K. Rao

Forest Research Institute, Dehra Dun, India

1. Introduction. Using characteristic functions Lukacs [3] has shown that a necessary and sufficient condition for the independence of the sample mean and variance is that the parent population be normal. Geisser [2] has derived a similar theorem concerning the sample mean and the first order mean square successive difference. In section 2 of this note a general theorem of which Lukacs' and Geisser's results are particular cases has been proved.

Lukacs [3] has extended his theorem to the multivariate case, namely, that a necessary and sufficient condition that the sample mean vector is distributed independently of the variance-covariance matrix is that the parent population be multivariate normal. In section 3, the general theorem of section 2 is extended to the multivariate population of which Lukacs' theorem for the multivariate population is a particular case. To prove the necessity of this theorem, we extend, to the multivariate case, Daly's [1] result that if $f(x)$ is the normal density, then the sample mean and $g(x_1 \cdots x_n)$ are independently distributed where $g(x_1 \cdots x_n) = g(x_1 + a, \cdots, x_n + a)$.

2. Univariate case. Let $x_1, \cdots, x_n$ be independent and identically distributed with density function $f(x)$ and mean μ and variance σ^2.

Let,

$$\bar{x} = n^{-1} \sum_{j=1}^{n} x_j \cdots \tag{2.1}$$

and

$$\delta^2 = \left(\sum_{t=1}^{m} \sum_{j=1}^{n} l_{tj}^2\right)^{-1} \sum_{t=1}^{m} (l_{t1} x_1 + \cdots + l_{tn} x_n)^2, \qquad m \geqq 1 \tag{2.2}$$

where

$$\sum_{j=1}^{n} l_{tj} = 0 \quad \text{for} \quad t = 1, \cdots, m.$$

The following theorem is proved.

Received April 8, 1957; revised January 17, 1958.

[1] Supported by a Senior Research Training Scholarship from the Government of India.

THEOREM 1. *A necessary and sufficient condition that $f(x)$ be the normal density is that $\bar{x}$ and δ^2 are independent.*

PROOF. Following Lukacs [3] we derive the sufficiency. Now,

$$E(\delta^2) = \left(\sum_{t=1}^{m}\sum_{j=1}^{n} l_{tj}^2\right)^{-1}\left\{\sum_{t}^{m}\sum_{j}^{n} l_{tj}^2 E(x_j^2) + \sum_{t=1}^{m}\sum_{j\neq j'}^{n} l_{tj}l_{tj'}E(x_j x_{j'})\right\}$$
$$= \sigma^2$$

The joint characteristic function of $\bar{x}$ and δ^2 is

$$\phi(t_1, t_2) = \int\int \cdots \int e^{it_1\bar{x}}e^{it_2\delta^2}f(x_1)\cdots f(x_n)\, dx_1\cdots dx_n.$$

Therefore

$$\frac{\partial}{\partial t_2}\phi(t_1, t_2)\,|_{t_2=0} = \phi_1(t_1)\frac{\partial}{\partial t_2}\phi_2(t_2)\,|_{t_2=0}, \tag{2.3}$$

where

$$\phi_1(t_1) = [\psi(t_1/n)]^n$$

and

$$\psi(t_1) = \int e^{it_1x}f(x)\, dx,$$

$$\begin{aligned}\frac{\partial}{\partial t_2}\phi(t_1, t_2)\,|_{t_2=0} &= i\,\left(\sum_t\sum_j l_{tj}^2\right)^{-1}\left\{\left(\sum_t\sum_j l_{tj}^2\right)[\psi(t_1/n)]^{n-1}\int x^2e^{it_1x/n}f(x)\, dx\right.\\ &\quad\left. + 2\left(\sum_t\sum_{j\neq j'} l_{tj}l_{tj'}\right)[\psi(t_1/n)]^{n-2}\left[\int xe^{it_1x/n}f(x)\, dx\right]^2\right\}\\ &= i\left\{[\psi(t_1/n)]^{n-1}\int x^2e^{it_1x/n}f(x)\, dx\right.\\ &\quad\left. - [\psi(t_1/n)]^{n-2}\left[\int xe^{it_1x/n}f(x)\, dx\right]^2\right\},\end{aligned} \tag{2.4}$$

and

$$\frac{\partial}{\partial t_2}\phi_2(t_2)\,|_{t_2=0} = i\sigma^2.$$

Hence, Eq. (2.3) reduces to

$$-\psi(t)\frac{d^2\psi(t)}{dt^2} + \left[\frac{d\psi(t)}{dt}\right]^2 = \sigma^2[\psi(t)]^2, \tag{2.5}$$

the solution of which is the characteristic function of the normal distribution.

The necessary condition follows from Daly [1] who has proved that $\bar{x}$ and $g(x_1 \cdots x_n)$ are independent in the normal case, if

$$g(x_1 \cdots x_n) = g(x_1 + a, \cdots, x_n + a).$$

Since δ^2 is invariant under a translation, the theorem is proved.

In fact, the above result can easily be extended[2] to cover a more general class of quadratic forms, namely those which are invariant under a translation and have non-zero expected values. For, Lukacs' method can be applied even when δ^2 is defined as follows:

$$\delta^2 = \left(\sum_{t=1}^{m}\sum_{j=1}^{n} a_{tij}\right)^{-1}\left[\sum_{t=1}^{m}\sum_{i,j=1}^{n}\sum a_{tij}x_ix_j\right], \qquad m \geqq 1, \tag{2.6}$$

where $\sum_{j=1}^{n} a_{tij} = 0$ ($t = 1, \cdots, m$, $i = 1, \cdots, n$), provided

$$\sum_{j=1}^{n} a_{tij} \neq 0 \; (t = 1, \cdots, m).$$

It will be noted that δ^2 defined in (2.2) above is a special case of δ^2 defined in (2.6) by putting $a_{tij} = l_{ti}l_{tj}$.

Particular Cases.

(a) To obtain Lukacs' result, put

$$l_{tj} = 1 - \frac{1}{n} \quad \text{for} \quad t = j$$
$$= \frac{-1}{n} \quad \text{for} \quad t \neq j$$
$$\text{and } m = n.$$

(b) To get Geisser's result, put

$$l_{tj} = 1 \quad \text{when} \quad j = t + k$$
$$= -1 \quad \text{when} \quad j = t$$
$$= 0 \text{ for other values of } j$$
$$\text{and } m = n - k.$$

(c) An interesting extension of Geisser's result is: a necessary and sufficient condition for the independence of the sample mean and *any* order mean square successive difference is that the parent population be normal.

The rth order mean square successive difference is given by,

$$\delta_r^2 = (n - r)^{-1}\left\{\binom{r}{0}^2 + \cdots + \binom{r}{r-1}^2 + \binom{r}{r}^2\right\}^{-1}\sum_{t=1}^{n-r}(\Delta^r x_t)^2, \qquad r \geqq 1.$$

where

$$\Delta^r x_t = \binom{r}{0}x_{t+r} - \binom{r}{1}x_{t+r-1} + \cdots + (-1)^r\binom{r}{r}x_t.$$

To get the above result, put

$$l_{tj} = (-1)^{t+r-j}\binom{r}{t+r-j} \quad \text{when} \quad t \leqq j \leqq t + r$$
$$= 0 \quad \text{when} \quad 1 \leqq j \leqq t - 1 \quad \text{and} \quad t + r + 1 \leqq j \leqq n,$$
$$\text{and} \quad m = n - r.$$

[2] I am indebted to the referee for pointing this out.

3. Multivariate case. The same reasoning applies also to the multivariate case. Denote by $x_{\alpha i}$ ($\alpha = 1, \cdots, n$; $i = 1, \cdots, p$) the α observation on the ith variate, by $\bar{x}_i$, the sample mean of the ith variate,

$$\delta_{ij} = \left[\left(\sum_{t=1}^{m}\sum_{\alpha=1}^{n} l_{t\alpha}^2\right)\right]^{-1} \sum_{t=1}^{m}\left\{\sum_{\alpha,\alpha'}^{n} l_{t\alpha}\, l_{t\alpha'}\, x_{\alpha i}\, x_{\alpha' j}\right\}, \tag{3.1}$$

or more generally,

$$\delta_{ij} = \left[\left(\sum_{t=1}^{m}\sum_{\alpha=1}^{n} a_{t\alpha\alpha}\right)\right]^{-1} \sum_{t=1}^{m}\left\{\sum_{\alpha,\alpha'}^{n} a_{t\alpha\alpha'}\, x_{\alpha i}\, x_{\alpha' j}\right\} \qquad (i, j = 1, \cdots, p), \tag{3.2}$$

where $\sum_{\alpha'=1}^{n} a_{t\alpha\alpha'} = 0$ ($t = 1, \cdots, m$; $\alpha = 1, \cdots, n$), provided

$$\sum_{\alpha=1}^{n} a_{t\alpha\alpha} \neq 0 \ (t = 1, \cdots, m).$$

Assuming that the distribution of $[\delta_{ij}]_{p \times p}$ is independent of the joint distribution of the p sample means $(\bar{x}_1, \cdots, \bar{x}_p)$ one obtains the equation,

$$\frac{\psi_{ij}}{\psi} - \frac{\psi_i \psi_j}{\psi^2} = -\lambda_{ij}, \tag{3.3}$$

where λ_{ij} is population covariance of the variates x_i and x_j,

$$\psi = \psi(t_1, \cdots, t_p) = \int\int \cdots \int e^{\iota(t_1 x_1 + \cdots + t_p x_p)} f(x_1 \cdots x_p)\, dx_1 \cdots dx_p .$$

$$\psi_i = \frac{\partial \psi}{\partial t_i}, \qquad \psi_{ij} = \frac{\partial^2 \psi}{\partial t_i\, \partial t_j}$$

If (3.3) is true for $i, j = 1, \cdots, p$, one has a set of partial differential equations which leads to the characteristic function to the multivariate normal distribution.

To prove the necessity, we give an extension of Daly's [1] lemma of which it is a particular casc.

THEOREM 2. *Let* $g_l(x_{11}, \cdots, x_{n1}; \cdots; x_{1p}, \cdots, x_{np})$, $l = 1, \cdots, r$, *be functions of* $(x_{11}, \cdots, x_{n1}); \cdots, (x_{1p}, \cdots, x_{np})$ *and are such that*

$$g_l(x_{11} + a_1, \cdots, x_{n1} + a_1; \cdots; x_{1p} + a_p, \cdots, x_{np} + a_p) = g_l(x_{11}, \cdots, x_{n1}; \cdots; x_{1p}, \cdots, x_{np}).$$

The sample means $(\bar{x}_1, \cdots, \bar{x}_p)$ *are independently distributed of these r functions if* $f(x_1 \cdots x_p)$ *has a p-variate normal distribution.*

PROOF. The joint characteristic function is

$$\phi(t_1, \cdots, t_p; \xi_1, \cdots, \xi_r) = \frac{1}{(2\pi)^{np/2} \mid \lambda_{ij} \mid^{n/2}} \int \cdots \int \exp\left\{\iota \sum_{\alpha=1}^{n} t_i\, x_{\alpha i}/n\right\} \exp\left\{\iota \sum_{l=1}^{r} \xi_l\, g_l\right\} \exp\left\{-\frac{1}{2}\sum_{\alpha=1}^{n}\sum_{i,j=1}^{p} \lambda^{ij}\, x_{\alpha i}\, x_{\alpha j}\right\} \times \prod_{\alpha=1}^{n} [dx_{\alpha 1} \cdots dx_{\alpha p}],$$

where $(\iota)^2 = -1$.

Make the contragradient transformation

$$x_{\alpha i} = \sum_{j=1}^{p} c_{ij}\, y_{\alpha j}\,, \qquad t_i = \sum_{j=1}^{p} c_{ij}\, u_j \qquad\qquad i = 1, \cdots, p;\ \alpha = 1, \cdots, n.$$

Then,

$$\phi(t_1, \cdots, t_p; \xi_1, \cdots, \xi_r)$$

$$= \frac{1}{(2\pi)^{np/2}\, |\lambda_{ij}|^{n/2}} \int \cdots \int \exp\left\{\iota \sum_{\alpha=1}^{n} \sum_{i=1}^{p} u_i\, y_{\alpha i}/n\right\} \exp\left\{\iota \sum_{l=1}^{r} \xi_l\, g_l'\right\}$$

$$\exp\left\{-\frac{1}{2} \sum_{\alpha=1}^{n} \sum_{i=1}^{p} y_{\alpha i}^2/\rho_i\right\} \times \prod_{\alpha=1}^{n} [dy_{\alpha 1} \cdots dy_{\alpha p}],$$

where $\rho_1, \cdots, \rho_p$ are latent roots of the variance-covariance matrix and

$$g_l'(y_{11} + a_1, \cdots, y_{n1} + a_1; \cdots; y_{1p} + a_p, \cdots, y_{np} + a_p)$$

$$= g_l'(y_{11}, \cdots, y_{n1}; \cdots; y_{1p}, \cdots, y_{np}).$$

Put

$$\frac{y_{\alpha i}}{\sqrt{\rho_i}} - \frac{\iota u_i \sqrt{\rho_i}}{n} = Z_{\alpha i};$$

then

$$\phi(t_1, \cdots, t_p; \xi_1, \cdots, \xi_r)$$

$$= \frac{1}{(2\pi)^{np/2}} \int \cdots \int \exp\left\{-\frac{1}{2n} \sum_{i,j=1}^{p} \lambda_{ij}\, t_i\, t_j\right\} \exp\left\{\iota \sum_{l=1}^{r} \xi_l\, g_l''\right\}$$

$$\exp\left\{-\frac{1}{2} \sum_{\alpha=1}^{n} \sum_{i=1}^{p} Z_{\alpha i}^2\right\} \times \prod_{\alpha=1}^{n} [dZ_{\alpha 1} \cdots dZ_{\alpha p}],$$

where

$$g_l'' = g_l'(Z_{11}\sqrt{\rho_1}, \cdots, Z_{n1}\sqrt{\rho_1}; \cdots; Z_{1p}\sqrt{\rho_p}, \cdots, Z_{np}\sqrt{\rho_p})$$

and hence is a function of $(Z_{11}, \cdots, Z_{n1}); \cdots; (Z_{1p}, \cdots, Z_{np})$ only. Therefore,

$$\phi(t_1, \cdots, t_p; \xi_1, \cdots, \xi_r)$$

$$= \exp\left\{-\frac{1}{2n} \sum_{i,j=1}^{p} \lambda_{ij}\, t_i\, t_j\right\} \times (\text{a function of } \xi_1, \cdots, \xi_r \text{ only}).$$

Hence the theorem.

Particular case. The sample mean vector $(\bar{x}_1 \cdots \bar{x}_p)$ is independently distributed of products moments of any order if $f(x_1 \cdots x_p)$ has a p-variate normal density.

4. Acknowledgement. My thanks are due to Dr. K. R. Nair for his valuable suggestions and to the referee for his helpful comments.

REFERENCES

[1] J. F. DALY, "On the use of the sample range in an analogue of student's t-test," *Ann. Math. Stat.*, Vol. 17 (1946), pp. 71–74.

[2] S. GEISSER, "A note on the normal distribution," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 858–859.

[3] E. LUKACS, "A characterization of the normal distribution," *Ann. Math. Stat.*, Vol. 13 (1942), pp. 91–93.

A NOTE ON P.B.I.B. DESIGN MATRICES

BY W. A. THOMPSON, JR.

University of Delaware

Summary. The notation P.B.I.B. (m) will mean partially balanced incomplete block design with m associative classes.

It is found that the C matrix of a P.B.I.B (m) may be expressed as a linear function of $m + 1$ commutative and linearly independent matrices. The author feels that this decomposition may be of interest to those studying the properties of P.B.I.B. designs.

1. The C matrix of a P.B.I.B. design. The reader should review the definition of partially balanced designs, and the relations among the parameters. See, for example, Bose and Shimamoto [2], or Bose [1], or Connor and Clatworthy [3].

The matrix

$$C = (c_{ij}),$$

where

$$c_{ii} = r(1 - 1/k),$$
$$c_{ij} = -\lambda_{ij}/k, \qquad i \neq j$$

is of special interest in incomplete block design theory.

In the case of a P.B.I.B. (m), the C matrix may be written in a particular form. We may write

$$kC = r(k - 1)I - \sum_{i=1}^{m} \lambda_i B_i, \tag{1.1}$$

where $B_s = [b_{tj}^{(s)}]$ for $s = 1, \cdots, m$, where $b_{tt}^{(s)} = 0$ and $b_{tj}^{(s)} = 1$ or 0 according as the treatments t and j are or are not sth associates. Note that $I, B_1, B_2, \cdots, B_m$ form a linearly independent set of matrices since a one in the (i, j)th position of any of them implies a zero in the (i, j)th position of all the others. $b_{hj}^{(s)} b_{ht}^{(s)}$ equals 1 if treatment j and treatment t are both sth associates of treatment h, but equals 0 otherwise. If $j \neq t$ then $\sum_i b_{ij}^{(s)} b_{it}^{(s)}$ is the number of treatments which are sth associates of both treatments j and t. But if j and t

Received May 10, 1955; revised April 8, 1958.