# EFFECT OF LINEAR TRUNCATION ON A MULTINORMAL POPULATION[1]

By Z. W. Birnbaum[2]

*University of Washington*

**1. Introduction.** In admission to educational institutions, personnel selection, testing of materials, and other practical situations, the following mathematical model is frequently encountered: A $(k + l)$-dimensional random variable $(X_1, X_2, \cdots, X_k, Y_1, Y_2, \cdots, Y_l) = (\mathbf{X}, \mathbf{Y})$ is considered, with a joint probability-distribution assumed to be non-singular multi-normal. The $Y_1, Y_2, \cdots, Y_l$ are scores in admission tests, the $X_1, X_2, \cdots, X_k$ scores in achievement tests. The admission tests are administered to all individuals in the $(\mathbf{X}, \mathbf{Y})$ population to decide on admission or rejection, and (usually at some later time) the achievement tests are administered to those admitted. A set of weights $a_j \geq 0, j = 1, 2, \cdots, l$ is used to define a composite admission test score $U = \sum_{j=1}^{l} a_j Y_j$ and a "cutting score" $\tau$ is chosen so that an individual is admitted if $U \geq \tau$, and rejected if $U < \tau$. We will refer to this procedure as *linear truncation of* $(\mathbf{X}, \mathbf{Y})$ *in* $\mathbf{Y}$ *to the set* $U \geq \tau$.

A linear truncation in $\mathbf{Y}$ clearly will change the absolute distribution of $\mathbf{X}$, except in the case of independence. In this paper a study is made of the absolute distribution of $\mathbf{X}$ after linear truncation in $\mathbf{Y}$ in the case $k = 1$; in particular, the possibility is investigated of choosing the $a_j$ and $\tau$ in such a way that the distribution of $\mathbf{X}$ after truncation has certain desirable properties. The case $k > 1$ leads to a considerable diversity of problems which are being studied and, it is hoped, will be the subject of a separate paper.

Throughout this paper it will be assumed that all the parameters of $(\mathbf{X}, \mathbf{Y})$, that is the expectations, variances and covariances before truncation, are known. In practical situations it often happens that only the parameters of $Y_1, Y_2, \cdots, Y_l$ before truncation are known, while the first and second moments involving $X_1, X_2, \cdots, X_k$ are only known for the joint distribution after truncation. It can be shown [1] that in such situations the expectations, variances and covariances of $(\mathbf{X}, \mathbf{Y})$ before truncation can always be reconstructed if $(\mathbf{X}, \mathbf{Y})$ has a multinormal distribution.

In the simplest case $k = l = 1$ the probability-density of the original bi-normal random variable $(X, Y)$ may be, without loss of generality, assumed equal to

$$f(X, Y, \rho) = \frac{1}{2\pi\sqrt{1 - \rho^2}} e^{-(X^2 - 2\rho XY + Y^2)/2(1-\rho^2)}. \tag{1.1}$$

By truncating this distribution in $Y$ to the set $Y \geq \tau$ one obtains the probability-density

$$g(X, Y; \rho, \tau) = \begin{cases} \psi^{-1}(\tau) f(X, Y; \rho), & \text{for } Y \geq \tau, \\ 0, & \text{for } Y < \tau, \end{cases} \tag{1.2}$$

[1] Presented to the Institute of Mathematical Statistics on June 18, 1949.

[2] Research done under the sponsorship of the Office of Naval Research.