ON EXACT PROBABILITIES OF RANK ORDERS FOR TWO WIDELY SEPARATED NORMAL DISTRIBUTIONS[1]

BY ROY C. MILTON

University of Wisconsin

The purpose of this note is to give numerical evidence of the adequacy of the Hodges-Lehmann [1] asymptotic formula for the probability of rank orders from two widely separated normal distributions. A table is given of the probabilities of the most probable and second most probable rank orders from two normal distributions (with unit variances) with means differing by $D = 4$, 5, and 6 units.

Assume that random variables $X_1, \cdots, X_m(Y_1, \cdots, Y_n)$ are normally and independently distributed with mean $O(D)$ and variance 1. Let $\mathbf{U} = (U_1, \cdots, U_{m+n})$, $U_1 < \cdots < U_{m+n}$, denote the order statistics of the random variables $(X_1, \cdots, X_m, Y_1, \cdots, Y_n)$, and let $\mathbf{Z} = (Z_1, \cdots, Z_{m+n})$ denote a random vector of zeros and ones where the ith component Z_i is $O(1)$ if U_i is an $X(Y)$. Denote by $f(x \mid D)$ the normal density with mean D and variance 1. If $\mathbf{z} = (z_1, \cdots, z_{m+n})$ is a fixed vector of zeros and ones, then the probability of the rank order $\mathbf{z}$, $P_{m,n}(\mathbf{z} \mid D)$, is given by

$$P_{m,n}(\mathbf{z} \mid D) = m!n! \int \cdots \int_{\mathcal{R}} \prod_{i=1}^{m+n} f(t_i \mid D z_i)\, dt_i, \tag{1}$$

where the region of integration $\mathcal{R}$ is $-\infty < t_1 \leqq \cdots \leqq t_{m+n} < \infty$.

It is clear that $P_{m,n}(\mathbf{z}^0 \mid D) \to 1$ as $D \to \infty$ for $\mathbf{z}^0 = (0 \cdots 01 \cdots 1)$, and consequently that the probability of all other rank orders tends to zero. Hodges and Lehmann [1] have presented a method describing this tendency for $\mathbf{z} \neq \mathbf{z}^0$, as follows. The rank order

$$\mathbf{z} = (\underbrace{0\cdots0}_{r_0}\underbrace{1\cdots1}_{s_1}\underbrace{0\cdots0}_{r_1}\underbrace{1\cdots1}_{s_2}\underbrace{0\cdots0}_{r_2}\cdots\underbrace{1\cdots1}_{s_c}\underbrace{0\cdots0}_{r_c}\underbrace{1\cdots1}_{s_0})$$

is characterized by the number of variables in the successive groups, a set of integers $(r_0, s_1, r_1, \cdots, s_c, s_0)$ with $\sum r_i = m$, $\sum s_j = n$. Here $r_0 = 0$ if $z_1 = 1$ and $s_0 = 0$ if $z_{m+n} = 0$. The "graph" of a rank order $\mathbf{z}$ may be obtained by representing each 0 by a horizontal and each 1 by a vertical unit segment. Figure 1 illustrates the graph for $\mathbf{z} = (001101010011101)$ with $(r_0, s_1, \cdots, r_c, s_0) = (2, 2, 1, 1, 1, 1, 2, 2, 1, 1)$.

The segments r_0 and s_0 are desregarded and the lower convex hull of the graph is formed with $k + 1$ corner points $(A_0, B_0), (A_1, B_1), \cdots, (A_k, B_k)$. In Figure 1, $k = 2$ and the convex hull represented by the dotted line has three corner points: (2, 0), (6, 4), (7, 6). Hodges and Lehmann prove that, for $\mathbf{z} \neq \mathbf{z}^0$,

$$\lim_{D\to\infty} [P_{m,n}(\mathbf{z} \mid D)^{D^{-2}}] = \exp[-\tfrac{1}{2} \textstyle\sum_{i=1}^{k} a_i b_i/(a_i + b_i)] \tag{2}$$

Received 12 April 1967.

[1] Research supported by the Office of Naval Research under Contract NONR-710(31), NR 042-003 and by a grant from the Society of the Sigma Xi and RESA Research Fund.