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

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

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 $\mathbf{z} \neq \mathbf{z}^0$,

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

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.

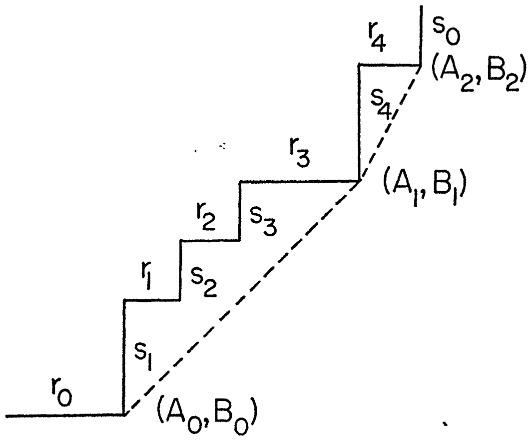

FIG. 1

TABLE 1

m	n	$P_{m,n}(\underset{\sim}{z}^0 \mid D)$ $[\underset{\sim}{z}^0 = (0\ldots01\ldots1)]$			$P_{m,n}(\underset{\sim}{z}^1 \mid D)$ $[\underset{\sim}{z}^1 = (0\ldots0101\ldots1)]$			$P_{m,n}(\underset{\sim}{z}^1 \mid D)^{1/D^2}$ Hodges-Lehmann value: .7788		
		D = 4	5	6	D = 4	5	6	D = 4	5	6
1	1	.99766113	.99979652	.99998895	.00233887	.00020348	.00001105	.6848	.7118	.7283
2	1	.99549652	.99959919	.99997803	.00432922	.00039466	.00002186	.7117	.7309	.7422
3	1	.99347006	.99940723	.99996720	.00607938	.00057588	.00003246	.7269	.7420	.7504
2	2	.99133986	.99921066	.99995628	.00797576	.00076486	.00004325	.7394	.7505	.7564
4	1	.99155761	.99922005	.99995649	.00764982	.00074871	.00004288	.7374	.7498	.7563
3	2	.98745787	.99883286	.99993476	.01115405	.00111521	.00006423	.7550	.7619	.7648
5	1	.98974182	.99903719	.99994586	.00907895	.00091431	.00005312	.7454	.7559	.7608
4	2	.98380234	.99846462	.99991344	.01398348	.00144887	.00008483	.7658	.7699	.7707
3	3	.98185533	.99827457	.99990264	.01554262	.00162491	.00009537	.7709	.7735	.7732
6	1	.98800962	.99885826	.99993533	.01039315	.00107356	.00006321	.7517	.7607	.7645
5	2	.98033865	.99810499	.99989230	.01653969	.00176816	.00010508	.7739	.7761	.7753
4	3	.97659023	.99773060	.99987082	.01942203	.00210972	.00012595	.7817	.7816	.7792
7	1	.98635083	.99868295	.99992488	.01161155	.00122717	.00007314	.7569	.7648	.7676
6	2	.97704073	.99775323	.99987135	.01887454	.00207485	.00012502	.7803	.7811	.7791
5	3	.97161075	.99719955	.99983929	.02290414	.00257309	.00015601	.7898	.7878	.7839
4	4	.96982514	.99701565	.99982861	.02419873	.00273755	.00016632	.7925	.7898	.7853
8	1	.98475726	.99851099	.99991451	.01274854	.00137571	.00008294	.7614	.7683	.7703
7	2	.97388823	.99740869	.99985057	.02102549	.00237033	.00014466	.7855	.7852	.7822
6	3	.96687784	.99668026	.99980802	.02606545	.00301767	.00018559	.7962	.7928	.7877
5	4	.96343809	.99631791	.99978677	.02846077	.00333694	.00020599	.8006	.7960	.7900
9	1	.98322223	.99834214	.99990422	.01381529	.00151965	.00009260	.7652	.7714	.7726
8	2	.97086486	.99707083	.99982995	.02302071	.00265573	.00016402	.7900	.7888	.7850
7	3	.96236108	.99617180	.99977702	.02896113	.00344556	.00021473	.8014	.7971	.7909
6	4	.95737714	.99563584	.99974530	.03230857	.00391144	.00024503	.8069	.8011	.7938
5	5	.95573461	.99545774	.99973474	.03339125	.00406546	.00025512	.8086	.8024	.7947
10	1	.98174016	.99817620	.99989401	.01482068	.00165938	.00010214	.7686	.7741	.7747
9	2	.96795727	.99673919	.99980949	.02488200	.00293202	.00018312	.7939	.7919	.7874
7	4	.95160183	.99496817	.99970417	.03581441	.00446382	.00028348	.8121	.8054	.7970
6	5	.94843554	.99461715	.99968315	.03781927	.00476302	.00030344	.8149	.8074	.7985
11	1	.98030635	.99801300	.99988387	.01577192	.00179526	.00011156	.7716	.7765	.7766
10	2	.96515438	.99641337	.99978918	.02662667	.00319997	.00020197	.7972	.7947	.7895
7	5	.94149041	.99379454	.99963199	.04183370	.00543312	.00035103	.8201	.8117	.8017
6	6	.93997547	.99362192	.99962154	.04274392	.00557768	.00036090	.8212	.8126	.8024
12	1	.97891677	.99785237	.99987379	.01667495	.00192757	.00012088	.7742	.7788	.7784
11	2	.96244679	.99609300	.99976901	.02826871	.00346027	.00022059	.8002	.7972	.7915
7	6	.93193653	.99264824	.99956045	.04718750	.00635958	.00041747	.8263	.8168	.8056
12	2	.95982651	.99577777	.99974898	.02981965	.00371350	.00023898	.8029	.7995	.7932
7	7	.92286990	.99152705	.99948950	.05199646	.00724805	.00048287	.8313	.8211	.8089

where $a_i = A_i - A_{i-1}$ and $b_i = B_i - B_{i-1}$, $(i = 1, \cdots, k)$. For the rank order shown in Figure 1, the right-hand side of (2) is $e^{-\frac{4}{3}} = .2636$. Using the table of $P_{m,n}(\mathbf{z} \mid D)$ given by Milton [2], the values of $P_{m,n}(\mathbf{z} \mid D)^{D^{-2}}$ for $D = 2$ and 3 are seen to be .0550 and .1341, respectively. A table of values of $P_{m,n}(\mathbf{z} \mid D)$ to 9 decimal places for all $\mathbf{z}$ for $1 \leqq n \leqq m \leqq 7$ and $n = 1$, $m = 8(1)12$; $D = 0(0.2)1, 1.5, 2, 3$ has been deposited in the UMT repository.

Table 1 gives values to 8 decimal places of $P_{m,n}(\mathbf{z} \mid D)$ for $\mathbf{z}^0 = (0 \cdots 01 \cdots 1)$ and $\mathbf{z}^1 = (0 \cdots 0101 \cdots 1)$; $1 \leqq n \leqq m \leqq 7$ and $n = 1$ and 2, $m = 8(1)12$; $D = 4, 5, 6$. These values were computed as described by Milton ([2], [3]) and may be considered an abbreviated extension of the table in [2], since for fixed m and n the probability of any other rank order $\mathbf{z}^i$ $(i \neq 0, 1)$ is less than both $1 - P_{m,n}(\mathbf{z}^0 \mid D) - P_{m,n}(\mathbf{z}^1 \mid D)$ and $P_{m,n}(\mathbf{z}^1 \mid D)$ ([4], Theorem 3). The tabled values give an indication of the rate at which $P_{m,n}(\mathbf{z}^0 \mid D) \rightarrow 1$ and $P_{m,n}(\mathbf{z}^1 \mid D) \rightarrow 0$. The last three columns of the table contain values of $P_{m,n}(\mathbf{z}^1 \mid D)^{D^{-2}}$. For sample sizes (m, n) as covered by the table, it is seen that, for example, $P_{m,n}(\mathbf{z}^1 \mid 6)^{1/36}$ ranges from .7283 to .8089, compared with the Hodges-Lehmann value $e^{-\frac{1}{4}} = .7788$ (which, it is interesting to note, is independent of m and n). The maximum departures above and below the Hodges-Lehmann value occur for the largest and smallest symmetric sample sizes tabled, namely $(m, n) = (7, 7)$ and $(1, 1)$, for all values of D.

REFERENCES

[1] Hodges, J. L., Jr. and Lehmann, E. L. (1962). Probabilities of ranking for two widely separated normal distributions, 146–151. *Studies in Mathematical Analysis and Related Topics* (Edited by Gilbarg, Solomon and others). Stanford Univ. Press.

[2] Milton, Roy C. (1965). Exact properties of rank order procedures under normal shift alternatives. Ph.D. thesis, Univ. of Minnesota.

[3] Milton, Roy C. (1967). Quadrature of high dimensional integrals: applications to nonparametric statistics. Submitted to *Math. Comp.*

[4] Savage, I. R., Sobel, M. and Woodworth, G. (1966). Fine structure of the ordering of probabilities of rank orders in the two sample case. *Ann. Math. Statist.* **37** 98–112.