FINE STRUCTURE OF THE ORDERING OF PROBABILITIES OF RANK ORDERS IN THE TWO SAMPLE CASE[1]

By I. Richard Savage, Milton Sobel and George Woodworth

Florida State University, University of Minnesota and University of Minnesota

1. Introduction. In constructing admissible two sample rank order tests one needs information on the ordering of probabilities of rank orders. Specifically, if, under some restriction of the class of alternatives, the rejection region of a test contains the rank order z then it should contain all rank orders more probable than z.

This paper contains several theorems on such orderings under various alternatives, especially the location parameter case for symmetric distributions.

2. Notation and assumptions. $X = (X_1, \cdots, X_m)$ and $Y = (Y_1, \cdots, Y_n)$ are samples drawn from absolutely continuous populations with densities $f(\cdot)$ and $g(\cdot)$, respectively. $F(\cdot)$ and $G(\cdot)$ denote the corresponding distributions. $W = (W_1, \cdots, W_{m+n})$ denotes the order statistics of the combined sample, $(X, Y) = (X_1, \cdots, X_m, Y_1, \cdots, Y_n)$, and $Z = (Z_1, \cdots, Z_{m+n})$ is a random vector of zeros and ones whose ith component, Z_i, is 0 if W_i comes from $f(\cdot)$ and 1 if W_i comes from $g(\cdot)$.

Let $z = (z_1, \cdots, z_{m+n})$ be a fixed vector of zeros and ones; we define the *complement of* z, $z^c = (z_1^c, \cdots, z_{m+n}^c)$ and the *transpose of* z, $z^t = (z_1^t, \cdots, z_{m+n}^t)$, to be the vectors whose ith components are $1 - z_i$ and $z_{m+n+1-i}$, respectively. $P(z) = \Pr\{Z = z\}$ denotes the probability of the rank order z.

Since the following restrictions of f and g are assumed in several results below, we list them now along with a shorthand notation.

Restrictions. ST: $f(x) = f(-x)$ and $g(x) = f(x - \theta)$, where θ is a non-negative constant.

U: $f(x) \geqq f(x')$ if $0 \leqq x < x'$ or $x' < x \leqq 0$.

MLR: $g(y)/f(y) \geqq g(x)/f(x)$ if $x \leqq y$.

N: $f(\cdot)$ and $g(\cdot)$ are normal densities with common variance 1 and means 0 and θ, respectively, where $\theta \geqq 0$.

Note. ST stands for *S*ymmetry and *T*ranslation and U implies that $f(\cdot)$ is *U*nimodal. It is assumed, without loss of generality, that the mode of $f(\cdot)$ is the origin. MLR stands for *M*onotone *L*ikelihood *R*atio and N stands for *N*ormality. Of course N is the strongest and implies the other three.

3. Theorems on the ordering of rank order probabilities. The general expression for $P(z)$ is

$$P(z) = m!\,n! \int \cdots \int \textstyle\prod_{i=1}^{m+n} h_{z_i}(t_i)\, dt_i, \tag{3.1}$$

Received 2 June 1965.

[1] Work supported by Nonr-710(31), NR 042 003 of the Office of Naval Research and Grant Number G-19126 of the National Science Foundation.

where

$$h_{z_i}(t_i) = f(t_i), \qquad z_i = 0,$$
$$= g(t_i), \qquad z_i = 1,$$

and the region of integration is $-\infty < t_1 \leqq \cdots \leqq t_{m+n} < \infty$. In particular, under ST

$$h_{z_i}(t_i) = f(t_i - \theta z_i). \tag{3.2}$$

THEOREM 1. *If* ST *holds, then for all* θ
(i) $P(z \mid \theta) = P(z^t \mid -\theta)$ *and*
(ii) $P(z \mid \theta) = P(z^c \mid -\theta)$.

PROOF. Recall the definition of z^t and z^c and note that $f(x) = f(-x)$. In the integral (3.1) (using (3.2)) make the transformation
(i′) $t_i = -t'_{m+n+1-i} \quad (i = 1, 2, \cdots, m + n)$ or
(ii′) $t_i = \theta + t_i' \quad (i = 1, 2, \cdots, m + n)$
and (i) or (ii) follows at once.

THEOREM 2. (Savage (1957) p. 975.) *If* ST *holds, then for all* θ, $P(z \mid \theta) = P(z^{tc} \mid \theta)$.

PROOF. Note that $z^{tc} = (z^t)^c$. Thus, by Theorem 1, $P(z^{tc} \mid \theta) = P(z^t \mid -\theta) = P(z \mid \theta)$.

THEOREM 3. (Savage (1956) p. 594.) *If* MLR *holds and* z *and* z' *differ only in their* ith *and* $i + 1$st *components with* $(z_i, z_{i+1}) = (0, 1)$ *while* $(z_i', z_{i+1}') = (1, 0)$, *then* $P(z) > P(z')$.

The following discussion and Corollary will give a method for moving from the partial ordering induced by Theorem 3 for sample sizes m and n to the corresponding partial ordering with sample sizes $m - 1$ and n or m and $n - 1$. Repeated use of the Corollary would allow one to work from any combination of sample sizes down to all smaller combinations of sample sizes. If z is a rank order with m 0's and n 1's, let $z_0(k)$ be formed by deleting the kth 0 in z where $1 \leqq k \leqq m$. As an example consider $z = (010011)$ then $z_0(2) = z_0(3) = (01011)$. One defines $z_1(k)$ in an analogous manner.

COROLLARY 1. *Under the conditions of Theorem* 3, *one obtains:* $P(z_0(k)) > P(z_0'(k))$ *provided* $z_0(k) \neq z_0'(k)$ *and* $P(z_1(k)) > P(z_1'(k))$ *provided* $z_1(k) \neq z_1'(k)$. *The partial orderings of the* $z_0(k)$ *and* $z_1(k)$ *are implied by Theorem* 3.

PROOF. z and z' are of the form: $z = (z^1 01 z^2)$, $z' = (z^1 10 z^2)$. Clearly $z_0(k)$ and $z'_0(k)$ are of one of the following forms: $z_0(k) = (z_0^1(k)01z^2)$, $z_0'(k) = (z_0^1(k)10z^2)$ or $z_0(k) = (z^1 1 z^2)$, $z_0'(k) = (z^1 1 z^2)$ or $z_0(k) = (z^1 01 z_0^2(k^*))$, $z_0'(k) = (z^1 10 z_0^2(k^*))$, where $k^* = k - 1 - \#$ of 0's in z^1. In any case the conclusion holds.

REMARK. If z and z' have a common number of zeros and ones and are such that $\sum_{j=1}^{i} (z_j' - z_j) \geqq 0$, for $i = 1, \cdots, m + n$, then there exist $z^1, z^2, \cdots, z^p$, where $z^k = (z_1^k, z_2^k, \cdots, z_{m+n}^k)$ for $k = 1, \cdots, p$ and $z^1 = z'$, $z^p = z$, such that for $k = 2, \cdots, p$, z^{k-1} and z^k differ in exactly two adjacent components, i_k and $i_k + 1$ with $(z_{i_k}^k, z_{i_k+1}^k) = (0, 1)$ and $(z_{i_k}^{k-1}, z_{i_k+1}^{k-1}) = (1, 0)$.

For example: $z = z^4 = (0010101)$, $z^3 = (0010110)$, $z^2 = (0011010)$, $z' = z^1 = (0101010)$. Therefore we have the following result:

COROLLARY 2. *If* MLR *holds and z and z' have the same number of zeros and ones and are such that $\sum_{j=1}^{i} (z_j' - z_j) \geqq 0$, for $i = 1, \cdots, m+n$, then $P(z) \geqq P(z')$.*

The properties of the orderings implied by Theorem 3 are discussed in detail in Savage (1964).

In succeeding pages we employ the notation (z, w), where $w = (w_1, \cdots, w_{p+q})$ and $z = (z_1, \cdots, z_{m+n})$, to denote the combined vector $(z_1, \cdots, z_{m+n}, w_1, \cdots, w_{p+q})$.

THEOREM 4. *If* ST *and* U *hold, $\theta > 0$, and z contains the same number, r, of zeros and ones, then*

(i) $P(001, z^{tc}) > P(z, 001)$ *and*

(ii) $P(100, z^{tc}) > P(z, 100)$.

PROOF. (i) By Theorem 1, a necessary and sufficient condition for the conclusion is that $P(z, 011) > P(z, 001)$, which, by (3.1) and (3.2) is equivalent to the inequality

$$(3.3) \quad I = \int_{-\infty}^{\infty} H(x)F(\theta - x)[f(x - \theta) - f(x)]\, dx > 0, \qquad \text{for all} \quad \theta > 0,$$

where

$$H(x) = (r + 2)!\,(r + 1)! \int \cdots \int_{R_{2r+1}} [\textstyle\prod_{i=1}^{2r} f(t_i - z_i\theta)\, dt_i] f(t_{2r+1})\, dt_{2r+1},$$

where $R_{2r+1} = \{t_i : -\infty < t_i \leqq \cdots \leqq t_{2r+1} \leqq x\}$.

We note here for future use that

$$\begin{aligned} H'(x) &= (d/dx)H(x) \\ (3.4) \qquad &= f(x)(r + 1)!\,(r + 2)! \int \cdots \int_{R_{2r}} \textstyle\prod_{i=1}^{2r} f(t_i - \theta z_i)\, dt_i \\ &= f(x)G(x), \quad \text{say,} \end{aligned}$$

where $R_{2r} = \{t_i : -\infty < t_1 \leqq \cdots \leqq t_{2r} \leqq x\}$.

Let $i(x)$ denote the integrand in (3.3) and let

$$I_1 = \int_{\theta/2}^{\infty} i(x)\, dx, \qquad I_2 = \int_{-\infty}^{\theta/2} i(x)\, dx.$$

In I_2 make the change of variable $x = \theta - x'$. This yields, after replacing x' by x in the transformed I_2 and adding I_1 and I_2,

$$I = \int_{\theta/2}^{\infty} [H(x)/F(x) - H(\theta - x)/F(\theta - x)] \cdot F(x)F(\theta - x)[f(x - \theta) - f(x)]\, dx.$$

It is easily seen that STU implies $[f(x - \theta) - f(x)] \geqq 0$ for $x \geqq \theta/2$. Therefore, a sufficient condition for I to be non-negative is that

$$H(x)/F(x) \geqq H(\theta - x)/F(\theta - x) \qquad \text{for} \quad x \geqq \theta/2.$$

Since $x \geqq \theta/2$ implies $x \geqq \theta - x$ it suffices to show that $H(x)/F(x)$ is non-decreasing for all x. And for this to be true it is sufficient that $H(x)/F(x)$ has a non-negative derivative, i.e., that $H'(x)F(x) - H(x)f(x) \geqq 0$ or, by (3.4) that $G(x)F(x) - H(x) \geqq 0$. Since $G(x)F(x) - H(x) = 0$ at $x = -\infty$ it suffices to show that $G(x)F(x) - H(x)$ is non-decreasing for all x, or that $G'(x)F(x) + f(x)G(x) - H'(x) \geqq 0$ or, by (3.4), that $G'(x) \geqq 0$, which is clearly so.

(ii) The proof is analogous to that of (i).

The proofs of the next three theorems have several features in common which we note here. They all state that if N holds then $P(z) > P(z')$. Equivalent conclusions are $P(z^{tc}) > P(z')$ and $P(z) > P(z'^{tc})$; one or the other is noted in each theorem and is in fact what is proved.

The first step is to replace each $P(\cdot)$ with its equivalent under (3.1) and (3.2) and to change the order of integration so that a particular pair of variables is integrated last. For the convenience of the reader this pair of variables will be indicated by adding primes (′) to the corresponding entries in the z-vectors the first time they appear.

At this point we have an inequality of the form $\int_{-\infty}^{\infty} \int_{x}^{\infty} S(x, y)\, dy\, dx > 0$ as a necessary and sufficient condition for the inequality $P(z) \geqq P(z')$. By making the transformation $y - x' = w$, $x = x'$ we obtain the equivalent inequality (omitting primes) $\int_{0}^{\infty} \int_{-\infty}^{\infty} S(x, x + w)\, dx\, dw > 0$.

A sufficient condition for this inequality is that the inner integral is non-negative for $w \geqq 0$, i.e., that

$$I(w) = \int_{-\infty}^{\infty} S(x, x + w)\, dx \geqq 0, \qquad \text{for} \quad w \geqq 0.$$

Let $I_1(w) = \int_{-\infty}^{(\theta-w)/2} S(x, x + w)\, dx$ and $I_2(w) = I(w) - I_1(w)$. In $I_1(w)$ make the transformation $x' = \theta - x - w$; this makes the ranges of integration of $I_1(w)$ and $I_2(w)$ coincide.

By adding $I_1(w)$ and $I_2(w)$, we obtain $I(w) = \int_{(\theta-w)/2}^{\infty} T(x, w)\, dx$, where $T(x, w) = [S(x, x + w) + S(\theta - x - w, \theta - x)]$.

In each case we show that $T(x, w) \geqq 0$ for $x \geqq (\theta - w)/2$ and $w \geqq 0$, which, of course, implies $I(w) \geqq 0$ for $w \geqq 0$.

In the proof of Theorem 5 we shall repeat in detail the argument just outlined. By Theorem 1 (i) it is necessary to consider only $\theta > 0$ in proving Theorems 5 and 6.

Theorem 5. *If* N *holds and* $\theta \neq 0$, *then* $P(100^r01) > P(010^r10)$ *or, equivalently,* $P(10'0^r0'1) > P(10'1^r0'1)$. ($x^r$ *denotes a vector of* r x*'s.*)

Proof. By (3.1) and (3.2), the inequality $P(100^r01) - P(101^r01) > 0$ is equivalent to the inequality

$$\int \cdots \int_{R_{r+4}} f(t_1 - \theta) f(t_2) f(t_{r+3}) f(t_{r+4} - \theta) [\textstyle\prod_{i=3}^{r+2} f(t_i) - \prod_{i=3}^{r+2} f(t_i - \theta)]\, dt_1 \cdots dt_{r+4} > 0,$$

where $R_{r+4} = \{t_i : -\infty < t_1 \leqq \cdots \leqq t_{r+4} < \infty\}$. Integration of the above with respect to all the variables but t_2 and t_{r+3} (call them x and y, respectively), yields the equivalent inequality

$$\int_{-\infty}^{\infty} \int_{x}^{\infty} F(x - \theta) F(\theta - y) \{[F(y) - F(x)]^r - [F(y - \theta) - F(x - \theta)]^r\} f(x) f(y)\, dy\, dx > 0.$$

If we transform the integral by letting $x = x'$ and $y = x' + w$ we get the inequality

$$(3.5)\quad \int_0^\infty \int_{-\infty}^\infty F(x-\theta)F(\theta-x-w)\{[F(x+w)-F(x)]^r - [F(x+w-\theta)-F(x-\theta)]^r\}f(x+w)f(x)\,dx\,dw > 0.$$

Let $I(w)$ denote the inner integral. If $I(w) \geqq 0$ for all $w \geqq 0$, then it is clearly so that $\int_0^\infty I(w)\,dw > 0$, therefore a sufficient condition for (3.5) (hence for the conclusion of the theorem) is that $I(w) \geqq 0$ for $w \geqq 0$.

Let $S(w, x)$ be the integrand in (3.5). Then

$$I(w) = \int_{(\theta-w)/2}^\infty S(w, x)\,dx + \int_{-\infty}^{(\theta-w)/2} S(w, x)\,dx.$$

In the second integral let $x' = \theta - x - w$; the result, after combining the two integrals, is

$$\begin{aligned} I(w) = \int_{(\theta-w)/2}^\infty & F(x)F(\theta-x-w)f(x+w)f(x-\theta) \\ & \cdot\{[F(x+w-\theta)-F(x-\theta)]^r - [F(x+w)-F(x)]^r\} \\ & \cdot[F(-x-w)f(x+w-\theta)/f(x+w)F(\theta-x-w) \\ & - F(x-\theta)f(x)/f(x-\theta)F(x)]\,dx. \end{aligned}$$

Clearly, a sufficient condition for $I(w)$ to be non-negative for $w \geqq 0$ is that the integrand above, call it $T(x, w)$, is non-negative for $w \geqq 0$ and $x \geqq (\theta - w)/2$.

The expression in braces in $T(x, w)$ is non-negative if and only if $[F(x + w - \theta) - F(x - \theta)] - [F(x + w) - F(x)] \geqq 0$, for $w \geqq 0$ and $x \geqq (\theta - w)/2$. This is clearly so since the left member of the inequality is the difference of the probability contents of two intervals one of which is more central than the other.

By the Corollary to Lemma 1 of Appendix I, the term in square brackets in $T(x, w)$ is non-negative for $x \geqq (\theta - w)/2$ and $w \geqq 0$. Therefore $T(x, w) \geqq 0$.

THEOREM 6. *If* N *holds and* $\theta \neq 0$, *then* $P(0^r01100^r) > P(0^r10010^r)$ *or, equivalently,* $P(1^r1'001'1^r) > P(0^r1'001'0^r)$.

PROOF. Proceeding as was outlined above, one obtains, as a sufficient condition for the inequality $P(1^r1'001'1^r) > P(0^r1'001'0^r)$, the inequality

$$\begin{aligned} T(x, w) = {} & [(2r+2)!/2(r!)]f(x-\theta)f(x+w-\theta)f(x)f(x+w)[G^2(x, w) \\ & - G^2(x-\theta, w)]\{[F(x-\theta)F(\theta-x-w)]^r - [F(x)F(-x-w)]^r\} \geqq 0, \end{aligned}$$

for $\theta \geqq 0$, $w \geqq 0$, and $x \geqq (\theta - w)/2$, where

$$(3.6)\quad G(x, w) = (2\pi)^{-\frac{1}{2}}[F(x+w) - F(x)]/f[(x+w)/2^{\frac{1}{2}}]f(x/2^{\frac{1}{2}}).$$

By the Corollary to Lemma 2 of Appendix I, with $r = w/2$, $y = x - r$, the term in square brackets (in the expression for $T(x, w)$) is non-negative for $\theta \geqq 0$, $w \geqq 0$, and $x \geqq (\theta - w)/2$.

Therefore $T(x, w)$ is non-negative provided the term in braces is non-negative. This term is non-negative if and only if

TABLE 1

$P(z)$ for selected values of θ under condition N *when $m = 4$ and $n = 2$ [from Klotz, (1962)]*

z	θ								
	.25	.50	1.0	1.5	2.0	3.0	4.0	5.0	6.0
000011	.10454	.15548	.29662	.47430	.65377	.90079	.98380	.99847	.999914
000101	.09313	.12192	.17290	.19240	.17020	.07084	.01398	.00145	$.0_4 85$
001001	.08400	.09871	.11104	.09580	.06412	.01411	.00132	$.0_4 57$	$.0_5 121$
000110	.07955	.08871	.09032	.07094	.04340	.00804	.00064	$.0_4 24$	$.0_6 46$
010001	.07521	.07903	.07080	.04837	.02547	.00341	.00019	$.0_5 47$	0
001010	.07170	.07158	.05712	.03397	.01510	.00127	$.0_4 35$†	$.0_6 32$†	0
100001	.06450	.05843	.03941	.02055	.00835	.00068	$.0_4 23$†	$.0_6 37$†	0
001100	.06433	.05781	.03771	.01853	.00688	.00041	$.0_5 8$	$.0_7 6$	0
010010	.06415	.05717	.03607	.01674	.00572	.00027	$.0_5 4$	$.0_7 1$	0
010100	.05753	.04606	.02357	.00892	.00250	$.0_4 79$*	$.0_6 75$	0	0
100010	.05500	.04219	.01992	.00698	.00181	$.0_4 48$	$.0_6 38$	0	0
011000	.05317	.03810	.01647	.00537	.00131	$.0_4 33$	$.0_6 26$	0	0
100100	.04929	.03392	.01291	.00366	.00077	$.0_4 13$	$.0_7 7$	0	0
101000	.04467	.02798	.00893	.00215	.00039	$.0_5 53$	$.0_7 1$	0	0
110000	.04022	.02290	.00620	.00130	.00021	$.0_5 23$	$.0_7 1$	0	0

* Subscripts on the first zero after the decimal indicate the number of zeros to be entered; for example, $.0_4 85$ stands for .000085.

† This appears to be a crossover but it is not clear that the eighth decimal is correct in this calculation.

$$[F(x-\theta)F(\theta-x-w) - F(x)F(-x-w)] \geqq 0, \qquad \text{for} \quad x \geqq (\theta - w)/2.$$

This inequality is proved in the Corollary to Lemma 3 of Section 5.

THEOREM 7. *If* N *holds and $\theta \neq 0$, then $P(0110, z) > P(1001, z)$ for any z or, equivalently, $P(z, 1'001') > P(z, 0'110')$.*

PROOF. Proceeding as was outlined above, one obtains as a sufficient condition for the inequality $P(z, 0110) < P(z, 1001)$, the inequality

$$T(x, w) = [H(x) - H(\theta - w - x)]\{G^2(x, w) - G^2(x - \theta, w)\} \cdot f(x)f(x+w)f(x-\theta)f(x+w-\theta) \geqq 0,$$

for $w \geqq 0$ and $x \geqq (\theta - w)/2$, where $G(x, w)$ is defined by (3.6) and

$$H(x) = [(m+2)!\,(n+2)!/2] \int \cdots \int_{R_{m+n}} \prod_{i=1}^{m+n} f(t_i - \theta z_i)\, dt_i,$$

where $R_{m+n} = \{t_i : -\infty < t_1 \leqq \cdots \leqq t_{m+n} < x\}$, m and n being the number of zeros and ones in z.

It is shown by the Corollary to Lemma 2 of Appendix I that the term in braces in $T(x, w)$ is non-negative for $w \geqq 0$ and $x \geqq (\theta - w)/2$.

Clearly, $H(x)$ is everywhere non-decreasing. Since $x \geqq (\theta - w)/2$ implies $x \geqq \theta - w - x$, we have $H(x) \geqq H(\theta - w - x)$, for $w \geqq 0, x \geqq (\theta - w)/2$. Thus, the term in square brackets is non-negative, and, therefore, so is $T(x, w)$.

$m = n = 2$

$$\begin{array}{c} 0011 \\ \Big\downarrow \text{MLR} \\ 0101 \\ \Big\downarrow \text{MLR} \\ 0110 \overset{\text{ST}}{\,|\,} 1001 \\ \Big\downarrow \text{MLR} \\ 1010 \\ \Big\downarrow \text{MLR} \\ 1100 \end{array}$$

$m = 3, n = 2$

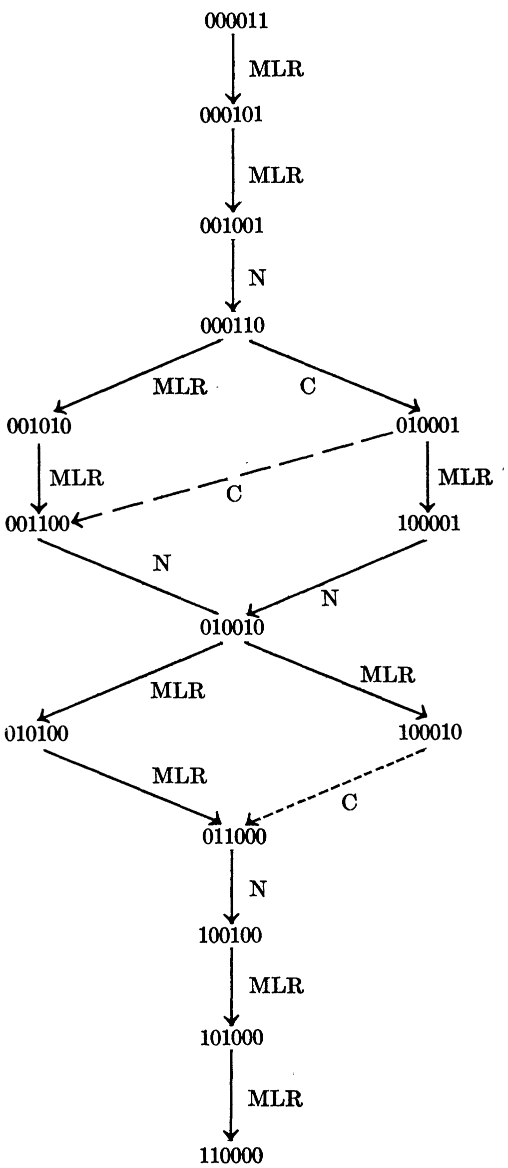
$$\begin{array}{c} 00011 \\ \Big\downarrow \text{MLR} \\ 00101 \\ \Big\downarrow \text{MLR} \\ 01001 \\ \Big\downarrow \text{N} \\ 00110 \\ \Big\downarrow \text{ST and U} \\ 10001 \\ \Big\downarrow \text{N} \\ 01010 \\ \Big\downarrow \text{MLR} \\ 01100 \\ \Big\downarrow \text{N} \\ 10010 \\ \Big\downarrow \text{MLR} \\ 10100 \\ \Big\downarrow \text{MLR} \\ 11000 \end{array}$$

4. Examples and conjectures. Theorems 2 through 7 are "independent," i.e., each of the theorems implies results not obtainable from any of the other theorems. The point is illustrated by the following list: Theorem 2 implies $P(1001) = P(0110)$, Theorem 3 implies $P(01010) > P(01100)$, Theorem 4 (i) implies $P(00110) > P(10001)$, Theorem 4 (ii) implies $P(10001) > P(01100)$, Theorem 5 implies $P(10001) > P(01010)$, Theorem 6 implies $P(001100) > P(010010)$, Theorem 7 implies $P(011000) > P(100100)$.

The following obvious conjecture is *not* true: If N and $P(z) > P(z')$ then $P(zz'') > P(z'z'')$. As a counter example consider: $z = 01001$, $z' = 00110$, and $z'' = 0$, then $P(z) > P(z')$ and $P(zz'') < P(z'z'')$.

The following diagrams illustrate the theorems of Section 3. The symbol $z \rightarrow_{abc} z'$ means $P(z) > P(z')$ for $\theta > 0$ (under ST, STU, or N) under conditions *abc*. C stands for *C*onjecture. The numerical results accompanying some of the diagrams are extracted from tables of probabilities of rank orders under N computed by Jerome Klotz (1962), (1964) and Roy C. Milton (1965).

$m = 4, n = 2$



	Ordering from Klotz (1962) and Milton (1965)
	000011
	000101
	001001
	000110
	010001
*	001010
*	100001
	001100
	010010
	010100
	100010
	011000
	100100
	101000
	110000

* Crossover.

NOTE. MLR implies a simple ordering of the $P(z)$ for $n = 1$. Hence the first interesting case is $m = n = 2$. All diagrams derived are distributive lattices [see Savage, 1964]. The conjectured diagram for $m = 4$ and $n = 2$ is not distributive;

$m = n = 3$

$$
\begin{tikzcd}
& 000111 \arrow[d, "\mathrm{MLR}"] & \\
& 001011 \arrow[d, "\mathrm{MLR}"] & \\
& 010011 \;\mathrm{ST}\; 001101 \arrow[dl, "\mathrm{MLR}"'] \arrow[dr, "\mathrm{MLR}"] & \\
001110 \;\mathrm{ST}\; 100011 \arrow[dr, dashed, "\mathrm{C}"] & & 010101 \arrow[dl, "\mathrm{MLR}"] \\
& 011001 \arrow[d, "\mathrm{N}"] & \\
& 100101 \;\mathrm{ST}\; 010110 \arrow[d, "\mathrm{MLR}"] & \\
& 101001 \;\mathrm{ST}\; 011010 \arrow[d, "\mathrm{N}"] & \\
& 100110 \arrow[dl, "\mathrm{MLR}"'] \arrow[dr, dashed, "\mathrm{C}"] & \\
101010 \arrow[dr, "\mathrm{MLR}"] & & 011100 \;\mathrm{ST}\; 110001 \arrow[dl, "\mathrm{MLR}"] \\
& 110010 \;\mathrm{ST}\; 101100 \arrow[d, "\mathrm{MLR}"] & \\
& 110100 \arrow[d, "\mathrm{MLR}"] & \\
& 111000 &
\end{tikzcd}
$$

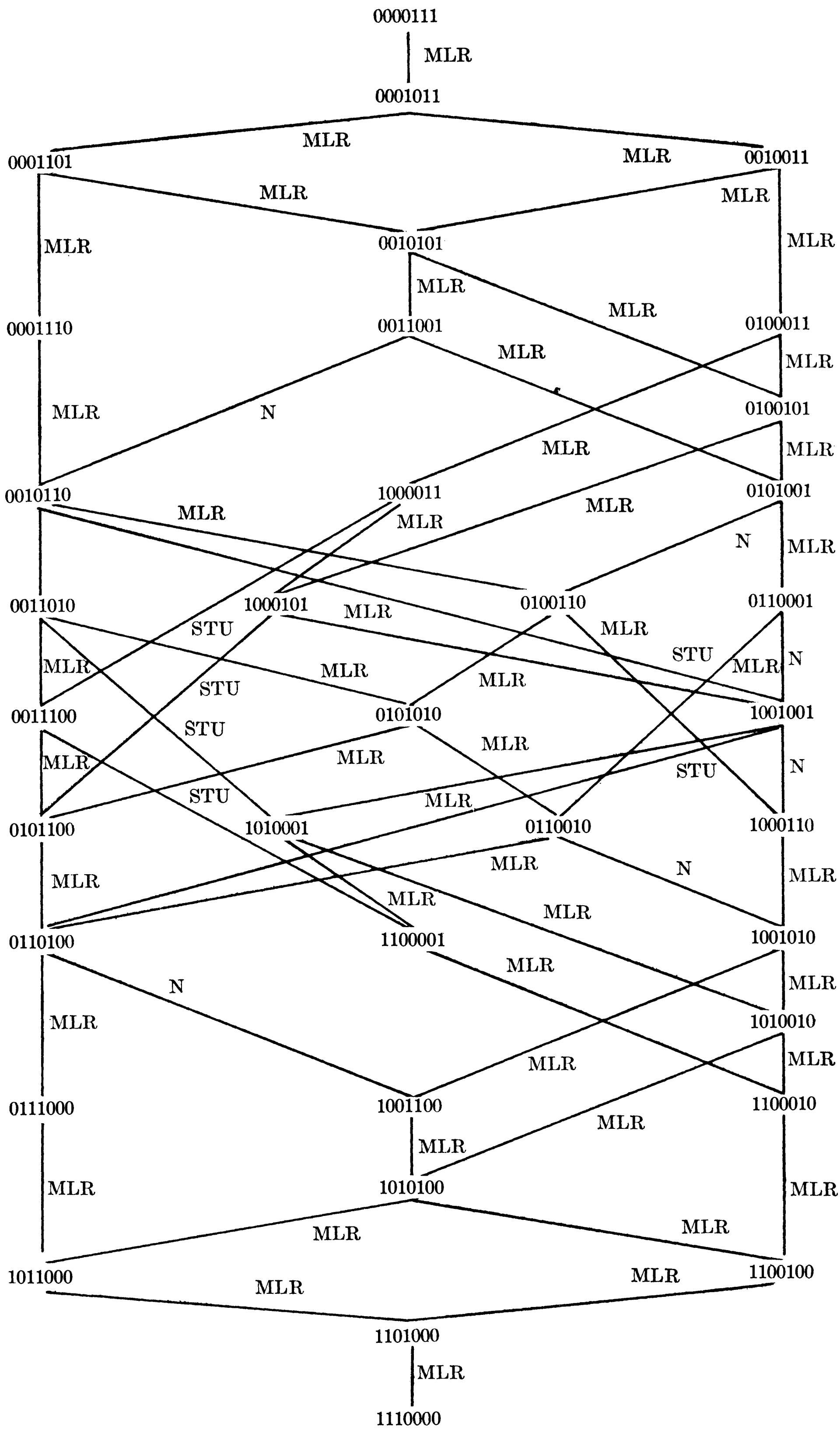
Ordering from Klotz (1964) and Milton (1965)
000111
001011
001101 = 010011
010101 $\}$*
100011 = 001110 $\}$*
011001
100101 = 010110
101001 = 011010
100110
110001 = 011100
101010
110010 = 101100
110100
111000

* Crossover.

in particular it does not satisfy the Jordan-Dedekind chain condition, i.e., not all chains from an arbitrary fixed z to (say) the least probable rank order are of the same length.

The notation, crossover $\begin{cases} z \\ z' \end{cases}$, denotes the existence of two values of $\theta, \theta_1 > \theta_2 > 0$, such that $P(z \mid \theta_1) > P(z' \mid \theta_1)$ but $P(z \mid \theta_2) < P(z' \mid \theta_2)$.

$m = 4, n = 3$


0000111
MLR
0001011
MLR
MLR
0001101
0010011
MLR
MLR
MLR
0010101
MLR
MLR
MLR
0001110
0011001
0100011
MLR
MLR
MLR
N
0100101
MLR
MLR
0010110
1000011
0101001
MLR
MLR
MLR
N
MLR
0011010
1000101
MLR
0100110
0110001
STU
MLR
STU
MLR
N
MLR
MLR
STU
MLR
0011100
0101010
1001001
STU
MLR
MLR
MLR
N
STU
STU
MLR
0101100
1010001
0110010
1000110
MLR
MLR
N
MLR
MLR
MLR
0110100
1100001
1001010
MLR
N
MLR
MLR
1010010
MLR
MLR
0111000
1001100
1100010
MLR
MLR
MLR
1010100
MLR
MLR
MLR
1011000
1100100
MLR
MLR
1101000
MLR
1110000

Ordering from Klotz (1962) and Milton (1965)

$m = 4, n = 3$

0000111	0101010
0001011	1000110 } *
0010011	1010001 } *
0001101	0110010
0100011	0101100
0010101 } *	* { 1001010 } *
0001110 } *	* { 1100001 } *
* { 0011001	* { 0110100
* { 0100101 } *	1010010 } *
* { 1000011 } *	1001100 } *
0010110	0111000 } *
0101001	1100010
0011010	1010100
1000101	1011000
0100110 } *	1100100
0110001 } *	1101000
0011100	1110000
1001001	

* Crossover.

The ordering from the numerical tables cannot be considered definitive since less than 10 values of θ have been used in the preparation of these tables. The diagram lines marked C, correspond to the conjecture that a rank order with a larger value of the Wilcoxon statistic is more probable than one with a smaller value. In tables prepared by Milton (1965), there are many cases for larger sample sizes where this conjecture is shown to be false. In fact if the indicated crossover in this situation is verified, the conjecture would already be false for $m = 4$ and $n = 2$.

For fixed m and n we define $N(z)$ and $N'(z)$ to be the number of rank orders less probable and more probable than z, respectively. The ideal situation for constructing tests of hypothesis is to have $N(z) + N'(z) = \binom{m+n}{n} - 1$, i.e., the rank orders form a chain. The Table 2 gives $N(z)$, $N'(z)$ and $N(z) + N'(z)$ for $m = 4$, $n = 3$ for the ordering implied by Theorem 3 alone and for the ordering implied by Theorems 1 through 7. Note that the second ordering is an improvement over the first in the sense that $N(z) + N'(z)$ for the second ordering is not smaller than that for the first. In particular there is considerable improvement in $z = (1000011), (0010110), (0110001)$ and (0011100).

APPENDIX I

Properties of some functions related to the normal density function. In the following $f(\cdot)$ and $F(\cdot)$ denote the standard normal density and distribution functions, respectively.

LEMMA 1. *If θ is a positive constant, then $G(x) = F(x - \theta)f(x)/F(x)f(x - \theta)$ is non-increasing for all x.*

TABLE 2*

z	MLR			MLR or STU or N		
	N(z)	N′(z)	N(z) + N′(z)	N(z)	N′(z)	N(z) + N′(z)
0000111	34	0	34	34	0	34
0001011	33	1	34	33	1	34
0010011	30	2	32	30	2	32
0001101	29	2	31	29	2	31
0100011	24	3	27	25	3	28
0010101	27	4	31	28	4	32
0001110	19	3	22	22	3	25
0011001	21	5	26	24	5	29
0100101	22	6	28	22	6	28
1000011	14	4	18	19	5	24
0010110	18	6	24	21	7	28
0101001	20	8	28	20	8	28
0011010	15	8	23	17	8	25
1000101	13	8	21	16	8	24
0100110	15	9	24	15	11	26
0110001	12	9	21	16	9	25
0011100	9	9	18	11	11	22
1001001	11	11	22	12	12	24
0101010	13	13	26	13	13	26

* To complete the table note that N(z^t) = N′(z).

Proof. A sufficient condition for the monotonicity of $G(x)$ is that its derivative is non-positive for all x, or, equivalently, that

$$f(x - \theta)/F(x - \theta) - f(x)/F(x) \leqq \theta$$

for all x. It follows from (3) of Sampford (1953) that the derivative of $f(x)/F(x)$ is in $(-1, 0)$; hence the result follows from the mean value theorem.

Corollary. *If w and θ are non-negative constants, then $G(\theta - x - w) \geqq G(x)$ for $x \geqq (\theta - w)/2$.*

Proof. The conclusion follows at once from the fact that $\theta - x - w \leqq x$ whenever $x \geqq (\theta - w)/2$.

Lemma 2. *If r is a positive constant, then $H(y, r) = [F(y + r) - F(y - r)]/f[(y + r)/2^{\frac{1}{2}}]f[(y - r)/2^{\frac{1}{2}}]$ is non-decreasing for $y \geqq 0$.*

Proof. A sufficient condition for $H(y, r)$ to be non-decreasing for $y \geqq 0$ is that its first derivative is non-negative, or, equivalently, that

$$H_1(y, r) = f(y + r) - f(y - r) + y[F(y + r) - F(y - r)] \geqq 0.$$

Since $H_1(y, 0) = 0$ it is sufficient to show that $H_1(y, r)$ is increasing in r for $r \geqq 0$ and fixed $y \geqq 0$, or that

$$(d/dr)H_1(y, r) = r[f(y - r) - f(y + r)] \geqq 0,$$

for $y \geqq 0$ and $r \geqq 0$, which is clearly so.

COROLLARY. *Let θ be a positive constant, then* $[H^2(y, r) - H^2(y - \theta, r)] \geqq 0$ *for all* $y \geqq \theta/2$ *and* $r \geqq 0$.

PROOF. The result follows easily from Lemma 2 and the fact that $H(y, r)$ is symmetric about $y = 0$ (for fixed $r \geqq 0$) and is non-negative.

LEMMA 3. *Let r be a positive constant, then $F(y - r)F(-y - r)$ is a non-increasing in y for $y \geqq 0$.*

PROOF. It is enough to show that the first derivative is non-positive, or equivalently, that $f(y + r)/F(y + r) \leqq f(r - y)/F(r - y)$, for $r \leqq 0$, $y \geqq 0$.

Since equality holds for $y = 0$ it is enough to use that $f(t)/F(t)$ is non-increasing for all t [Sampford (1953)].

COROLLARY. *If θ and w are non-negative constants, then* $F(x - \theta)F(\theta - x - w) - F(x)F(-x - w) \geqq 0$ *for* $x \geqq (\theta - w)/2$.

PROOF. Let $J(y, r) = -F(y - r)F(-y - r)$. Using this notation we are to show that

$$J(y, r) - J(y - \theta, r) \geqq 0$$

for $y \geqq \theta/2$ and $r \geqq 0$ where $J(y, r)$ is non-decreasing for $y \geqq 0$ and symmetric about $y = 0$. This is proved as in the corollary to Lemma 2.

APPENDIX II

Some nonlinear relationships between rank order probabilities. In Savage (1960), p. 520, linear relationships like the following have been presented:

$$P(011) = [P(0110) + P(0101) + 2P(0011)]/2.$$

A pair of non-linear relationships is obtained below. Note that no restriction is made of the densities $f(\cdot)$ and $g(\cdot)$.

THEOREM 8.

$$A\colon\ P(0110) = 2P(011) - 2P^2(01),$$
$$B\colon\ P(0101) = 2P^2(01) - 2P(0011),$$

PROOF. We note first that

$$\begin{aligned} P^2(01) &= [\textstyle\int_{-\infty}^{\infty} F(x)g(x)\,dx]^2 \\ &= \textstyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x)g(x)F(y)g(y)\,dx\,dy \\ &= 2\textstyle\int_{-\infty}^{\infty}\int_{x}^{\infty} F(x)F(y)g(x)g(y)\,dy\,dx. \end{aligned}$$

Then to prove A, one has

$$\begin{aligned} P(0110) &= 4\textstyle\int_{-\infty}^{\infty}\int_{x}^{\infty} F(x)[1 - F(y)]g(x)g(y)\,dy\,dx \\ &= 4\textstyle\int_{-\infty}^{\infty}\int_{x}^{\infty} F(x)g(x)g(y)\,dy\,dx - 2P^2(01) \\ &= 2P(011) - 2P^2(01). \end{aligned}$$

And to prove B, one has

$$P(0101) = 4 \int_{-\infty}^{\infty} \int_{x}^{\infty} F(x)[F(y) - F(x)]g(x)g(y)\, dy\, dx$$
$$= 2P^2(01) - 4 \int_{-\infty}^{\infty} \int_{x}^{\infty} F^2(x)g(x)g(y)\, dy\, dx = 2P^2(01) - 2P(0011).$$

COROLLARY 1.

$$A'\colon\ P(1001) = 2P(100) - 2P^2(10),$$
$$B'\colon\ P(1010) = 2P^2(10) - 2P(1100).$$

PROOF. Interchange F and G (f and g) in the proofs of A and B.

COROLLARY 2. $P(0011) + P(1100) = 2[P(011) + P(100)] - 1$.

PROOF. The set of all possible rank orders for $m = n = 2$ is an exhaustive set of mutually exclusive events. Therefore

$$P(0011) + P(1100) + P(0101) + P(1010) + P(1001) + P(0110) = 1.$$

Substituting the right members of A, B, A' and B' for $P(0110)$, $P(0101)$, $P(1001)$ and $P(1010)$, we obtain

$$P(0011) + P(1100) = 2(P(011) + P(100)] - 1.$$

Probabilities for all of the rank orders with $m = n = 2$ can be evaluated in terms of the probabilities for smaller sample sizes and $P(0011)$ or $P(1100)$. More generally, for $n = 2$ and arbitrary fixed $m = M$ let

$$z = (0, \cdots, 0, 1, 0, \cdots, 0, 1, 0, \cdots, 0),$$

where there are r_1, r_2 and r_3 0's respectively in the above sequence of 0's and where $r_1 + r_2 + r_3 = M$. Then

$$P(z) = [M!2!/r_1!r_2!r_3!]\int \cdots \int_{-\infty<x\leqq y<\infty} F^{r_1}(x)[F(y) - F(x)]^{r_2}[1 - F(y)]^{r_3}$$
$$g(x)g(y)\, dx\, dy$$
$$= [M!2!/r_1!r_2!r_3!]\sum_{1\leqq i_1\leqq r_2, 1\leqq i_2 \leqq r_3} \binom{r_2}{i_1}\binom{r_3}{i_2}(-1)^{r_2+r_3-(i_1+i_2)}$$
$$\cdot\int \cdots \int_{-\infty<x\leqq y<\infty} F^{r_1+r_2-i_1}(x)\cdot F^{r_3+i_1-i_2}(y)g(x)g(y)\, dx\, dy.$$

For $a + b < M$, the integral

$$\int \cdots \int_{-\infty<x\leqq y<\infty} F^{a}(x)F^{b}(y)g(x)g(y)\, dx\, dy$$

can be expressed as a linear combination of probabilities of rank orders for $m < M$ and $n = 1$ or 2. Therefore if all rank order probabilities for $m < M$ and $n = 1$ and 2 have been computed, the only new integrals required are of the form

$$A_i = \int \cdots \int_{-\infty<x\leqq y<\infty} F^{i}(x)F^{M-i}(y)g(x)g(y)\, dx\, dy, \qquad i = 0, \cdots, M.$$

Since

$$\begin{aligned}A_i + A_{M-i} &= \int \cdots \int_{-\infty<x\leqq y<\infty} [F^i(x)F^{M-i}(y) + F^{M-i}(x)F^i(y)]g(x)g(y)\,dx\,dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F^i(x)F^{M-i}(y)g(x)g(y)\,dx\,dy \\ &= [\int_{-\infty}^{\infty} F^i(x)g(x)\,dx][\int_{-\infty}^{\infty} F^{M-i}(y)g(y)\,dy],\end{aligned}$$

one needs to compute only one of the pair (A_i, A_{M-i}).

For $n > 2$ we must consider n-fold integrals of the form

$$A(i_1, \cdots, i_n) = \int \cdots \int_{-\infty<x_1\leqq\cdots\leqq x_n<\infty} \prod_{j=1}^{n} F^{i_j}(x_j)g(x_j)\,dx_j,$$

as above only those integrals for which $\sum_{j=1}^{n} i_j = M$ need be evaluated. Then if any of the $i_j = 0$ the dimensionality of the integral is easily decreased. Generally,

$$\sum A(i_1, \cdots, i_n) = \prod_{j=1}^{n} P(0, \cdots, 0, 1)$$

where $P(0, \cdots, 0, 1)$ contains i_j 0's, $j = 1, \cdots, n$, and where the summation is over all permutations of $(1, 2, \cdots, n)$.

REFERENCES

HODGES, J. L., JR. and LEHMANN, E. L. (1962). Probabilities of rankings for two widely separated normal distributions. *Studies in Mathematical Analysis and Related Topics* (edited by Gilbarg, D., Solomon H., and others). Stanford Univ. Press.

KLOTZ, J. H. (1962). Private communication.

KLOTZ, J. H. (1964). On the normal scores two-sample rank. *J. Amer. Statist. Assoc.* **59** 652–664.

MILTON, ROY C. (1965). Rank order probabilities: two-sample normal shift alternatives. Technical Report No. 53, Department of Statistics, University of Minnesota.

SAMPFORD, M. R. (1953). Some inequalities on Mill's ratio and related functions. *Ann. Math. Statist.* **24** 130–132.

SAVAGE, I. RICHARD (1956). Contributions to the theory of rank order statistics: the two sample case. *Ann. Math. Statist.* **27** 590–615.

SAVAGE, I. RICHARD (1957). Contributions to the theory of rank order statistics: the "trend" case. *Ann. Math. Statist.* **28** 968–977.

SAVAGE, I. RICHARD (1960). Contributions to the theory of rank order statistics: computation rules for probabilities of rank orders. *Ann. Math. Statist.* **31** 519–520.

SAVAGE, I. RICHARD (1964). Contributions to the theory of rank order statistics: applications of lattice theory. *Rev. Internat. Statist. Inst.* **32** 52–64.