*The Annals of Mathematical Statistics*
1969, Vol. 40, No. 4, 1449-1466

# THE SMIRNOV TWO SAMPLE TESTS AS RANK TESTS[1]

By G. P. Steck

*Sandia Laboratory*

**1. Introduction.** Small sample distributional problems associated with the Smirnov two-sample statistics have been of very great difficulty. These statistics, based on the maximum difference, minimum difference, and maximum absolute difference between the empirical distribution functions of two independent samples, were proposed for the two sample problem by Smirnov in 1939. The first two statistics are "one-sided" and the third is "two-sided". Despite their apparent simplicity of form it was over a decade before their distributions were found even for equal sample sizes. These results are due to Gnedenko and Korolyuk (1951), and independently to Drion (1952). Proofs of these first results were based on a random walk model which has been used for most subsequent results as well.

Closed form expressions for the distribution of the one-sided statistic have been given by Korolyuk (1955) for the case where one sample size is an integer multiple of the other, and by Hodges (1957) for the case where the sample sizes differ by one. Korolyuk (1955, page 85) also gives a formula for the general case, but it does not appear to be correct.

Expressions for the distribution of the two-sided statistic given by Korolyuk (1955) and Blackman (1956 and correction 1958) for the integer multiple case and by Depaix (1962) for the general case are extremely complicated and poorly suited for computation.

However, useful algorithms do exist for computing the small sample distributions of both statistics. Massey (1951) constructed a small table for the distribution of the two-sided statistic for $1 \leqq m = n \leqq 40$ and for $1 \leqq m \leqq n \leqq 10$ and certain other selected values of $m, n \leqq 20$ (Massey (1952)); however, Hodges (1957) could not check all of Massey's values in the latter case. Kim (personal communication) confirms Hodges figures and reports additional errors. Much larger tables of this distribution have been prepared by Marliss and Zayachkowski (1962) who tabulate the complete distribution for $1 \leqq m \leqq n \leqq 20$, and by Kim and Jennrich (1967) who tabulate the tail above the 80th percentile for $1 \leqq m \leqq n \leqq 100$. In addition, Borovkov et al (1964) tabulate the percentiles straddling the 95th, 98th, and 99th percentiles of both the one-sided and two-sided distributions for $1 \leqq m \leqq n \leqq 50$.

The limiting distributions, given by Smirnov in his original papers (Smirnov (1939a, b)) when the ratio of the sample sizes converges to a constant bounded away from zero and infinity, are approached most erratically (see, for example,

Received 29 July 1968; revised 14 January 1969.

[1] This work was supported by the United States Atomic Energy Commission.

Hodges (1957)) which makes the study of the exact distributions (and their approximations) more important than would otherwise be the case.

For fine summaries of results concerning the Smirnov and related statistics see Hájek and Sídak (1967), or Barton and Mallows (1965) and Darling (1957).

It should be emphasized at this point that all of the above results assume the usual null hypothesis of equality between the underlying distributions of the two populations. Nothing appears to have been published concerning the distributions, asymptotic or otherwise, under alternative hypotheses. For completeness, although they are not distributional results, it should be noted that Massey (1950) gives a lower bound to the asymptotic power of the two-sided test and Capon (1965) and Klotz (1967) give some asymptotic relative efficiencies.

The principal results of this paper are in two directions. First, the Smirnov statistics are expressed explicitly in terms of the ranks of one sample and their distributions are then expressed in terms of the joint distribution of those ranks. Following Lehmann (1953), we are then able to derive the distribution of the one-sided statistic when one underlying distribution is a power of the other. This distribution is expressed as a determinant.

Second, we show that the frequency content under the null hypothesis of any parallelepiped in the sample space of the ranks of one sample is expressible as a determinant. This result is easily specialized to give a similar expression for the null joint distribution of the one-sided statistics and the null distribution of the two-sided statistic for arbitrary sample sizes.

In addition, Korolyuk's formula for the integer multiple case is shown to hold whenever a certain sequence of numbers is an arithmetic progression. This condition is met, for example, sufficiently far in the upper tail of the distribution for arbitrary sample sizes and periodically all through the distribution when one sample size is congruent to one modulo the other. This result gives Korolyuk's translated formula wider applicability and suggests its use as an approximation to the exact distribution in the general case. Although it is a more complicated approximation than the one suggested by Hodges (1957) its relative error for the cases studied (sample sizes up to twenty-five) is usually more than an order of magnitude smaller.

We also apply some of these results to statistics related to the Smirnov statistics. First, we note certain simplifications in the computation of the null distribution of the sum of the two one-sided statistics for arbitrary sample sizes. And, second, we show that any test based on the two sample statistic recommended by Vincze (1957, 1959, 1961) consisting of considering jointly a one-sided statistic and the place position where that extremum first occurred is equivalent to the test based on the one-sided statistic alone when the sample sizes are relatively prime.

**2. The Smirnov statistics and their distributions in terms of the ranks of one sample and their joint distribution.** Let $X_1 \leqq X_2 \leqq \cdots \leqq X_m$ be the order statistics from a sample of $m$ independent identically distributed random variables with a continuous distribution function $F$, and let $Y_1 \leqq Y_2 \leqq \cdots \leqq Y_n$ be the order statistics from a sample of $n$ independent identically distributed random variables with a continuous distribution function $G$. Let

$$F_m(z) = 0 \quad z < X_1 \qquad G_n(z) = 0 \quad z < Y_1$$
$$= i/m \quad X_i \leqq z < X_{i+1} \qquad = i/n \quad Y_i \leqq z < Y_{i+1}$$
$$= 1 \quad X_m \leqq z \quad , \qquad = 1 \quad Y_n \leqq z$$

be the corresponding empirical distribution functions. Finally, let $Z_1 \leqq Z_2 \leqq \cdots \leqq Z_{m+n}$ denote the ordered combined sample and let $R_i$ and $S_j$ denote the ranks of $X_i$ and $Y_j$, respectively, in the ordered combined sample.

The Smirnov statistics are

$$D^+(m, n) = \sup_z (F_m(z) - G_n(z))$$
$$D^-(m, n) = \sup_z (G_n(z) - F_m(z))$$
$$D(m, n) = \max (D^+(m, n), D^-(m, n))$$

and their possible values are integer multiples of $1/mn$.

THEOREM 2.1.

$$P(mnD^+(m, n) \leqq r) = P[R_i \geqq (i(m + n) - r)/m, 1 \leqq i \leqq m]$$
$$P(mnD^+(m, n) < r) = P[R_i > (i(m + n) - r)/m, 1 \leqq i \leqq m].$$

The proof follows immediately from Maag and Stephens (1968) who observed that

$$mnD^+(m, n) = \sup_i \{(m + n)i - mR_i\}.$$

The basic idea can also be found in Anderson (1962).

THEOREM 2.2.

$$P(mnD^-(m, n) \leqq r) = P[R_i \leqq (i(m + n) - n + r)/m, 1 \leqq i \leqq m]$$
$$P(mnD^-(m, n) < r) = P[R_i < (i(m + n) - n + r)/m, 1 \leqq i \leqq m].$$

PROOF.

$$D^-(m, n) = \sup_z (G_n(z) - F_m(z)) = \sup_z [(1 - F_m(z)) - (1 - G_n(z))]$$
$$= \sup_z (\tilde{F}_m(z) - \tilde{G}_n(z)) = \tilde{D}^+(m, n),$$

where $\tilde{F}_m$, $\tilde{G}_n$ and $\tilde{D}^+$ are $F_m$, $G_n$ and $D^+$, respectively, computed for the sample ordered from largest to smallest; that is, for $\tilde{Z}_1 \geqq \tilde{Z}_2 \geqq \cdots \geqq \tilde{Z}_{m+n}$.
Thus, by Theorem 2.1, letting $a_i = (i(m + n) - r)/m$

$$P(mnD^-(m, n) \leqq r) = P(mn\tilde{D}^+(m, n) \leqq r)$$
$$= P(\tilde{R}_1 \geqq a_1, \tilde{R}_2 \geqq a_2, \cdots, \tilde{R}_m \geqq a_m).$$

But $\tilde{R}_i = m + n + 1 - R_{m-i+1}$ so that

$$P(mn\tilde{D}^+(m, n) \leqq r) = P(R_1 \leqq m + n + 1 - a_m,$$
$$R_2 \leqq m + n + 1 - a_{m-1}, \cdots, R_m \leqq m + n + 1 - a_1)$$

and hence

$$P(mnD^-(m, n) \leqq r) = P(R_i \leqq (i(m + n) - n + r)/m, 1 \leqq i \leqq m).$$

The second part of the theorem follows similarly. Theorems 2.3 and 2.4 follow trivially from Theorems 2.1 and 2.2.

THEOREM 2.3.

$$P(mnD^+(m, n) \leqq r, mnD^-(m, n) \leqq s)$$
$$= P[(i(m + n) - r)/m \leqq R_i \leqq (i(m + n) - n + s)/m, 1 \leqq i \leqq m].$$

If one of the inequalities for $D^+$ or $D^-$ is replaced by strict inequality then the corresponding inequality for $R_i$ is similarly changed.

THEOREM 2.4.

$$P(mnD(m, n) \leqq r)$$
$$= P[(i(m + n) - r)/m \leqq R_i \leqq (i(m + n) - n + r)/m, 1 \leqq i \leqq m]$$
$$P(mnD(m, n) < r)$$
$$= P[(i(m + n) - r)/m < R_i < (i(m + n) - n + r)/m, 1 \leqq i \leqq m].$$

In general the bounds on $R_i$ will not be integers. In fact $r$ is a possible value of $D$, $D^+$, or $D^-$ if and only if at least one of the corresponding bounds of $R_i$ is an integer. This follows since if no bound is an integer then $P(mnD < r) = P(mnD \leqq r)$. Kim and Jennrich (1967) give a result equivalent to Theorem 2.4 when $F = G$.

COROLLARY 2.1. *If $n = mp$, $p$ an integer greater than or equal to one, then*

$$P(nD^+(m, n) \leqq t) = P(R_i \geqq i(p + 1) - t, 1 \leqq i \leqq m)$$
$$P(nD(m, n) \leqq t) = P(i(p + 1) - t \leqq R_i \leqq i(p + 1) + t - p, 1 \leqq i \leqq m).$$

COROLLARY 2.2. *The Smirnov statistics are expressible in terms of the ranks of one sample as follows.*

$$mnD^+(m, n) = \sup_{1 \leqq i \leqq m} ((m + n)i - mR_i)$$
$$mnD^-(m, n) = \sup_{1 \leqq i \leqq m} (mR_i - (m + n)i + n)$$
$$mnD(m, n) = n/2 + \sup_{1 \leqq i \leqq m} |mR_i - (m + n)i + n/2| .$$

2.1. *Summary of notation.* For easy reference we collect here some of the notation used in this paper.

$[x]$ denotes the largest integer less than or equal to $x$.

$\langle x \rangle = -[-x]$ denotes the smallest integer greater than or equal to $x$.

$x!$ denotes $\Gamma(x + 1)$ for non-integer as well as integer $x$.

$r_i$ denotes a possible value of $R_i$.

$p$ denotes an integer such that in the integer multiple case $n = mp$.

$r$ denotes a possible value of $D^+$.

$s$ denotes a possible value of $D^-$.

$a_i = (i(m + n) - r)/m$.

$\alpha_i = \max(i, [a_i] + 1)$ denotes the larger of $i$ and the smallest integer exceeding $a_i$.

$b_i = \alpha_i - 1 = \max(i - 1, [a_i])$.

$c_i = \min(n + i + 1, \langle(i(m + n) - n + s)/m\rangle)$.

$h = \langle r/n\rangle$ denotes the smallest index such that $a_i \geqq i$.

$k$ denotes the exponent in the hypothesis $G^k = F$.

$A_i = \Gamma(\alpha_i + ki - i)/\Gamma(\alpha_i - i) = (b_i + ki - i)!/(b_i - i)!$.

$M_i = n!\binom{m}{i}/(n + ki)!$.

**3. The non-null distribution of $D^+_{mn}$. Exact formulae for $P(mnD^+(m, n) \geqq r)$ when $G^k = F$.** Theorem 2.1 implies

$$(3.1)\quad P(mnD^+(m, n) < r) = \sum_{r_m > a_m}^{m+n} \sum_{r_{m-1} > a_{m-1}}^{r_m - 1} \cdots \sum_{r_1 > a_1}^{r_2 - 1} P(R_1 = r_1, R_2 = r_2, \cdots, R_m = r_m),$$

where $a_i = (i(m + n) - r)/m$ as before.

Lehmann (1953) gives the following expression for the joint frequency function of the $\{R_i\}$ under the assumption $G^k = F$, $k \geqq 0$,

$$(3.2)\quad P(R_1 = r_1, \cdots, R_m = r_m \mid G^k = F) = k^m m! n!/\Gamma(n + km + 1) \cdot\prod_{j=1}^{m} \Gamma[r_j + j(k - 1)]/\Gamma[r_j + (j - 1)(k - 1)].$$

Although the following derivation is true for any $k \geqq 0$, it makes statistical sense only if $0 \leqq k \leqq 1$, because $D^+$ is the appropriate test statistic for testing the hypothesis $F = G$ against the alternative $G^k = F$ only in that case.

Substituting (3.2) into (3.1), we begin evaluating (3.1) as follows. Let

$$S_i(k) = \sum_{r_i = \alpha_i}^{r_{i+1} - 1} \Gamma[r_i + i(k - 1)]/\Gamma[r_i + (i - 1)(k - 1)] \cdot \sum_{r_{i-1} = \alpha_{i-1}}^{r_i - 1} \Gamma[r_{i-1} + (i - 1)(k - 1)]/\Gamma[r_{i-1} + (i - 2)(k - 1)] \cdots \sum_{r_1 = \alpha_1}^{r_2 - 1} \Gamma(r_1 + k - 1)/\Gamma(r_1)$$

where $\alpha_i$ is the larger of $i$ and the smallest integer exceeding $a_i$.

Then, since

$$\sum_{i=b}^{c} \Gamma(i + a + 1)/\Gamma(i + 1) = a! \sum_{i=b}^{c} \binom{i+a}{a} = a!\left[\binom{a+c+1}{a+1} - \binom{a+b}{a+1}\right],$$

we have

$$S_1(k) = (r_2 - 1) \cdots (r_2 + k - 2)/k - (\alpha_1 - 1) \cdots (\alpha_1 + k - 2)/k = h_1(r_2) - h_1(\alpha_1)$$

where $h_1(x) = \Gamma(x + k - 1)/(k\Gamma(x - 1))$.

Similarly,

$$S_2(k) = \sum_{r_2 = \alpha_2}^{r_3 - 1} \Gamma(r_2 + 2k - 2)/\Gamma(r_2 + k - 1) \cdot[\Gamma(r_2 + k - 1)/k\Gamma(r_2 - 1) - h_1(\alpha_1)] = h_2(r_3) - h_2(\alpha_2),$$

where

$$h_2(x) = (2k^2)^{-1}\Gamma(x + 2k - 2)/\Gamma(x - 2) \\ - k^{-1}\Gamma(x + 2k - 2)/\Gamma(x + k - 2)h_1(\alpha_1).$$

By induction

$$S_i(k) = h_i(r_{i+1}) - h_i(\alpha_i) \tag{3.3}$$

where

$$h_i(x) = \Gamma(x + ki - i)/[i!k^i\Gamma(x - i)] \\ - \textstyle\sum_{j=1}^{i-1} \Gamma(x + ki - i)/[(i - j)!k^{i-j}\Gamma(x + kj - i)]h_j(\alpha_j). \tag{3.4}$$

If $B_j(k) = j!k^jh_j(\alpha_j)$ and $C_j$ denotes $B_j(1)/j!$, then (3.3) and (3.4) give

$$S_i(k) = [(i!k^i)^{-1}\Gamma(r_{i+1} + ki - i)/\Gamma(r_{i+1} - i) \\ - \textstyle\sum_{j=1}^{i} \binom{i}{j}\Gamma(r_{i+1} + ki - i)/\Gamma(r_{i+1} + kj - i)B_j(k)] \tag{3.5}$$

and

$$S_i(1) = \binom{r_{i+1}-1}{i} - \textstyle\sum_{j=1}^{i} \binom{r_{i+1}-1}{i-j}C_j .$$

Note from the equations for $h_1(x)$ and $h_2(x)$ that $h_1(1) = 0$ and that $h_2(2) = 0$ if $\alpha_1 = 1$. Repeated use of (3.4) shows that $h_i(i) = 0$ for all $i < h$. Therefore $B_i(k) = 0$ for $i < h$, where $h = \langle r/n\rangle$ is the smallest index such that $a_i \geqq i$ (i.e., $\alpha_i > i$).

Hence, using $r_{m+1} = m + n + 1$, (3.5) implies the following theorem.

THEOREM 3.1.

$$\begin{aligned} P(mnD^+(m, n) \geqq r) &= 1 - P(mnD^+(m, n) < r) \\ &= 1 - k^m m!n!/\Gamma(n + km + 1)S_m(k) \\ &= n!\textstyle\sum_{j=h}^{m} \binom{m}{j}B_j(k)/\Gamma(n + kj + 1), \end{aligned} \tag{3.6}$$

*where the $\{B_i(k)\}$ are as defined above and are determined recursively by*

$$B_i(k) = \Gamma(\alpha_i + ki - i)/\Gamma(\alpha_i - i) \\ - \textstyle\sum_{j=h}^{i-1} \binom{i}{j}\Gamma(\alpha_i + ki - i)/\Gamma(\alpha_i + kj - i)B_j(k). \tag{3.6a}$$

For the special case $k = 1$,

$$\binom{m+n}{m}P(mnD^+(m, n) \geqq r) = \textstyle\sum_{j=h}^{m} \binom{m+n}{m-j}C_j \tag{3.7}$$

where, letting $b_i = \alpha_i - 1 = [a_i]$,

$$C_i = \binom{b_i}{i} - \textstyle\sum_{j=h}^{i-1} \binom{b_i}{i-j}C_j , \qquad i = h + 1, h + 2, \cdots , m \tag{3.7a}$$

and

$$C_h = \binom{b_h}{h}.$$

Specializing further to the case $k = 1$, $n = mp$, it can be shown that, letting $r = nt$,

$$C_{h+k} = b_{h+k}/(h + k)\cdot\text{coefficient of } w^{kp+\beta} \text{ in } (1 + w + \cdots + w^{p-1})^{h+k},$$

where $\beta$ is defined by $t = hp - \beta$, $0 \leqq \beta < p$. Since $C_{h+k} = 0$ if $kp + \beta > (h + k)(p - 1)$, it follows that there are at most $t - h + 1$ nonzero $C$'s when $n = mp$. In particular, if $p = 1$ then $h = \langle t/p\rangle = t$, $C_h = 1$ and $C_j = 0$ for $j > h$, and (3.7) becomes,

$$\binom{2n}{n}P(nD^{+}(n, n) \geqq t) = \binom{2n}{n-t}$$

as is well known.

3.1. *A determinant form for* $P(mnD^{+} \geqq r)$ *when* $G^k = F$. Writing (3.6a) as

$$\sum_{j=h}^{i} \binom{i}{j}\Gamma(\alpha_i + ki - i)/\Gamma(\alpha_i + kj - i)B_j(k)$$
$$= \Gamma(\alpha_i + ki - i)/\Gamma(\alpha_i - i) = A_i \text{ (say) } i = h, h + 1, \cdots, m$$

and following Wald and Wolfowitz (1939), we can write $B_k = D^{-1}A$ (remember $0 \leqq k \leqq 1$ and that $\alpha_i = b_i + 1$) where $A' = (A_h, A_{h+1}, \cdots, A_m)$, $B_k' = (B_h(k), B_{h+1}(k), \cdots, B_m(k))$, and $D$ is a square matrix such that the element in the $(i - h + 1)$th row and $(j - h + 1)$th column is

$$d_{ij} = \binom{i}{j}(b_i + ki - i)!/(b_i + kj - i)!, \qquad h \leqq i, j \leqq m.$$

It then follows from (3.6) that

$$P(mnD^{+}(m, n) \geqq r \mid G^k = F) = M'D^{-1}A, \tag{3.1.1}$$

where, letting $x!$ denote $\Gamma(x + 1)$ for non-integer $x$ as well as integer $x$,

$$M' = (M_h, M_{h+1}, \cdots, M_m)$$
$$= \left(\frac{n!\binom{m}{h}}{(n + kh)!}, \frac{n!\binom{m}{h+1}}{(n + kh + k)!}, \cdots, \frac{n!\binom{m}{m}}{(n + km)!}\right),$$

But it is easily verified that

$$\begin{pmatrix} D^{-1} & -D^{-1}A \\ 0 & 1 \end{pmatrix}\begin{pmatrix} D & A \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

so that

$$\begin{pmatrix} D & A \\ M' & 0 \end{pmatrix}\begin{pmatrix} D^{-1} & -D^{-1}A \\ 0 & 1 \end{pmatrix}\begin{pmatrix} D & A \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} D & A \\ M' & 0 \end{pmatrix}$$

or, multiplying the first two factors,

$$\begin{pmatrix} I & 0 \\ M'D^{-1} & -M'D^{-1}A \end{pmatrix}\begin{pmatrix} D & A \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} D & A \\ M' & 0 \end{pmatrix}.$$

Taking determinants on both sides of the above enables (3.1.1) to be rewritten as

$$P(mnD^{+}(m, n) \geqq r \mid G^k = F) = -\begin{vmatrix} D & A \\ M' & 0 \end{vmatrix}. \tag{3.1.2}$$

Writing out the matrix on the right-hand side of (3.1.2) and rearranging it slightly gives

$$P(mnD^{+} \geqq r \mid G^{k} = F) = (-1)^{m-h+2}$$

$$\cdot \begin{vmatrix} A_h & A_{h+1} & A_{h+2} & \cdots & A_{m-1} & A_m & 0 \\ 1 & d_{h+1,h} & d_{h+2,h} & \cdots & d_{m-1,h} & d_{m,h} & M_h \\ 0 & 1 & d_{h+2,h+1} & \cdots & d_{m-1,h+1} & d_{m,h+1} & M_{h+1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & d_{m,m-1} & M_{m-1} \\ 0 & 0 & 0 & \cdots & 0 & 1 & M_m \end{vmatrix}.$$

Shorack (1967) tabulates the distribution of the Wilcoxon statistic for these alternatives for $4 \leqq m, n \leqq 8$. A similar tabulation of the distribution of $D_{mn}^{+}$ statistic was also made in order to compare the power of these two tests. The comparison was made for randomized tests with $\alpha = .10$ against alternatives: $1/k = 2, 3, 4, 6, 9, 12$.

The results were that the Wilcoxon test was generally slightly more powerful. The Smirnov test was slightly better for the larger sample sizes when $1/k$ was large and $m = n + 1$. The differences in power never exceeded 0.05 although the percent difference was occasionally as high as 10 percent.

**4. New results for the null distributions of the Smirnov statistics.** Let $b_1 \leqq b_2 \leqq \cdots \leqq b_m$ and $c_1 \leqq c_2 \leqq \cdots \leqq c_m$ be increasing sequences of integers such that $i - 1 \leqq b_i < c_i \leqq n + i + 1$. Let

$$\begin{aligned} d_{ij} &= 0 && \text{if } i - j > 1 \text{ or if } c_i - b_j \leqq 1 \\ &= \binom{c_i - b_j + j - i - 1}{j - i + 1} && \text{otherwise,} \end{aligned}$$

and let $N(b_i, \cdots, b_j; c_i, \cdots, c_j)$ denote the number of ways $X_i, \cdots, X_j$ can occur in the combined sample so that $b_\nu < R_\nu < c_\nu$ for every $i \leqq \nu \leqq j$. When $i = 1$ and $j = m$, we will denote this function by $N_m(b; c)$.

THEOREM 4.1. $N(b_1, \cdots, b_m; c_1, \cdots, c_m) = \det(d_{ij})$. This formula was found by comparing exact expressions for the number of such arrangements when $m$ was up to 4 with corresponding expressions for the one-sided case. Their similarity suggested that perhaps a determinant would provide the answer since one had already been obtained for the one-sided case, and since one is implied by a recurrence relation given by Kemperman (1957) for the two-sided case when $n = mp$. A simple determinant was constructed to cover the cases $m \leqq 4$ and the extension was obvious and checked numerically for a few cases with $m = 5$. The proof came last and consisted of showing that the determinant satisfied the recurrence relations and boundary condition required by the function $N_m(b; c)$.

The recurrence relations are obtained by letting each $b_i$ and $c_i$ in turn be changed by 1. If $c_i$ is increased to $c_i + 1$, $N_m(b; c)$ is increased by the number of ways the event ($b_j < R_j < c_j$ all $j \neq i$, $R_i = c_i$) can occur. Since the set of ranks with indices less than $i$ is conditionally independent, given $R_i$, of the set of ranks with indices greater than $i$, the following difference equation holds.

$$
\begin{aligned}
&N(b_1, \cdots, b_m; c_1, \cdots, c_i + 1, \cdots, c_m) \\
(4.1)\quad &= N(b_1, \cdots, b_m; c_1, \cdots, c_m) + N(b_1, \cdots, b_{i-1}; c_1, \cdots, c_{i-1}) \\
&\qquad \cdot N(b_{i+1}^*, \cdots, b_m^*; c_{i+1}, \cdots, c_m),
\end{aligned}
$$

where $b_j^* = \max(b_j, c_i)$, $j = i + 1, \cdots, m$. Similarly, if $b_i$ is decreased to $b_i - 1$, the following difference equation holds.

$$
\begin{aligned}
&N(b_1, \cdots, b_i - 1, \cdots, b_m; c_1, \cdots, c_m) \\
&\quad = N(b_1, \cdots, b_m; c_1, \cdots, c_m) + N(b_1, \cdots, b_{i-1}; c_1^*, \cdots, c_{i-1}^*) \\
&\qquad \cdot N(b_{i+1}, \cdots, b_m; c_{i+1}, \cdots, c_m),
\end{aligned}
$$

where $c_j^* = \min(b_i, c_j)$, $j = 1, 2, \cdots, i - 1$.

To prove that the determinant solution satisfies (4.1) observe first that $c_i + 1$ occurs only in the $i$th row of the matrix and second that each element in that row satisfies

$$\binom{c_i+1-b_j+j-i-1}{j-i+1} = \binom{c_i-b_j+j-i-1}{j-i+1} + \binom{c_i-b_j+j-i-1}{j-1}.$$

Thus the determinant for $c_i + 1$, call it $D_m(c_i + 1)$, equals the sum of two determinants one of which is $D_m(c_i)$ and the other is of the form

$$\left|\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}\right| = |A| \cdot |C| \qquad A: (i-1) \times (i-1)$$

$$C: (m - i + 1) \times (m - i + 1)$$

where $|A| = N(b_1, \cdots, b_{i-1}; c_1, \cdots, c_{i-1})$ and

$$C = \begin{bmatrix}
1 & \binom{c_i-b_{i+1}}{1} & \binom{c_i-b_{i+2}+1}{2} & \cdots & \binom{c_i-b_m+m-i-1}{m-i} \\
1 & \binom{c_{i+1}-b_{i+1}-1}{1} & \binom{c_{i+1}-b_{i+2}}{2} & \cdots & \binom{c_{i+1}-b_m+m-i-2}{m-i} \\
0 & 1 & \binom{c_{i+2}-b_{i+2}-1}{1} & \cdots & \binom{c_{i+2}-b_m+m-i-3}{m-i-1} \\
\vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \cdots & \binom{c_m-b_m-1}{1}
\end{bmatrix}.$$

The determinant of $C$ is seen to be $N(b_{i+1}^*, \cdots, b_m^*; c_{i+1}, \cdots c_m)$ by the following argument. Let $q \geqq 1$ denote the largest index such that $c_i - b_{i+q} > 0$. That is, $q$ denotes the number of non-zero entries in the first row of $C$. The proof is to modify $C$, without changing its determinant, until the first row of $C$ is $(1\ 0\ 0\ 0 \cdots 0)$. At that point $C$ itself will be changed in such a way that it will be obvious that its determinant is the required quantity. The proof makes repeated use of the Vandermonde convolution formula in the form of the following lemma.

LEMMA 4.1.

$$
(4.2)\qquad
\begin{bmatrix}
1 & \binom{a}{1} & \binom{a}{2} & \cdots & \binom{a}{j} \\
1 & \binom{b}{1} & \binom{b}{2} & \cdots & \binom{b}{j} \\
0 & 1 & \binom{c}{1} & \cdots & \binom{c}{j-1} \\
\vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \cdots & \binom{d}{2} \\
0 & 0 & 0 & \cdots & \binom{e}{1}
\end{bmatrix}
\cdot
\begin{bmatrix}
1 & 0 & 0 & \cdots & \binom{x}{j} \\
0 & 1 & 0 & \cdots & \binom{x}{j-1} \\
0 & 0 & 1 & \cdots & \binom{x}{j-2} \\
\vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \cdots & \binom{x}{1} \\
0 & 0 & 0 & \cdots & 1
\end{bmatrix}
=
\begin{bmatrix}
1 & \binom{a}{1} & \binom{a}{2} & \cdots & \binom{a}{j-1} & \binom{a+x}{j} \\
1 & \binom{b}{1} & \binom{b}{2} & \cdots & \binom{b}{j-1} & \binom{b+x}{j} \\
0 & 1 & \binom{c}{1} & \cdots & \binom{c}{j-2} & \binom{c+x}{j-1} \\
\vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \binom{d}{1} & \binom{d+x}{2} \\
0 & 0 & 0 & \cdots & 1 & \binom{e+x}{1}
\end{bmatrix}.
$$

The proof is immediate through use of the Vandermonde convolution formula

$$\textstyle\sum_{k=0}^{j} \binom{a}{k}\binom{x}{j-k} = \binom{a+x}{j} \quad a, x \text{ real.}$$

It is convenient to have symbols for the matrices in (4.2). Let them be, from left to right, $C_j(0)$, $V_j(x)$ and $C_j(x)$, respectively. Note that the number of leading zeroes in the rows of $C_j(0)$ is irrelevant and note, too, that the determinants of $C_j(0)$ and $C_j(x)$ are equal.

We will also need the inverse of $V_j(x)$; call it $W_j(x)$. It is easy to verify that

$$
W_j(x) =
\begin{bmatrix}
1 & 0 & 0 & \cdots & -\binom{x}{j} \\
0 & 1 & 0 & \cdots & -\binom{x}{j-1} \\
\vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \cdots & -\binom{x}{1} \\
0 & 0 & 0 & \cdots & 1
\end{bmatrix}.
$$

Hence,

$$(4.3)\qquad C_j(x)W_j(x) = C_j(0).$$

Now let $V_j^*(x)$ denote a partitioned matrix $\begin{pmatrix} V_j(x) & 0 \\ 0 & I \end{pmatrix}$ where $V_j(x)$ is the square matrix defined above and $I$ is an identity matrix of some appropriate dimension. Define $W_j^*(x)$ similarly. Then postmultiplication of $C$ by $V_j^*(x)$ or $W_j^*(x)$ changes only the $(j+1)$th column of $C$.

If $x_j = b_{i+1} - b_{i+j} + j - 1$ then the matrix $C_{q-1}^* = C \cdot W_2^*(x_2) \cdot W_3^*(x_3) \cdots W_{q-1}^*(x_{q-1})$ has an upper left hand $q \times q$ minor which has the general form of

$C_{q-1}(0)$ as defined in Lemma 4.1; that is, the "numerators" of the binomial coefficients in any row of this minor are all equal. That this is, indeed, accomplished note that the upper left $3 \times 3$ principal minor of $C$ is of the form $C_2(x_2)$; hence the corresponding minor of $C \cdot W_2^*(x_2)$ is of the form $C_2(0)$ by (4.3) and the upper left $4 \times 4$ minor is of the form $C_3(x_3)$ which is ready for postmultiplication by $W_3^*(x_3)$, and so on. Now, let $y_j = -(c_i - b_{i+1} - j + 2)$, and consider the matrix

$$C^* = C_{q-1}^* V_{q-1}^*(y_{q-1}) \cdot V_{q-2}^*(y_{q-2}) \cdots V_2(y_2) V_1(y_1).$$

Each of these matrix multiplications is like using (4.2) with $x = j - 1 - a$; and, since $\binom{a+x}{j} = \binom{j-1}{j} = 0$, a zero is created in the upper right element of the product matrix.

The result of all this is that we have created a matrix $C^*$ whose determinant is the same as that of $C$ and which has $(1\ 0\ 0 \cdots 0)$ for its first row. What are its other elements? Each successive multiplication by $W$ subtracted $b_{i+1} - b_{i+j} + j - 1$ from the "numerators" of each binomial coefficient in the $(j + 1)$th column $(j \leqq q - 1)$ and each successive multiplication by $V$ subtracted $c_i - b_{i+1} - j + 1$ from each "numerator" in the $(j + 1)$th column $(j \leqq q - 1)$. The net result is that $c_i - b_{i+j}$ is subtracted from each "numerator" in the $(j + 1)$th column $(j \leqq q - 1)$. Thus the element in the $(k + 1)$th row and $(j + 1)$th column of $C^*$ is $\binom{c_i + k - c_i + j - k - 1}{j - k + 1}$ for $j \leqq q$. Since no $V$ or $W$ changed any element in any column except the first $q$, the element in the $(k + 1)$th row and $(j + 1)$th column is still $\binom{c_i + k - b_{i+j} + j - k - 1}{j - k + 1}$ for $j > q$. Letting $b_{i+j}^* = \max(c_i, b_{i+j})$, $j = 1, 2, \cdots, m - i$, it follows from the definition of $q$ that $b_{i+j} < c_i$ for $j \leqq q$ and $b_{i+j} \geqq c_i$ for $j > q$. Hence $b_{i+j}^* = c_i$ for $j \leqq q$ and $b_{i+j}^* = b_{i+j}$ for $j > q$. It then follows that $|C| = N(b_{i+1}^*, \cdots, b_m^*; c_{i+1}, \cdots, c_m)$ as desired. Thus $N_m(b; c) = |(d_{ij})|$ is seen to satisfy the required recurrence relation on $c_i$. A very similar argument shows that it also satisfies the required recurrence relation on $b_i$.

All that remains of the proof is to show that the determinant $|(d_{ij})|$ satisfies the boundary condition

$$N(b, \cdots, b; c, \cdots, c) = \binom{c - b - 1}{m}.$$

With $L = c - b - 1 \geqq m$ we write

$$|(d_{ij})| = \begin{vmatrix} \binom{L}{1} & \binom{L+1}{2} & \binom{L+2}{3} & \cdots & \binom{L+m-2}{m-1} & \binom{L+m-1}{m} \\ 1 & \binom{L}{1} & \binom{L+1}{2} & \cdots & \binom{L+m-3}{m-2} & \binom{L+m-2}{m-1} \\ 0 & 1 & \binom{L}{1} & \cdots & \binom{L+m-4}{m-3} & \binom{L+m-3}{m-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & & 1 & \binom{L}{1} \end{vmatrix}.$$

Using a familiar result on reciprocal series (see, for example, Riordan (1958, page 45)), we see this determinant is $(-1)^m$. (The coefficient of $t^m$ in

$\{1 + \binom{L}{1}t + \binom{L+1}{2}t^2 + \cdots\}^{-1} = (1 - t)^L$). This coefficient is $\binom{L}{m}$; hence

$$|(d_{ij})| = \binom{L}{m} = \binom{c-b-1}{m}.$$

This completes the proof of the theorem since we have shown that $|(d_{ij})|$ satisfies the recurrence relations and boundary conditions required by $N_m(b; c)$.

Determinant forms for the distributions of $D^+$, $D^-$ and $D$ are obtained by proper choices of the sequences $\{b_i\}$ and $\{c_i\}$.

In the following theorems we let $[x]$ denote the largest integer less than or equal to $x$ and let $\langle x \rangle = -[-x]$ denote the smallest integer greater than or equal to $x$.

THEOREM 4.2. $\binom{m+n}{m}P(mnD^+ < r, mnD^- < s) = |(d_{ij})|$ *where*

$$\begin{aligned} d_{ij} &= 0, && \text{if } i - j > 1 \text{ or } c_i - b_j \leqq 1; \\ &= \binom{c_i - b_j + j - i - 1}{j - i + 1}, && \text{otherwise}; \\ b_i &= \max\,(i - 1, [(i(m + n) - r)/m]) \\ c_i &= \min\,(n + i + 1, \langle (i(m + n) - n + s)/m \rangle). \end{aligned}$$

*Moreover, if* $r = s$ *the matrix* $(d_{ij})$ *is symmetric about the non-principal diagonal; that is,* $d_{ij} = d_{m-j+1,m-i+1}$. *This can be directly verified making use of* $\langle -x \rangle = -[x]$ *and* $\max\,(-x, -y) = -\min\,(x, y)$.

If $r = s$ in Theorem 4.2 one has $P(mnD < r)$; if $s = mn + 1$ one has $P(mnD^+ < r)$ and if $r = mn + 1$ one has $P(mnD^- < s)$. In the latter two cases $c_i = n + i + 1$ and $b_i = i - 1$, respectively. This gives

THEOREM 4.3. $\binom{m+n}{m}P(mnD^+ < r) = |(d_{ij})|$ *where*

$$\begin{aligned} d_{ij} &= 0, && \text{if } i - j > 1; \\ &= \binom{n - b_j + j}{j - i + 1}, && \text{otherwise}; \\ b_i &= \max\,(i - 1, [(i(m + n) - r)/m]). \end{aligned}$$

4.1. *Extension of a formula due to Korolyuk.* Korolyuk (1955) gives a formula for $P(D^+(m, n) \geqq a)$ when $F = G$ for the special case $n = mp$. In our notation his formula can be written as

$$\binom{m+n}{m}P(nD^+(m, n) \geqq t) = \textstyle\sum_{j=h}^{m} t/(n - b_j + j)\binom{m+n-b_j-1}{m-j}\binom{b_j}{j}.$$

Writing $mt = r$ and replacing $t$ by $\langle r/m \rangle = m + n - b_m$ gives

$$\begin{aligned} (4.1.1)\quad \binom{m+n}{m}&P(mnD^+(m, n) \geqq r) \\ &= \textstyle\sum_{j=h}^{m} (m + n - b_m)/(m + n - b_j)\binom{m+n-b_j}{m-j}\binom{b_j}{j}. \end{aligned}$$

We will show this "translated" form of Korolyuk's formula holds for all $m$, $n$, and $r$ such that $b_h, b_{h+1}, \cdots, b_m$ form an increasing arithmetic progression. Using (3.7a) to express (4.1.1) in terms of the $\{C_j\}$, one sees that the more general form of Korolyuk's formula is equivalent to

$$\begin{aligned} \binom{m+n}{m}&P(mnD^+(m, n) \geqq r) \\ &= \textstyle\sum_{i=h}^{m} C_i \sum_{j=i}^{m} (m + n - b_m)/(m + n - b_j)\binom{m+n-b_j}{m-j}\binom{b_j}{j-i} \end{aligned}$$

and the equivalence of (3.7) and (4.1.1) follows from the equalities

$$(4.1.2)\quad \textstyle\sum_{j=i}^{m} (m + n - b_m)/(m + n - b_j)\binom{m+n-b_j}{m-j}\binom{b_j}{j-i} = \binom{m+n}{m-i}, \qquad i = h, \cdots, m.$$

If $b_j = \eta + \rho j$, $j \geqq i \geqq h$, that is, if $b_h, b_{h+1}, \cdots, b_m$ form an increasing (or decreasing) arithmetic progression, then these equalities follow from a formula (see problem 16, page 169) in Riordan (1968) using the substitutions (Riordan's symbol given first): $k \to n - m + j$, $n \to m - i$, $\alpha = p \to \eta + \rho i$, $q = \beta \to \rho$, $\gamma \to m + n - b_m$.

The question now arises as to when these $b$'s are in arithmetic progression so that (4.1.1) is exact. This is a difficult question to answer in general, but (4.1.1) is exact at least for the cases:

(i) any $m$ and $n$ when $r \geqq mn - 2n + 1$, since $h \geqq m - 1$ and $b_{m-1}, b_m$ is an arithmetic progression in a trivial sense;

(ii) $n = mp + 1$, $r \equiv j(\text{mod } m)$, $1 \leqq j \leqq h$ since $[a_i] = i(p + 1) - [r/m] + [(i - j)/m]$ and the last term is zero for $h \leqq i \leqq m$ if $j \leqq h$;

(iii) $n = mp - 1$, $r \equiv -j(\text{mod } m)$, $0 \leqq j \leqq h - 1$ since $[a_i] = i(p + 1) - [r/m] - \langle (i - j)/m \rangle$ and the last term is $-1$ for $h \leqq i \leqq m$ if $j \leqq h - 1$.

4.2. *Formulae for $P(mnD^+ \geqq r)$ when $n = mp \pm 1$.* It is easy to show that the sequence $\{b_i\}$, $i = h, \cdots, m$, forms no more than two arithmetic progressions when $n = mp \pm 1$. In that case at most one increment in the sequence equals $(p + 1) \pm 1$ and the others are all equal to $p + 1$. For example, if $m = 5$, $n = 16$, $r = 15$ then $h = 1$ and the 5 $b$'s are 1, 5, 9, 13, 18; and if $m = 5$, $n = 14$, $r = 17$ then $h = 2$ and the 4 $b$'s are 4, 8, 11, 15.

Suppose now that $n = mp + 1$ and that $b_h, b_{h+1}, \cdots, b_{h+\alpha}$ form the lower arithmetic sequence $(0 \leqq \alpha \leqq m - h - 1)$ and $b_{h+\alpha+1}, \cdots, b_m$ form the upper. Let

$$E_i = b_i + 1, \qquad \text{if } h \leqq i \leqq h + \alpha;$$
$$= b_i, \qquad \text{otherwise.}$$

Then the sequence $\{E_i\}$, $i = h, \cdots, m$ is an arithmetic progression and

$$P(mnD^+ < r)$$
$$= P(R_i > b_i, \text{all } i)$$
$$= P(R_i > E_i, \text{all } i)$$
$$+ \textstyle\sum_{j=0}^{\alpha} P(R_i > b_i(i < h + j), R_{h+j} = E_{h+j}, R_i > E_i(i > h + j)).$$

Hence

$$(4.2.1)\quad \binom{m+n}{m} P(mnD^+ \geqq r) = \textstyle\sum_{j=h}^{m} (m + n - E_m)/(m + n - E_j)\binom{m+n-E_j}{m-j}\binom{E_j}{j} - \sum_{j=0}^{\alpha} Q_j$$

where, since sets of ranks above and below a given rank are conditionally independent and since the $b$'s involved with each set of ranks are in arithmetic progression,

$$(4.2.2)\quad Q_j = \{\tbinom{b_{h+j}}{h+j-1} - \textstyle\sum_{i=h}^{h+j-1} (b_{h+j} - b_{h+j-1})/(b_{h+j} - b_i)\tbinom{b_{h+j}-b_i}{h+j-i-1}\tbinom{b_i}{i}\}$$
$$\cdot\{\tbinom{m+n-E_{h+j}}{m-h-j} - \textstyle\sum_{i=1}^{m-h-j} (m+n-E_m)/(m+n-E_{h+j+i})$$
$$\cdot\tbinom{m+n-E_{h+j+i}}{m-h-j-i}\tbinom{E_{h+j+i}-E_{h+j}}{i}\}.$$

An analogous formula for $n = mp - 1$ would follow from

$$P(mnD^+ < r) = P(R_i > d_i\,,\ \text{all } i)$$
$$+ \textstyle\sum_{j=\alpha+1}^{m-h} P(R_i > d_i(i < h+j),\ R_{h+j} = d_{h+j}\,,$$
$$R_i > b_i(i > h+j)),$$

where

$$\begin{aligned} d_i &= b_i\,, && \text{if } h \leqq i \leqq h + \alpha; \\ &= b_i + 1, && \text{otherwise.} \end{aligned}$$

While apparently unrelated, (4.2.1) is reminiscent of a formula due to Hodges (1957) for the case $n = m + 1$. It is also possible to interpret the $Q_j$ in (4.2.1) as point probabilities for $D^+$. What happens is best illustrated by an example.

In Table I we give, for each $i \geqq h$, $t_i = i(m+n) - r$, $b_i = [t_i/m]$ (in bold face) and $E_i$ (in italics) for the case $m = 5$, $n = 16$, and $r = 17(1)24$.

Using $b_1 = 0$ plus the $b$'s given in Table I, we have, by Theorem 4.3,

$$\tbinom{21}{5}P(80D^+ < 17) = \begin{vmatrix} \binom{17}{1} & \binom{13}{2} & \binom{10}{3} & \binom{7}{4} & 0 \\ 1 & \binom{13}{1} & \binom{10}{2} & \binom{7}{3} & \binom{4}{4} \\ 0 & 1 & \binom{10}{1} & \binom{7}{2} & \binom{4}{3} \\ 0 & 0 & 1 & \binom{7}{1} & \binom{4}{2} \\ 0 & 0 & 0 & 1 & \binom{4}{1} \end{vmatrix} = 8053.$$

TABLE I

*Table of $t_i = i(m+n) - r$, $b_i = [t_i/m]$ (bold face) and $E_i$ (in italics) $i = h, \cdots, m$, for $m = 5$, $n = 16$, and $r = 17(1)24$*

| $i$ | $r$ = 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|
| 2 | 25 **5** *5* | 24 **4** *5* | 23 **4** *5* | 22 **4** *5* | 21 **4** *4* | 20 **4** *4* | 19 **3** *4* | 18 **3** *4* |
| 3 | 46 **9** *9* | 45 **9** *9* | 44 **8** *9* | 43 **8** *9* | 42 **8** *8* | 41 **8** *8* | 40 **8** *8* | 39 **7** *8* |
| 4 | 67 **13** *13* | 66 **13** *13* | 65 **13** *13* | 64 **12** *13* | 63 **12** *12* | 62 **12** *12* | 61 **12** *12* | 60 **12** *12* |
| 5 | 88 **17** *17* | 87 **17** *17* | 86 **17** *17* | 85 **17** *17* | 84 **16** *16* | 83 **16** *16* | 82 **16** *16* | 81 **16** *16* |

Similarly,

$$\binom{21}{5}P(80D^+ < 18) = 8613, \qquad \binom{21}{5}P(80D^+ < 19) = 9097, \qquad \binom{21}{5}P(80D^+ < 20) = 9657,$$

from which it follows that

$$(4.2.3) \quad \binom{21}{5}P(80D^+ = 17) = 560, \; \binom{21}{5}P(80D^+ = 18) = 484, \; \binom{21}{5}P(80D^+ = 19) = 560.$$

These values will be compared to those given by the $Q$'s.

Since the $\{b_i\}$ are in arithmetic progression for $r = 17$, 21, and 22, the first term on the right-hand side of (4.2.1) is the only term and $P(80D^+ \geqq 17)$ is given by Korolyuk's formula. Now when $r = 18$ then $\alpha = 0$ but the $E$'s have not changed and

$$\binom{21}{5}P(80D^+ \geqq 18) = \binom{21}{5}P(80D^+ \geqq 17) - Q_0 .$$

Thus

$$Q_0 = \binom{21}{5}P(80D^+ = 17).$$

But from Table I and (4.2.2) using $r = 18$, we find as a check that

$$Q_0 = 4\cdot\left\{\binom{16}{3} - \tfrac{4}{12}\binom{12}{2}\binom{4}{1} - \tfrac{4}{8}\binom{8}{1}\binom{8}{2} - \binom{4}{4}\binom{4}{0}\binom{12}{3}\right\} = 560$$

which equals the appropriate value in (4.2.3), as it should.

Similarly, when $r = 19$ then $\alpha = 1$, but again the $E$'s are the same and, moreover, $Q_0$ when $r = 19$ is equal to $Q_0$ when $r = 18$. Hence it follows that

$$Q_1 = \binom{21}{5}P(80D^+ = 18).$$

Again, from Table I and (4.2.2) using $r = 19$, we find

$$Q_1 = \left\{\binom{8}{2} - \tfrac{4}{4}\binom{4}{0}\binom{4}{2}\right\}\cdot\left\{\binom{12}{2} - \tfrac{4}{8}\binom{8}{1}\binom{4}{1}\right\} = 484$$

which checks with (4.2.3).

Finally, using $r = 20$ we can show that

$$Q_2 = \binom{21}{5}P(80D^+ = 19).$$

Again, by direct computation from (4.2.2) we find

$$Q_2 = \left\{\binom{12}{3} - \tfrac{4}{8}\binom{8}{1}\binom{4}{2} - \tfrac{4}{4}\binom{4}{0}\binom{8}{3}\right\}\cdot\left\{\binom{8}{1} - \binom{4}{1}\right\} = 560$$

which checks with (4.2.3).

As $r$ continues to increase, we find $P(80D^+ = 21) = 0$ since no $t_i/m$ is an integer and hence $P(80D^+ \geqq 21) = P(80D^+ \geqq 22)$. Both these probabilities are given by Korolyuk's formula with a new set of $E$'s. As $r$ continues to increase to 23 and 24 we find again $\alpha = 0$ and 1 and the new $Q_0$ and $Q_1$ will be $\binom{21}{5}P(80D^+ = 22)$ and $\binom{21}{5}P(80D^+ = 23)$, respectively.

The implication in the general case when $n = mp + 1$ is that Korolyuk's formula will give the required probability when $r \equiv j(\text{mod } m)$ and $j$ does not exceed $h$, and that as $r$ increases from such a value, say $r_0$, to a value such that $r \equiv h + i + 1(\text{mod } m)$, then

$$Q_i = P(mnD^+ = r_0 + i).$$

**5.0. Application to the null distributions of related statistics.**

5.1. *The $V_{mn}$ statistic.* Computational problems associated with the small sample distribution of the $V_{mn}$ statistic when $m \neq n$ and reported on by Maag and Stephens (1968) can be simplified by applying the above results.

In our notation

$$mnV_{mn} = mn\,D_{mn}^+ + mn\,D_{mn}^- = n + \sup_k \{(m+n)k - mR_k\} + \sup_k (mR_k - (m+n)k\}$$

which expresses $V_{mn}$ in terms of the $R_k$ only. Also, following Maag and Stephens, we can write

$$\begin{aligned} P(mnV_{mn} < r) &= \textstyle\sum_{a=0}^{r-1} P(mn\,D^+ = a, mn\,D^- < r - a) \\ &= \textstyle\sum_{a=0}^{r-1} P(mn\,D^+ < a + 1, mn\,D^- < r - a) \\ &\quad - \textstyle\sum_{a=1}^{r-1} P(mn\,D^+ < a, mn\,D^- < r - a). \end{aligned}$$

By Theorem 4.2

$$\binom{m+n}{m}P(mnV_{mn} < r) = \textstyle\sum_{a=0}^{r-1} |d_{ij}(a)| - \sum_{a=1}^{r-1} |d_{ij}^*(a)| \tag{5.1.1}$$

where, with or without the asterisk,

$$d_{ij}(a) = \begin{cases} 0, & \text{if } i - j > 1 \text{ or } c_i - b_j \leqq 1; \\ \binom{c_i - b_j + j - i - 1}{j - i + 1}, & \text{otherwise,} \end{cases}$$

and

$$b_i = \max(i - 1, [(i(m+n) - a - 1)/m]), \qquad b_i^* = \max(i - 1, [(i(m+n) - a)/m])$$

$$c_i = \min(n + i + 1, \langle(i(m+n) - n + r - a)/m\rangle), \qquad c_i^* = c_i.$$

The computation of the right-hand side of (5.1.1) is simplified by noting that $b_i$ and $b_i^*$ are frequently equal and by noting that one needs to sum only over those values of $a$ which are possible values of $mn\,D^+$. Let $g$ denote the greatest common divisor of $m$ and $n$ and consider the following cases.

Case I. $g = 1$. In this case $m$ and $n$ are relatively prime and $b_i \neq b_i^*$ if and only if the solution in $i$, say $i_0$, to $ni \equiv a(\text{mod } m)$ is greater than or equal to $\langle a/n\rangle$. This congruence has exactly one solution for every $a$. The pertinent result is: $na \equiv b(\text{mod } m)$ has a solution in $a$ if and only if $g$ divides $b$, in which case there are exactly $g$ solutions (see, for example, Jones (1955), page 46). Thus for every $a$, $b_i$ differs from $b_i^*$ for only one value of $i$ and the matrices $(d_{ij}(a))$ and $(d_{ij}^*(a))$ differ only in one column. Hence $|d_{ij}(a)| - |d_{ij}^*(a)| = |d_{ij}(a) - d_{ij}^*(a)|$ for every $a$. In order to tell which $a$'s are possible values of $mn\,D^+$, it is convenient to make a table of $a$, $i_0(a)$ and $\langle a/n\rangle$. The possible values of $mn\,D^+$ are those values of $a$ for which $i_0(a) \geqq \langle a/n\rangle$.

Hence, when $m$ and $n$ are relatively prime, we have expressed $\binom{m+n}{m}P(mnV_{mn} < r)$ as the sum of at most $r$ $\mu \times \mu$ determinants where $\mu = \min(m, n)$. While it is still prohibitive to compute the distribution of $V_{mn}$ by hand this way except for $\mu = 3$ or perhaps 4, there is a significant reduction for larger $\mu$ so that computing the distribution up to say $\mu = 25$ is not out of the question.

Case II. $g > 1$. In this case the possible values of $mn\,D^+$ and $mnV$ are integer multiples of $g$, and hence we can restrict ourselves to values of $r$ and $a$ which are integer multiples of $g$. Here, however, the congruence $ni \equiv a(\text{mod } m)$ has $g$ solutions when $g$ divides $a$ so that, in general $b_i \neq b_i^*$ for more than one value of $i$ which means we cannot subtract the matrices and then take the determinant of the difference to get the difference of the determinants.

In this case $\binom{m+n}{m}P(mnV_{mn} < r)$ is expressed as the difference of two sums of no more than $r/g$ $\mu \times \mu$ determinants each.

5.2. *Vincze's statistics* $(D, R)$, $(D^+, R^+)$. Vincze (1957, 1959, 1961) considers the use of pairs of statistics $(D^+(n, n), R^+(n, n))$ and $(D(n, n), R(n, n))$ for the two-sample problem, where $R^+$ (or $R$) is defined as the smallest value of $i$ such that $D^+$ (or $D$) attains its supremum at $z = Z_i$. If it should be the case that $m$ and $n$ are relatively prime, what we have shown implies that whatever the value of $D$ or $D^+$, say $r$, then there is exactly one solution in $k$ to each of the equations

$$(m + n)k - mR_k = r$$

$$n/2 + |mR_k - (m + n)k + n/2| = r.$$

This in turn implies that in this case there is precisely one value of $R^+$ (or $R$) to be associated with $D^+$ (or $D$). Hence when $m$ and $n$ are relatively prime, any test based on the statistics $(D^+, R^+)$ or $(D, R)$ is equivalent to one based on $D^+$ or $D$ alone.

## REFERENCES

Anderson, T. W. (1962). On the distribution of the two-sample Cramér-von Mises criterion. *Ann. Math. Statist.* **33** 1148–1159.

Barton, D. E. and Mallows, C. L. (1965). Some aspects of the random sequence. *Ann. Math. Statist.* **36** 236–260.

Blackman, J. (1956). An extension of the Kolmogorov distribution. *Ann. Math. Statist.* **27** 513–520.

Blackman, J. (1958). Correction to "An extension of the Kolmogorov distribution." *Ann. Math. Statist.* **29** 318–324.

Borovkov, A. A., Markova, N. P., and Sycheva, H. M. (1964). *Tables for N. V. Smirnov's two sample test of equality.* (in Russian). Akad. Nauk SSSR, Novosibirsk.

Capon, J. (1965). On the asymptotic efficiency of the Kolmogorov-Smirnov test. *J. Amer. Statist. Assoc.* **60** 843–853.

DARLING, D. A. (1957). The Kolmogorov-Smirnov, Cramér-von Mises tests. *Ann. Math. Statist.* **28** 823–838.

DEPAIX, M. (1962). Distributions de déviations maximales bilatérales entre deux echantillons independents de même loi continue. *Comptes Rendues Acad. Sci. Paris* **255** 2900–2902.

DRION, E. F. (1952). Some distribution-free tests for the difference between two empirical cumulative distribution functions. *Ann. Math. Statist.* **23** 563–574.

GNEDENKO, B. V. and KOROLYUK, V. S. (1951). On the maximum discrepancy between two empirical distributions. (in Russian). *Doklady Akad. Nauk SSSR* **80** 525–528.

HÁJEK, J. and SIDÁK, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

HODGES, J. L., JR. (1957). The significance probability of the Smirnov two-sample test. *Ark. Mat.* **3** 469–486.

JONES, B. W. (1955). *The Theory of Numbers.* Rinehart, New York.

KEMPERMAN, J. H. B. (1957). Some exact formulae for the Kolmogorov-Smirnov distributions. *Nederl. Akad. Wentensch. Proc. Ser. A* **60** 535–540.

KIM, P. J. and JENNRICH, R. I. (1967). Tables of the exact sampling distribution of the two-sample Kolmogorov-Smirnov criterion $D_{mn}(m \leqq n \leqq 25)$. Unpublished.

KLOTZ, J. (1967). Asymptotic efficiency of the two sample Kolmogorov-Smirnov test. *J. Amer. Statist. Assoc.* **62** 932–938.

KOROLYUK, V. S. (1955). On the deviation of empirical distributions for the case of two independent samples. (in Russian). *Izv. Akad. Nauk. SSSR Ser. Mat.* **19** 81–96.

LEHMANN, E. L. (1953). The power of rank tests. *Ann. Math. Statist.* **24** 23–43.

MAAG, U. R. and STEPHENS, M. A. (1968). The $V_{NM}$ two-sample test. *Ann. Math. Statist.* **39** 923–935.

MARLISS, G. S. and ZAYACHKOWSKI, W. (1962). The Smirnov two-sample statistic $D_{mn}(m \leqq n \leqq 20)$. Unpublished.

MASSEY, F. J., JR. (1950). A note on the power of a non-parametric test. *Ann. Math. Statist.* **21** 440–443.

MASSEY, F. J., JR. (1951). The distribution of the maximum deviation between two sample cumulative step functions. *Ann. Math. Statist.* **22** 125–128.

MASSEY, F. J., JR. (1952). Distribution table for the deviation between two sample cumulatives. *Ann. Math. Statist.* **23** 435–441.

RIORDAN, J. (1968). *Combinatorial Identities.* Wiley, New York.

SHORACK, R. A. (1967). Tables of the distribution of the Mann-Whitney-Wilcoxon U-statistic under Lehmann alternatives. *Technometrics* **9** 666–677.

SMIRNOV, N. (1939a). Sur les écarts de la courbe de distribution empirique. *Mat. Sb.* **6** 3–26.

SMIRNOV, N. (1939b). Estimating the deviation between the empirical distribution functions of two independent samples. (in Russian.) *Bull. Math. Univ. Moscow* **2** 3–16.

VINCZE, I. (1957). Einige Zweidimensionale Verteilungs und Grenverteilungssätze in der Theorie der geordneten Stichproben. *Publ. Math. Inst. Hungar. Acad. Sci.* **2** 183–209.

VINCZE, I. (1959). On some joint distributions and joint limiting distributions in the theory of order statistics. *Publ. Math. Inst. Hungar. Acad. Sci.* **4** 29–47.

VINCZE, I. (1961). On two-sample tests based on order statistics. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1** 695–705. Univ. of California Press.

WALD, A. and WOLFOWITZ, J. (1939). Confidence limits for continuous distribution functions. *Ann. Math. Statist.* **10** 105–118.