The Annals of Probability
1974, Vol. 2, No. 1, 155-160

A NEW FORMULA FOR $P(R_i \leqq b_i,\ 1 \leqq i \leqq m \mid m, n, F = G^k)$[1]

By G. P. Steck[2]

Sandia Laboratories

Let $X_1 \leqq X_2 \leqq \cdots \leqq X_m$ and $Y_1 \leqq Y_2 \leqq \cdots \leqq Y_n$ be independent samples of i.i.d. random variables from continuous distributions F and G, respectively, and suppose $F(x) = [G(x)]^k$ or $F(x) = 1 - [1 - G(x)]^k$, $k > 0$. Let R_i and S_j denote the ranks of X_i and Y_j, respectively, in the ordered combined sample. We express $P(R_i \leqq b_i$, all $i)$ as the determinant of a simple $m \times m$ matrix. We also show that for increasing sequences $\{a_i\}$ and $\{b_i\}$, $P(a_i \leqq R_i \leqq b_i$, all $i \mid F, G) = P(\alpha_j \leqq S_j \leqq \beta_j$, all $j \mid F, G)$, where $\{\alpha_j\} = \{b_i\}^c$ and $\{\beta_j\} = \{a_i\}^c$ and complementation is with respect to the set $\{i \mid 1 \leqq i \leqq m + n\}$, for *any* pair of continuous distributions F and G.

1. Introduction. Let $X_1 \leqq X_2 \leqq \cdots \leqq X_m$ and $Y_1 \leqq Y_2 \leqq \cdots \leqq Y_n$ be independent samples of i.i.d. random variables from continuous distributions F and G, respectively, and suppose $F(x) = [G(x)]^k$, $k > 0$. Let R_i and S_j denote the ranks of X_i and Y_j, respectively, in the ordered combined sample. We will prove for $k > 0$

Theorem 1.

$$(1)\qquad P(R_i \leqq b_i,\ 1 \leqq i \leqq m \mid m, n, F = G^k) = \frac{n!}{(n+km)!} \det\left\{\binom{j}{j-i+1} \frac{\Gamma(\theta_i + kj)}{\Gamma(\theta_i + ki - k)}\right\}_{m \times m},$$

where $\{b_i\}$ is an increasing sequence of integers and $\theta_i = b_i - i + 1$.

This is a generalization of a result appearing in Steck ((1969) (3.1.1)). These two results differ in notation and structure but the one given here seems considerably simpler. Since it depends only on k and not on F and G separately we will assume G to be the uniform distribution on $[0, 1]$.

The principal use of the results in this paper will probably be in carrying out one-sided Smirnov-type tests of hypotheses involving the distributions F, F^θ, $1 - (1 - F)^\theta$. The following remarks show that these are useful classes of distributions.

Harte and Pfanzagl (1969) have a biological problem that requires testing $H: G = 1 - (1 - F)^k$ against $A: G > 1 - (1 - F)^k$ (their inequality (2) should be reversed) for k a known integer. The essence of their problem is the following: let the X's be the times required for individual students to solve a

Received January 31, 1973; revised May 10, 1973.

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

[2] Most of this work was done while the author was on sabbatical leave as a Visiting Professor at Oregon State University.

AMS 1970 *subject classifications.* Primary 62E15; Secondary 62G99.

Key words and phrases. Distribution of ranks, Lehmann alternatives.

certain statistics problem and let the Y's be the times required for independent groups of k students each to solve the problem. The question is—is there collaboration among members of a group? If not, $G = 1 - (1 - F)^k$; if so, $G > 1 - (1 - F)^k$ (assuming collaboration aids solution). Harte and Pfanzagl use the Wilcoxon test but a Smirnov type test is reasonable, too. In this example our results are needed to determine the critical region.

Shorack (1967) furnishes another example. Suppose something fails when the last of k bonds fails, where k is not known exactly but is representative of the manufacturer—obviously large k's are desirable. If the bonds fail in an i.i.d. fashion according to a distribution F, the time-to-failure distribution is F^k. Suppose manufacturer A, with $k = \alpha$ and $G_A = F^\alpha$, has been the accepted supplier but that now there is a second manufacturer B, with $k = \beta$ and $G_B = F^\beta = (F^\alpha)^{\beta/\alpha}$. To test the hypothesis that both manufacturers are equally good against the alternative that B is better is to test $H\colon G = \tilde{F}$ against A: $G = \tilde{F}^\theta$, $\theta > 1$. In this example our results are needed to determine the power of Smirnov-type tests.

Finally, we note that Allen (1963) proved a characterization theorem for these classes of distribution: if F and G are absolutely continuous with common support and associated hazard functions $r(t)$ and $s(t)$, then the statements (i) $r(t) = \theta s(t)$, (ii) $1 - G = (1 - F)^\theta$, (iii) $P(X < Y \mid X < t)$ independent of t are equivalent. A similar result follows for $G = F^\theta$ if t is replaced by $-t$.

2. Proof of Theorem 1. Let E_t denote expectation with respect to the joint distribution of the t distinct random variables among the collection $Y_{\theta_1}, Y_{\theta_2}, \cdots, Y_{\theta_m}$ (not counting Y_0 or Y_{n+1} which we take as identically 0 and 1, respectively). Then we have

LEMMA 1.

$$P(R_i \leqq b_i, \quad all \quad i \mid m, n, F = G^k) = E_t[\det\{\tbinom{j}{j-i+1} Y_{\theta_i}^{k(j-i+1)}\}_{m\times m}]\,.$$

PROOF. Since $R_i \leqq b_i$ if and only if $X_i < Y_{b_i-i+1}$, we have $P(R_i \leqq b_i, \text{ all } i) = P(X_i < Y_{\theta_i}, \text{ all } i)$. Let the t distinct random variables among $\{Y_{\theta_i}\}$ be $U_1, U_2, \cdots, U_t$. Then

$$P(X_i < Y_{\theta_i}, \quad \text{all} \quad i) = \int_{0\leqq u_1\leqq\cdots\leqq u_t\leqq 1}\cdots\int P(X_i < y_{\theta_i}, \quad \text{all} \quad i)\, dF_{U_1,U_2,\cdots,U_t}(u_1, u_2, \cdots, u_t)\,.$$

Since $F(X) = X^k$ is $U(0, 1)$ the probability inside the integral equals $P(V_i < y_{\theta_i}^k, \text{ all } i)$ for uniform order statistics $V_1 \leqq V_2 \leqq \cdots \leqq V_m$ from a sample of m i.i.d. $U(0, 1)$ random variables. This probability is given by Steck (1971) as $P(V_i \leqq v_i, \text{ all } i) = \det\{\tbinom{j}{j-i+1} v_i^{\,j-i+1}\}_{m\times m}$ so that

$$P(X_i < Y_{\theta_i}, \quad \text{all} \quad i) = \int_{0\leqq u_1\leqq\cdots\leqq u_t\leqq 1}\cdots\int \det\{\tbinom{j}{j-i+1} y_{\theta_i}^{k(j-i+1)}\}\, dF_{U_1,\cdots,U_t}(u_1, \cdots, u_t)\,. \quad \square$$

We now evaluate the expectation in Lemma 1. Assume, for the moment, that all the numbers $\theta_1, \theta_2, \cdots, \theta_m$ are distinct. Let the random variables

$W_1, W_2, \cdots, W_m$ be defined by

$$
\begin{aligned}
Y_{\theta_1} &= W_1 W_2 \cdots W_m \\
Y_{\theta_2} &= \quad\; W_2 \cdots W_m \\
&\vdots \\
Y_{\theta_m} &= \qquad\qquad W_m .
\end{aligned}
$$

It is known that the W_i's are independently distributed beta variables and W_i is Beta $(\theta_i, \theta_{i+1} - \theta_i)$ with $\theta_{m+1} \equiv n + 1$—that is, $f_{W_i}(w) \propto w^{\theta_i - 1}(1 - w)^{\theta_{i+1} - \theta_i - 1}$. If one adopts the convention that for $r > 0$ a Beta $(r, 0)$ variable is identically one, then the same formulation works even though the numbers $\theta_1, \theta_2, \cdots, \theta_m$ are not all distinct. For example, if $\theta_h = \theta_{h+1}$ so that we want $Y_{\theta_h} \equiv Y_{\theta_{h+1}}$, then W_h is Beta $(\theta_n, 0)$ and $P(W_h = 1) = 1$. This leads to $Y_{\theta_h} \equiv Y_{\theta_{h+1}}$, as desired, since $Y_{\theta_h} = W_h W_{h+1} \cdots W_m = W_{h+1} \cdots W_m = Y_{\theta_{h+1}}$.

When expressed in terms of the $\{W_i\}$, Lemma 1 gives

$$
\begin{aligned}
(2) \qquad P(R_i \leqq b_i, \text{ all } i \mid m, n, F = G^k) \\
= E_t[\det \{\tbinom{j}{j-i+1}(W_i W_{i+1} \cdots W_m)^{k(j-i+1)}\}_{m \times m}] .
\end{aligned}
$$

Expanding the determinant on the RHS of (2) by its last column and inducting on the dimensionality using the fact $a_{ij} = 0$ for $i > j + 1$ implies $\det \{a_{ij} x^{j-i+1}\}_{m \times m} = x^m \det \{a_{ij}\}_{m \times m}$ (the problem is more notational than conceptual) shows that the RHS of (2) is a determinant whose entry in row i and column j is

$$
\begin{array}{ll}
\binom{j}{j-i+1} \prod_{q=i}^{m} EW_q^{\,kj} / EW_q^{\,k(i-1)} & j \geqq i - 1 \\
0 & \text{otherwise} .
\end{array}
$$

Since $EZ^t = \Gamma(r + s)\Gamma(r + t)/[\Gamma(r)\Gamma(r + s + t)]$ for Z a Beta (r, s) random variable, and since W_q is Beta (r_q, s_q) with $r_q = \theta_q$, $s_q = \theta_{q+1} - \theta_q$, $r_q + s_q = r_{q+1}$ we have, after much cancelling of factors,

$$
\begin{aligned}
P(R_i \leqq b_i, \text{ all } i \mid m, n, F = G^k) \\
= \det \left\{ \binom{j}{j - i + 1} \frac{\Gamma(\theta_i + kj)\Gamma(n + k(i-1) + 1)}{\Gamma(\theta_i + k(i-1))\Gamma(n + kj + 1)} \right\}_{m \times m} .
\end{aligned}
$$

A trivial induction shows the factors involving n can be brought to the outside of the determinant as the single multiplier $n!/(n + km)!$. Similarly the factors $j!/(i - 1)!$ can be brought outside the determinant as the factor $m!$. This completes the proof of (1). □

Note that if $k = 1$ then

$$\binom{m+n}{m} P(R_i \leqq b_i \mid F = G) = \det \{\tbinom{b_i - i + j}{j - i + 1}\}_{m \times m} .$$

This is a different expression from the one given by Steck ((1969) Theorem 4.1) which is

$$\binom{m+n}{m} P(R_i \leqq b_i \mid F = G) = \det \{\tbinom{b_i - i + 1}{j - i + 1}\}_{m \times m} ;$$

however, it is not difficult to show by repeated adding of adjacent columns that they are equivalent expressions.

If $k > 1$ then X is stochastically greater than Y and the X's will occur generally later than the Y's in the ordered combined sample. One reasonable test of the hypothesis $H: F = G$ against the alternative $A: F = G^k$, $k > 1$, is to reject if the X's are "too late" in the ordered combined sample in the sense that $R_i \geqq b_i$, all i. This means that if $k > 1$ we also interested in probabilities of the form $P(R_i \geqq b_i, \text{ all } i \mid m, n, F = G^k)$. Interchanging the roles of X and Y we see that this probability is the same as $P(S_j \geqq b_j, \text{ all } j \mid n, m, F = G^{1/k})$. To express it in terms of what we already know, we need the following theorem.

3. Relating the events $a_i \leqq R_i \leqq b_i$, all i and $\alpha_j \leqq S_j \leqq \beta_j$, all j. It is obvious that $R_i = b_i$, all i, is equivalent to $S_j = \beta_j$, all j, where $\beta_1 < \beta_2 < \cdots < \beta_n$ are the integers left over when $b_1, b_2, \cdots, b_m$ are deleted from the set of integers $1, 2, \cdots, m + n$. What follows shows that something very similar holds for the events $a_i \leqq R_i \leqq b_i$, all i, and $\alpha_j \leqq S_j \leqq \beta_j$, all j.

THEOREM 2. *Let $\alpha_1 < \alpha_2 < \cdots < \alpha_n$ be an increasing sequence such that $j \leqq \alpha_j \leqq m + j$. Then $\{\text{sample orderings} \mid R_i \leqq b_i, \text{ all } i\} = \{\text{sample orderings} \mid S_j \geqq \alpha_j, \text{ all } j\}$ where $\{b_i\} = \{1, 2, m + n\} - \{\alpha_j\} \equiv \{\alpha_j\}^c$ with $b_i < b_{i+1}$ and $i \leqq b_i \leqq n + i$.*

SKETCH OF PROOF. Since $S_j \geqq \alpha_j \Leftrightarrow R_{\alpha_j - j} \leqq \alpha_j - 1$, the set of orderings such that $S_j \geqq \alpha_j$, $1 \leqq j \leqq n$ is equal to the set of orderings such that $R_{\alpha_j - j} \leqq \alpha_j - 1$, $1 \leqq j \leqq n$. Since $j \leqq \alpha_j \leqq m + j$, we have $0 \leqq \alpha_j - j \leqq m$. Let a_i be the number of times a number $\leqq i$ appears in the set $\{\alpha_j - j \mid j = 1, 2, \cdots, n\}$, for $i = 0, 1, \cdots, m$. Clearly $a_m = n$ and $0 \leqq a_i \leqq a_{i+1} \leqq n$. Then for $i = 0, 1, 2, \cdots, m$ we have the following relations and implications (put $a_{-1} = 0$)

$$\begin{aligned}
\alpha_{a_{i-1}+1} - (a_{i-1} + 1) = i &\Longrightarrow S_{a_{i-1}+1} \geqq a_{i-1} + i + 1 \wedge R_i \leqq a_{i-1} + i \\
&\ \ \vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots \\
a_{a_i} - a_i = i &\Longrightarrow \quad S_{a_i} \geqq a_i + i \qquad \wedge R_i \leqq a_i + i - 1 \\
&\Longrightarrow R_i \leqq a_{i-1} + i\,, \\
&\quad \vdots \\
&\Longrightarrow R_i \leqq a_{i-1} + i\,.
\end{aligned}$$

Regardless of the values of the a_i, it is clear that $\{a_{i-1} + i \mid i = 1, 2, \cdots, m\} = \{\alpha_j\}^c$ where complementation is respect to the set $\{i \mid 1 \leqq i \leqq m + n\}$. Putting $b_i = a_{i-1} + i$ completes the sketching of the proof since $\{a_{i-1} + i\}$ is an increasing sequence and since $i \leqq a_{i-1} + i \leqq n + i$. □

COROLLARY. *If $\{\alpha_j\}$ and $\{\beta_j\}$ are increasing sequences with $j \leqq \alpha_j \leqq \beta_j \leqq m + j$, then $\{\text{sample orderings} \mid \alpha_j \leqq S_j \leqq \beta_j\} = \{\text{sample orderings} \mid a_i \leqq R_i \leqq b_i\}$ where $\{a_i\} = \{\beta_j\}^c$, $\{b_i\} = \{\alpha_j\}^c$ are increasing sequences with $i \leqq a_i \leqq b_i \leqq n + i$.*

COROLLARY. *$P(R_i \geqq b_i, 1 \leqq i \leqq m \mid m, n, F = G^k) = P(R_i \leqq a_i, 1 \leqq i \leqq n \mid n, m, F = G^{1/k})$ where $\{a_i\} = \{b_i\}^c$.*

A proof very similar to that of Theorem 1 will prove

THEOREM 3.

$$P(R_i \geqq b_i,\ 1 \leqq i \leqq m \,|\, m, n, F = 1 - (1 - G)^k)$$
$$= \frac{n!}{(n + km)!} \det \left\{ \binom{j}{j - i + 1} \frac{\Gamma(\varphi_j + k(m - i + 1))}{\Gamma(\varphi_j + k(m - j))} \right\}_{m \times m},$$

where $\{b_i\}$ *is an increasing sequence of integers and* $\varphi_j = n - b_j + j + 1$.

COROLLARY.

$$P(R_i \leqq b_i,\ 1 \leqq i \leqq m \,|\, m, n, F = 1 - (1 - G)^k)$$
$$= P(R_i \geqq a_i,\ 1 \leqq i \leqq n \,|\, n, m, F = 1 - (1 - G)^{1/k}),$$

where $\{a_i\} = \{b_i\}^c$.

4. Examples. Take $m = 2$, $n = 3$, $k = 3$. Then from Lehmann (1953) we have the following table of probabilities.

TABLE 1.
$1680 \cdot P(R_1 = r_1, R_2 = r_2 \,|\, F, G)$

(r_1, r_2)	(1,2)	(1,3)	(1,4)	(1,5)	(2,3)	(2,4)	(2,5)	(3,4)	(3,5)	(4,5)
$F = G^3$	20	30	42	56	90	126	168	252	336	560
$F = 1 - (1 - G)^3$	560	336	168	56	252	126	42	90	30	20

Note that

$$P(R_1 = r_1, R_2 = r_2 \,|\, F = G^3) = P(R_1 = 6 - r_2, R_2 = 6 - r_1 \,|\, F = 1 - (1 - G)^3).$$

EXAMPLE 1.

$$P(R_1 \leqq 2, R_2 \leqq 4 \,|\, 2, 3, F = G^3) = \frac{3!}{9!} \begin{vmatrix} \frac{\Gamma(5)}{\Gamma(2)} & \frac{\Gamma(8)}{\Gamma(2)} \\ 1 & \frac{2\Gamma(9)}{\Gamma(6)} \end{vmatrix} = \frac{11}{60} = \frac{308}{1680}.$$

This answer is also obtained from the table as 20 + 30 + 42 + 90 + 126.

EXAMPLE 2.

$$P(R_1 \geqq 2, R_2 \geqq 4 \,|\, 2, 3, F = G^3)$$
$$= P(R_1 \leqq 1, R_2 \leqq 3, R_3 \leqq 5 \,|\, 3, 2, F = G^{\frac{1}{3}})$$
$$= \frac{2!}{3!} \begin{vmatrix} \frac{\Gamma(\frac{4}{3})}{\Gamma(1)} & \frac{\Gamma(\frac{5}{3})}{\Gamma(1)} & \frac{\Gamma(\frac{6}{3})}{\Gamma(1)} \\ 1 & \frac{2\Gamma(\frac{8}{3})}{\Gamma(\frac{7}{3})} & \frac{3\Gamma(\frac{9}{3})}{\Gamma(\frac{7}{3})} \\ 0 & 1 & \frac{3\Gamma(\frac{12}{3})}{\Gamma(\frac{11}{3})} \end{vmatrix} = \frac{1}{3} \begin{vmatrix} \frac{\Gamma(\frac{1}{3})}{3} & \frac{2\Gamma(\frac{2}{3})}{3} & 1 \\ 1 & \frac{5\Gamma(\frac{2}{3})}{\Gamma(\frac{1}{3})} & \frac{27}{2\Gamma(\frac{1}{3})} \\ 0 & 1 & \frac{243}{40\Gamma(\frac{2}{3})} \end{vmatrix}$$
$$= \frac{103}{120} = \frac{1442}{1680}.$$

This answer is also obtained from the table as 126 + 168 + 252 + 336 + 560.

It is left to the reader to use Theorem 3 and its corollary to verify

EXAMPLE 3. $1680 \cdot P(R_1 \geqq 3, R_2 \geqq 4 \mid 2, 3, F = 1 - (1 - G)^3) = 140.$

EXAMPLE 4. $1680 \cdot P(R_1 \leqq 1, R_2 \leqq 4 \mid 2, 3, F = 1 - (1 - G)^3) = 1064.$

5. Applications to the distributions of one-sided Smirnov and Rényi statistics. The one-sided Smirnov statistics are $D_{mn}^+ = \sup_x [F_m(x) - G_n(x)]$ and $D_{mn}^- = \sup_x [G_n(x) - F_m(x)]$. Steck (1969) shows $P(mnD_{mn}^+ \leqq r) = P(R_i \geqq a_i, 1 \leqq i \leqq m)$ and $P(mnD_{mn}^- \leqq s) = P(R_i \leqq b_i, 1 \leqq i \leqq m)$ where $a_i = \langle\{i(m + n) - r\}/m\rangle$ and $b_i = [\{i(m + n) - n + s\}/m]$ with $[x]$ = largest integer $\leqq x$ and $\langle x\rangle$ = smallest integer $\geqq x$. Thus Theorem 1 together with the corollary to Theorem 3 give the distribution of D_{mn}^- for $F = G^k$ and $1 - (1 - G)^k$, $k > 0$, respectively. The corollary to Theorem 1 together with Theorem 3 do the same for D_{mn}^+.

The one-sided Rényi statistic is $R_t^+ = \sup \{N[F_m(x) - G_n(x)]/[mF_m(x) + nG_n(x)]\}$ where the supremum is taken over those x for which $mF_m(x) + nG_n(x) \geqq t$, $0 < t \leqq m + n \equiv N$. This is a modification of D_{mn}^+ which gives more weight to differences occurring for small X's and Y's. Since R_t^+ cannot increase unless $x = X_k$ for some k, it follows that

$$R_t^+ = \max\left[(m + n)\left\{\frac{k}{m} - \frac{R_k - k}{n}\right\}\Big/R_k\right]$$

where the maximum is over those k for which $R_k \geqq t$. Consequently, it can be shown that $P(R_t^+ \leqq a(m + n)/mn) = P(R_k \geqq c_k, 1 \leqq k \leqq m)$ where $c_k = k$ for $k < k_0$ and $c_k = k(m + n)/(m + a)$ for $k \geqq k_0$ with $k_0 = \langle t(m + a)/(m + n)\rangle$. Hence the remarks concerning the distribution of D_{mn}^+ apply to R_t^+ as well.

REFERENCES

[1] ALLEN, W. R. (1963). A note on conditional probabilities when hazards are proportional. *Operations Res.* **11** 658–659.

[2] HARTE, C. and PFANZAGL, J. (1969). A nonparametric comparison between two samples referring to individuals and clusters. *Biom. Zeit.* **11** 217–227.

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

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

[5] STECK, G. P. (1969). The Smirnov two sample tests as rank tests. *Ann. Math. Statist.* **40** 1449–1466.

[6] STECK, G. P. (1971). Rectangle probabilities for uniform order statistics and the probability that the empirical distribution function lies between two distribution functions *Ann. Math. Statist* **42** 1–11.

NUMERICAL MATHEMATICS DIVISION 5122
SANDIA LABORATORIES
SANDIA CORPORATION
ALBUQUERQUE, NEW MEXICO 87115