The Annals of Mathematical Statistics
1971, Vol. 42, No. 1, 413–414

A SHORT PROOF OF STECK'S RESULT ON TWO-SAMPLE SMIRNOV STATISTICS

BY S. G. MOHANTY

McMaster University

Let there be two independent random samples of sizes m and n respectively from a continuous population. Let R_i denote the ranks of the first sample in the ordered combined sample. Suppose $b = (b_1, b_2, \cdots, b_m)$ and $c = (c_1, c_2, \cdots, c_m)$ are two increasing sequences of integers such that $i-1 \leqq b_i \leqq c_i \leqq n+i+1$. Denote by $N(b; c)$ the number of ways the event $\{b_i < R_i < c_i, i = 1, 2, \cdots, m\}$ can occur in the ordered combined sample. It is well recognized that $N(b; c)$ determines the null distribution of Smirnov statistics in the two-sample case. In a recent paper [1], Steck has established that

$$N(b; c) = \det(d_{ij})_{m\times m} \tag{1}$$

where

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

Here
$$\binom{y}{z}_+ = 0 \quad \text{if} \quad z \neq 0 \quad \text{and} \quad y < z \quad \text{or if} \quad z < 0$$
$$= 1 \quad \text{if} \quad z = 0.$$

His proof, which is long and difficult to follow, consists of checking that the determinant satisfies the recurrence relations and boundary conditions required by $N(b; c)$. In this note, we provide a direct elementary proof.

Let $T_i = R_i - i$, $u_i = b_i - i + 1$ and $v_i = c_i - i - 1$. Then

$$\{b_i < R_i < c_i, i = 1, 2, \cdots, m\} \Leftrightarrow \{u_i \leqq T_i \leqq v_i, i = 1, 2, \cdots, m\}$$

and therefore $N(b; c)$ represent the number of vectors $(x_1, x_2, \cdots, x_m)$ which satisfies the following:

(i) x_i's are integers,
(ii) $0 \leqq x_1 \leqq x_2 \leqq \cdots \leqq x_m \leqq n$,
(iii) $u_i \leqq x_i \leqq v_i$ $i = 1, 2, \cdots, m$.

Clearly, we can write

$$N(b; c) = \sum_{x_1=u_1}^{v_1} \sum_{x_2=y_2}^{v_2} \cdots \sum_{x_{m-1}=y_{m-1}}^{v_{m-1}} \sum_{x_m=y_m}^{v_m} 1 \tag{2}$$

with $y_i = \max(u_i, x_{i-1})$, $i = 2, 3, \cdots, m$. Note that

$$\begin{vmatrix} \binom{x_m - u_m}{0}_+ & \binom{x_{m-1} - u_m}{1}_+ & \cdots & \binom{x_1 - u_m}{m-1}_+ \\ 0 & \binom{x_{m-1} - u_{m-1}}{0}_+ & \cdots & \binom{x_1 - u_{m-1}}{m-2}_+ \\ 0 & 0 & \cdots & \binom{x_1 - u_{m-2}}{m-3}_+ \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \binom{x_1 - u_1}{0}_+ \end{vmatrix} = 1. \tag{3}$$

Received May 25, 1970; revised September 2, 1970.

Furthermore, it is easy to verify that for $u_j \geqq u_i$ and $r \geqq 0$,

$$\sum_{x_i=y_i}^{v_i}\binom{x_i-u_j}{r}_+ = \binom{v_i-u_j+1}{r+1}_+ - \binom{x_{i-1}-u_j}{r+1}_+ \tag{4}$$

for each of the two cases $y_i = u_i$ and $y_i = x_{i-1}$. Replacing 1 in the right-hand side of (2) by the determinant in (3) and using (4), we see that

$N(b;c) =$

$$\sum_{x_1=u_1}^{v_1}\sum_{x_2=y_2}^{v_2}\cdots\sum_{x_{m-1}=y_{m-1}}^{v_{m-1}}\begin{vmatrix} \binom{v_m-u_m+1}{1}_+ - \binom{x_{m-1}-u_m}{1}_+ & \binom{x_{m-1}-u_m}{1}_+ & \cdots & \binom{x_1-u_m}{m-1}_+ \\ 1-1 & \binom{x_{m-1}-u_{m-1}}{0}_+ & \cdots & \binom{x_1-u_{m-1}}{m-2}_+ \\ 0 & 0 & \cdots & \binom{x_1-u_{m-2}}{m-3}_+ \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \binom{x_1-u_1}{0}_+ \end{vmatrix}$$

$$= \sum_{x_1=u_1}^{v_1}\sum_{x_2=y_2}^{v_2}\cdots\sum_{x_{m-1}=y_{m-1}}^{v_{m-1}}\begin{vmatrix} \binom{v_m-u_m+1}{1}_+ & \binom{x_{m-1}-u_m}{1}_+ & \cdots & \binom{x_1-u_m}{m-1}_+ \\ 1 & \binom{x_{m-1}-u_{m-1}}{0}_+ & \cdots & \binom{x_1-u_{m-1}}{m-2}_+ \\ 0 & 0 & \cdots & \binom{x_1-u_{m-2}}{m-3}_+ \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \binom{x_1-u_1}{0}_+ \end{vmatrix}.$$

With the help of (4), when the summation is continued to the end, we get

$$N(b,c) = \det(d_{ij}^*)_{m\times m}$$

where

$$d_{ij}^* = \binom{v_{m-j+1}-u_{m-i+1}+1}{j-i+1}_+ = \binom{c_{m-j+1}-b_{m-i+1}+i-j-1}{j-i+1}_+ .$$

It is easily checked that $\det(d_{ij}^*) = \det(d_{ij})$, and hence the proof of (1) is complete.

Finally, we offer the following remark. If in the starting determinant (3) $\binom{x_i-u_j}{r}_+$ is replaced by $(x_i-u_j)_+{}^r/r!$ (where $(x)_+ = \max(x, 0)$) and if the multiple sum (2) is replaced by the multiple integral

$$\int_{u_1}^{v_1}\int_{y_2}^{v_2}\cdots\int_{y_m}^{v_m} dx_m \cdots dx_1 ,$$

then it is possible to adapt the above proof to prove Steck's result on the joint distribution of uniform order statistics [2].

REFERENCES

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

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