number so denoted in (2), then there is a non-decreasing function $\sigma$, which is not an $s$-function, and which satisfies (4).

Conversely, given the $n$ equations (4), then by Lemma 2 (ii) we can choose $t_n$ sufficiently large so that

$$t_k = \int x^k dg \quad (k = 0, \ldots, n), \tag{8}$$

where $g$ is non-decreasing and is not an $s$-function. By case $E$, (8) implies

$$s_k = \int x^k f(x)\, dx \quad (k = 0, \ldots, n),$$

where $f$ is not an $s$-function equal to 0 or 1, $0 \leqslant f \leqslant 1$, and the $s_k$ are defined by

$$\exp\left(\sum_0^n \frac{s_k}{z^{k+1}}\right) = 1 + \sum_0^n \frac{t_k}{z^{k+1}} + \ldots.$$

But then $s_0, \ldots, s_{n-1}$ are the same as in (2). Hence equations (3) are true.

4. *Correction to* I. There is a passage in I § 1 which, though it does not affect what is subsequently proved, may give rise to a misunderstanding. It is there implied that in order that the equations I (3) shall be satisfied, it is necessary and sufficient that

$$\sum_{i,k=0}^{m} t_{i+k} x_i x_k$$

shall be positive definite or semi-definite, where $2m$ is the greatest even integer which does not exceed $n-1$. This implication is incorrect; and the mistake arose from making small changes in the enunciation of Achyèser and Krein's results.

Queen's University
Belfast

# RANK CORRELATION AND PERMUTATION DISTRIBUTIONS

By P. A. P. MORAN

*Received* 17 April 1947

Suppose that we have $n$ objects denoted by $1, \ldots, n$, and suppose them to be rearranged so that the rank of the object $i$ is $p(i)$. Let $s$ be the least number of interchanges of neighbours required to restore the permutation to the normal order. Now let $u(n, s)$ be the number of distinct arrangements which require $s$ interchanges. Then Haden (1) has proved that

$$u(n,s) = \frac{6n!}{\{\pi n(n-1)(2n+5)\}^{\frac{1}{2}}}\left[\exp\left\{\frac{-36t^2}{n(n-1)(2n+5)}\right\} + O\left(\frac{1}{n}\right)\right],$$

where $t = s - \frac{1}{4}n(n-1)$, and we have used the symbol $t$ where Haden uses $\sigma$, since the latter has statistical associations. This theorem is very similar to the theorem established by Kendall (2, 3) (see also Feller (4), Dantzig (5)), which states that his coefficient of rank correlation, $\tau$, between the order $1, 2, \ldots, n$ and $\{p(i)\}$, is such that

$$S = \tfrac{1}{2}n(n-1)\tau$$

is distributed about a zero mean with variance $\frac{1}{18}n(n-1)(2n+5)$ in a distribution

https://doi.org/10.1017/S0305004100024087 Published online by Cambridge University Press

which tends to normality as $n$ increases. The purpose of the present note is to point out that*

$$s = \tfrac{1}{4}n(n-1) - \tfrac{1}{2}S, \tag{1}$$

and so

$$\tau = 1 - \frac{4s}{n(n-1)} = -\frac{4t}{n(n-1)},$$

and Haden's theorem is therefore a strengthened form of Kendall's result. Dantzig (5) had already proved a theorem on interchanges of type similar to Haden's, but that of the latter is a stronger result because it applies to the individual ordinate, whereas Dantzig's theorem concerns the distributions as a whole. Moreover, (1) shows that Kendall's $\tau$ can be defined in terms of $s$, and so the theory of rank correlation can be linked up with the theory of permutations. To prove (1) we first prove $s \leqslant \frac{1}{4}n(n-1) - \frac{1}{2}S$, and then $s \geqslant \frac{1}{4}n(n-1) - \frac{1}{2}S$.

$p(i)$ is the rank of the object $i$ in the new arrangement. We define a function $m(i,j)$ such that

$$\begin{aligned} m(i,j) &= 1 \quad \text{if} \quad p(i) > p(j) \\ &= 0 \quad \text{if} \quad p(i) < p(j). \end{aligned}$$

In fact, we can write

$$m(i,j) = \frac{1}{2}\left\{\frac{p(i)-p(j)}{|\,p(i)-p(j)\,|} + 1\right\}.$$

Kendall's $\tau$ can be defined by writing

$$\begin{aligned} \tau = \frac{2S}{n(n-1)} &= \frac{2}{n(n-1)} \sum_{i>j} \frac{p(i)-p(j)}{|\,p(i)-p(j)\,|} \\ &= \frac{2}{n(n-1)} \sum_{i>j} (2m(i,j)-1). \end{aligned}$$

We first show that given the arrangement $\{p(i)\}$ we can reduce it to the normal order by $\frac{1}{4}n(n-1) - \frac{1}{2}S$ interchanges. The object 1 has rank $p(1)$. Transfer this to the extreme left of the arrangement. This will require $p(1)-1$ interchanges. This process of interchanges will have moved the object 2, $m(1,2)$ positions to the right. To transfer object 2 into the second position will therefore require $p(2)-2+m(1,2)$ interchanges. Similarly, we transfer objects 3, 4, 5, ... to the left until they are in their correct positions, the $i$th object requiring

$$p(i) - i + m(1,i) + m(2,i) + \ldots + m(i-1,i)$$

interchanges. The object $n$ will finally require

$$p(n) - n + m(1,n) + \ldots + m(n-1,n) = 0$$

interchanges. The total number of interchanges is then

$$\begin{aligned} &\sum_1^n p(n) - \tfrac{1}{2}n(n+1) + \sum_{i<j} m(i,j) \\ &\quad = \tfrac{1}{2}n(n+1) - \tfrac{1}{2}n(n+1) + \tfrac{1}{4}n(n-1) + \tfrac{1}{2}\sum_{i<j} \frac{p(i)-p(j)}{|\,p(i)-p(j)\,|} \\ &\quad = \tfrac{1}{4}n(n-1) - \tfrac{1}{2}S. \end{aligned}$$

We now show that

$$s \geqslant \tfrac{1}{4}n(n-1) - \tfrac{1}{2}S.$$

Let $T$ be a sequence of interchanges which reduces the arrangement $\{p(i)\}$ to the normal order. Each interchange in $T$ involves two objects. We classify the interchanges of $T$

https://doi.org/10.1017/S0305004100024087 Published online by Cambridge University Press

into $n$ groups $T_1, \ldots, T_n$ as follows. $T_1$ consists of all the interchanges which involve the object 1. $T_2$ consists of all the interchanges which involve the object 2 but not the object 1. Similarly, $T_i$ consists of all the interchanges which involve the object $i$ but not the objects $1, 2, \ldots, i-1$. $T_n$ is clearly empty. We now consider the number of interchanges in each of these groups. Let $A_i$ be the number of interchanges in $T_i$ which move the object $i$ to the right, and $B_i$ the number which move it to the left. The number of interchanges required in $T_1$ will be at least $p(1)-1$. These operations will move object 2, $m(1,2)$ places to the right. The net result of the operations in $T_2$ will be to move the object 2 $A_2 - B_2$ places to the right (taken algebraically). The total movement of the object 2 to the right is therefore

$$m(1,2) + A_2 - B_2,$$

and this must equal $2-p(2)$. So we have

$$B_2 - A_2 = p(2) - 2 + m(1,2),$$

and therefore

$$B_2 + A_2 \geqslant p(2) - 2 + m(1,2).$$

Similarly the operations in $T_1, \ldots, T_{i-1}$ will move the object $i$

$$m(1,i) + m(2,i) + \ldots + m(i-1,i)$$

places to the right and

$$m(1,i) + \ldots + m(i-1,i) + A_i - B_i$$

must equal $i - p(i)$. Therefore

$$B_i - A_i = p(i) - i + m(1,i) + \ldots + m(i-1,i),$$

and so

$$A_i + B_i \geqslant p(i) - i + m(1,i) + \ldots + m(i-1,i).$$

Taking the last results for $i = 1, \ldots, n-1$ and remembering that

$$p(n) - n + m(1,n) + \ldots + m(n-1,n) = 0$$

and adding, we obtain

$$s = \sum_1^{n-1} (A_i + B_i) \geqslant \sum_1^n p(n) - \tfrac{1}{2}n(n+1) + \sum_{i<j} m(i,j)$$

$$\geqslant \tfrac{1}{4}n(n-1) - \tfrac{1}{2}S,$$

so that

$$s = \tfrac{1}{4}n(n-1) - \tfrac{1}{2}S.$$

* I understand from Mr M. G. Kendall that this result has been known to him for some time, and the present note has therefore no claim to priority.

## REFERENCES

(1) Haden, H. G. *Proc. Cambridge Phil. Soc.* 43 (1947), 1.
(2) Kendall, M. G. *Biometrika*, 30 (1938), 81.
(3) Kendall, M. G. *The Advanced Theory of Statistics*, 1 (London, 1943).
(4) Feller, W. *Bull. American Math. Soc.* 51 (1945), 800.
(5) Dantzig, G. B. *Ann. Math. Statist.* 10 (1939), 247.

Institute of Statistics
Oxford University

https://doi.org/10.1017/S0305004100024087 Published online by Cambridge University Press