THE TREATMENT OF TIES IN RANKING PROBLEMS

By M. G. KENDALL

1. When a number of objects are presented for ranking by an observer there sometimes arise cases in which he is unable to express a preference in regard to certain of them and 'ranks them equal' or regards them as 'tying'. The effect may arise either because the objects really are indistinguishable, so far as the quality under consideration is concerned, or because the observer is unable to discern such differences as exist. Ties of this character are by no means uncommon—and indeed may be more the rule than the exception in some classes of work—and it is desirable to have a systematic method of dealing with them. In this paper I consider the effect of ties on coefficients of rank correlation, the estimation of rankings and the measurement of concordance in judges.

Rank correlations

2. The method of allocating ranking numbers to tied individuals in general use is to average the ranks which they cover. For instance, if the observer ties the third and fourth members each is allotted the number $3\frac{1}{2}$, and if he ties the second to the seventh inclusive, each is allotted the number $\frac{1}{6}(2+3+4+5+6+7) = 4\frac{1}{2}$. This is known as the mid-rank method and is the only one I shall consider. In fact I have seen only two other courses mentioned:

(a) 'Student' (1921) refers to a suggestion by Karl Pearson, as an alternative to mid-ranks, that the ties should all be ranked as if they were the highest member of the tie.

This is subject to the obvious disadvantages that it gives different results if one ranks from the other end of the scale and that it destroys the useful property that the mean rank of the whole series shall be $\frac{1}{2}(n+1)$. So far as I know it has never been used in problems involving the calculation of ranking coefficients.

(b) According to Woodbury (1940), DuBois (1939), in a paper which I have been unable to consult, has suggested allotting the ties an equal rank but proposes to determine it so that the sum of squares of the ranks shall be that of an untied ranking, namely, of the first n integers, $\frac{1}{6}n(n+1)(2n+1)$. This is rather troublesome, and it is not clear to me what advantages it possesses over the mid-rank method.

3. The effect of ties on the calculation of Spearman's ρ was considered by 'Student' (1921). ρ may be regarded as the product-moment correlation between two variates given by the two rankings. Since the variance of a ranking of n is given by $\frac{1}{12}(n^2-1)$, ρ is given by

$$\rho = \frac{12}{n^3-n}\sum_{i=1}^{n}\{X_i - \tfrac{1}{2}(n+1)\}\{Y_i - \tfrac{1}{2}(n+1)\}, \tag{1}$$

where X_i and Y_i represent the two rankings. This is easily reduced to the simpler and more familiar form

$$\rho = 1 - \frac{6\Sigma(d^2)}{n^3-n}, \tag{2}$$

where
$$d_i = X_i - Y_i. \tag{3}$$

Pursuing this analogy with the product-moment correlation 'Student' shows that, on the mid-rank method, the effect of a tie of t consecutive members is to lower the variance of the ranking by $\frac{1}{12n}(t^3-t)$. This is additive for any number of sets of ties in either the X- or the Y-variate, and if we write

$$T_X = \tfrac{1}{12}\Sigma(t^3-t), \tag{4}$$

the summation being over the ties of the X-variate, and T_Y for the similar sum for Y, we find for the product-moment correlation, say ρ_S,

$$\rho_S = \frac{1}{2}\frac{\operatorname{var} X + \operatorname{var} Y - \operatorname{var}(X-Y)}{\sqrt{(\operatorname{var} X \operatorname{var} Y)}}$$

$$= \frac{\frac{1}{6}(n^3-n)-(T_X+T_Y)-\Sigma(d^2)}{\sqrt{\{\frac{1}{6}(n^3-n)-2T_X\}\{\frac{1}{6}(n^3-n)-2T_Y\}}} \tag{5}$$

$$= \frac{\frac{1}{6}(n^3-n)-(T_X+T_Y)-\Sigma(d^2)}{\{\frac{1}{6}(n^3-n)-(T_X+T_Y)\}\sqrt{\left[1-\frac{(T_X-T_Y)^2}{\{\frac{1}{6}(n^3-n)-(T_X+T_Y)\}^2}\right]}} \tag{6}$$

'Student' notes that if T_X-T_Y is small, we have approximately

$$\rho_S = 1-\frac{\Sigma(d^2)}{\frac{1}{6}(n^3-n)-(T_X+T_Y)}. \tag{7}$$

It is also useful to note that if T_X and T_Y are small compared with $\frac{1}{6}(n^3-n)$, we have

$$\rho_S = 1-\frac{6\Sigma(d^2)}{n^3-n},$$

so that the correction to be applied to the ordinary formula is negligible for many practical purposes.

Example 1. Consider, for instance, the two rankings of 10:

X:	1	$2\frac{1}{2}$	$2\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	$6\frac{1}{2}$	$6\frac{1}{2}$	8	$9\frac{1}{2}$	$9\frac{1}{2}$
Y:	1	2	$4\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	8	8	8	10

In the first ranking there are four tied pairs and hence

$$T_x = \tfrac{4}{12}(2^3-2) = 2.$$

In the second there is one set of four ties and one of three, and hence

$$T_y = \tfrac{1}{12}(60+24) = 7.$$

We also have $\Sigma(d^2) = 13.$

Hence, in accordance with (5), $\rho_s = \frac{165-22}{\sqrt{(161.151)}} = \frac{143}{155\cdot 92}$

$= 0\cdot 9171.$

Calculation on the basis of (2) gives $\rho = 1-\frac{78}{990} = 0\cdot 9212.$

The value given by (7) is $\rho_s = 1-\frac{13}{165-9}$

$= 0\cdot 9167.$

4. There is another way of looking at this problem which 'Student' did not mention. Suppose we regard any set t of tied ranks as due to inability on the part of the observer to distinguish real differences; i.e. we assume that there does exist a set of integral ranks although we are ignorant of it on present evidence. Then we may ask, what is the *average* value of ρ over all the $t!$ possible ways of assigning integral ranks to the tied members?

5. If the t corresponding members in the Y-ranking are held fixed, then the average covariance for all $t!$ arrangements of the X-members is the covariance of the fixed Y-members and the average of the t X-members. But this latter gives the mean ranks of the tied members, and consequently the mean covariance of the two rankings is

$$\frac{1}{n}\{\tfrac{1}{12}(n^3-n)-\tfrac{1}{2}\Sigma(d^2)-\tfrac{1}{2}(T_X+T_Y)\}, \tag{8}$$

the effect of various sets of ties being additive. If now we divide by the *actual* variances of X and Y we arrive at equation (5). Thus 'Student's' formula may be regarded as giving a mean value of the coefficient which would be obtained if the ties were replaced in all possible ways by the integral ranks which they cover; always bearing in mind that we have not averaged the variances.

6. A similar point of view has been adopted by Woodbury (1940) who does not seem to have been aware of 'Student's' results; but Woodbury takes as his variance the quantity $\frac{1}{12}(n^2-1)$, that is to say, he determines the average ρ which would be obtained if the ties were replaced by appropriate integral rankings in all possible ways, the variance in each case being that of the first n integers. This results in ρ_W, say, where

$$\rho_W = 1-\frac{6\{\Sigma(d^2)+T_X+T_Y\}}{n^3-n}, \tag{9}$$

the difference from ρ_S of equation (5) lying, of course, in the denominators in the second term on the right.

Example 2. For example, in the illustration considered above Woodbury's value would be

$$\begin{aligned}\rho_W &= 1-\frac{6(13+2+7)}{990}\\ &= 0{\cdot}8667,\end{aligned}$$

against $\rho_S = 0{\cdot}9171$.

7. The question then arises, which is the better measure of rank correlation, ρ_S or ρ_W? It is useful in the first instance to consider some special cases.

(*a*) Suppose that the two rankings are both completely tied, i.e. that each rank is $\frac{1}{2}(n+1)$. We clearly have

$$\rho = 1,$$

indicating complete correlation. For the 'Student' form we have

$$\rho_S = \frac{0}{\sqrt{(0\times 0)}},$$

an indeterminate form which, however, may be regarded as unity as a limiting case in virtue of the next subsection. For Woodbury's form we find

$$\begin{aligned}\rho_W &= \frac{0}{\frac{1}{6}(n^3-n)}\\ &= 0,\end{aligned}$$

indicating zero correlation. In short, Woodbury's 'correction' has reduced ρ from 1 to 0.

(*b*) Suppose that both rankings are the same, that the last member in each is ranked n and that the others are all tied and hence have rank $\frac{1}{2}n$. Then it will be found that

$$\rho = 1,$$

$$\rho_S = 1,$$

$$\rho_W = \frac{3}{n+1} \quad (n \geqslant 2).$$

The crude form of ρ and 'Student's' corrected form are in agreement that the correlation is unity. Woodbury's form differs entirely and gives a correlation which is small for large n.

(*c*) Generally, if the two rankings are identical and there are ties giving a T-number of T in each

$$\rho = \rho_S = 1,$$

$$\rho_W = 1 - \frac{12T}{n^3 - n}.$$

(*d*) Suppose that one ranking is in the natural order $1, \ldots, n$ and has no ties, so that $T_X = 0$. If the other ranking has the last member ranked n and the others completely tied we find

$$\Sigma(d^2) = \tfrac{1}{12}n(n-1)(n-2),$$

$$\rho = \frac{n+4}{2(n+1)} \sim \frac{1}{2},$$

$$\rho_S = \sqrt{\frac{3}{n+1}},$$

$$\rho_W = \frac{3}{n+1}.$$

For large n, ρ tends to 0·5, whereas ρ_S and ρ_W tend to zero, the latter faster than the former.

8. It appears to me that the decision as to which of the coefficients ρ_S or ρ_W is preferable can only rest on the use to which they are to be put.

(*a*) Let us suppose in the first instance that the objects have a definite ranking $1, \ldots, n$ determined in some objective way. The purpose of correlating the order assigned by an observer is then to determine the observer's accuracy, not the real ranking. 'Student's' form of the coefficient would measure the product-moment correlation of ranks, giving weight to the fact that if the observer produces ties the variance of his estimates is reduced. Woodbury's form would measure the average correlation of all the results obtained if the observer allotted to the tied groups integral ranks determined at random. For instance in a ranking of 8

1	2	3	4	5	6	7	8
4	4	4	4	4	4	4	8

$\rho_S = \sqrt{\frac{1}{3}} = 0{\cdot}577$, $\rho_W = \frac{1}{3} = 0{\cdot}333$. There does not seem to me to be much to choose between the two, but on the whole Woodbury's form gives figures nearer to what one would expect. We may suppose the observer to be in a genuine state of indecision when considering the tied members, and the average of all the values given by guessing integral ranks at random seems a fair measure of his ability. 'Student's' form gives higher values because he divides the product-moment by the actual standard deviations, and hence gives the observer credit, so to speak, for clustering his values in spite of the fact that he ought not to do so *because there really is an objective order*. In a case of this kind I should therefore favour Woodbury's form.

(b) The situation is quite different if no objective order is given and we are measuring the concordance between two judges. In this case Woodbury's form seems to me to give the wrong answer. In the case where two rankings are identical, for instance, one is entitled to expect that a measure of correlation should be unity—agreement could not be better. Both judges may be wrong, but that is not the point. We are measuring their agreement, not their accuracy. It has been shown above that if all members are tied Woodbury's form would give a zero correlation between the two rankings, which on the face of it seems ridiculous. We are no longer entitled to assume an objective order, or, even if there are real differences in the objects, to suppose that they fall above the threshold of the discriminatory power of the judges. 'Student's' form appears to be far better.

9. It is, of course, undeniable that 'Student's' form is more troublesome to calculate. This is unimportant if only two or three rankings are to be compared, but might be more important if there were large numbers of rankings. In such a case, however, it is more usual to work out a single measure of joint correlation rather than many pairs. The problem of m rankings for tied variates is dealt with below.

10. I proceed to consider the appropriate method of dealing with ties in calculating the alternative coefficient of correlation known as τ (Kendall, 1938). In an elegant synthesis of the rank correlation problem Daniels (1944) has recently pointed out that τ may be defined as

$$\tau = \frac{\Sigma(a_{ij}b_{ij})}{\sqrt{\{\Sigma a_{ij}^2 \Sigma b_{ij}^2\}}}, \tag{10}$$

where $$a_{ij} = -a_{ji}, \quad b_{ij} = -b_{ji},$$

a_{ij} is a score allotted to each pair of ranks X_i, X_j as $+1$ if $j > i$ and -1 if $i < j$, b_{ij} similarly relating to the Y-ranking. In the ordinary ranking case, of course,

$$\Sigma a_{ij}^2 = \Sigma b_{ij}^2 = \tfrac{1}{2}n(n-1). \tag{11}$$

Daniels' form has the advantage of revealing τ as analogous to an ordinary product-moment coefficient.

11. To extend this definition to the case of tied ranks we have only to define the score a_{ij} for equal ranks, and this is easily done by defining it as zero, midway between the scores of $+1$ and -1 taken when the ranks are unequal. This, it will be noticed, affects the denominator in the definition of equation (10) as well as the numerator.

Example 3. Let us consider the rankings of Example 1, namely,

X:	1	$2\frac{1}{2}$	$2\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	$6\frac{1}{2}$	$6\frac{1}{2}$	8	$9\frac{1}{2}$	$9\frac{1}{2}$
Y:	1	2	$4\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	8	8	8	10

Considering the first member in association with the other 9, we see that the score in both rankings in each case is $+1$, so that the total score is 9; the second and third members in the X-ranking are tied and this pair therefore scores 0, whatever the Y-position. The score from pairs associated with the second member will be found to be 7; in the Y-ranking members 3–6 are tied and therefore the only non-vanishing scores arise from association of the third member with the seventh, eighth, ninth and tenth members, score 4. The total score will be found to be

$$9+7+4+4+4+3+1+1+0 = 33.$$

The sum Σa_{ij}^2 is found to be 41 and Σb_{ij}^2 is 36 and hence, writing τ_s for the corresponding quantity to ρ_s, we have

$$\tau_s = \frac{33}{\sqrt{(41.36)}} = 0{\cdot}8589.$$

The value of ρ_s is 0·9171 but the difference need cause no concern as the coefficients do give rather different results, having different scales.

The general rule for the formation of Σa_{ij}^2 will be clear. If there is a tie of extent t we calculate $\frac{1}{2}t(t-1)$ and sum for all ties. If this sum is U then

$$\Sigma a_{ij}^2 = \tfrac{1}{2}n(n-1) - U. \tag{12}$$

12. If we replace any tied set by the corresponding integral ranks in any order and average for all the $t!$ possible orders we get the same result as by replacing a_{ij} for the tied members by zero; for in the $t!$ arrangement each pair will occur an equal number of times in the order XY and the order YX, so that the allocation of $+1$ in the first case and -1 in the second is equivalent to the allocation of zero on the average. Thus we may regard our score for tied ranks as the mean of the values obtained by allotting integral ranks in all possible ways. On the analogy of Woodbury's treatment of ρ we could then define

$$\tau_W = \frac{\text{score}}{\frac{1}{2}n(n-1)}. \tag{13}$$

The choice between τ_S and τ_W is precisely the same as in the case of Spearman's ρ; that is to say we might use the latter where an objective order is known but the former where it is a question of measuring the concordance between judges.

13. For the purposes of comparison with the special cases considered in § 7 it may be worth while giving the corresponding values of T:

(*a*) Both rankings completely tied:

$$\tau_S \text{ indeterminate}, \quad \tau_W = 0.$$

(*b*) Rankings identical and all tied except last member:

$$\tau_S = 1, \quad \tau_W = \frac{2}{n}.$$

(*c*) Rankings identical, ties giving U-member U in each:

$$\tau_S = 1, \quad \tau_W = 1 - \frac{2U}{n(n-1)}.$$

(*d*) One ranking equal to the natural order $1, \ldots, n$, the other all tied except the last member:

$$\tau_S = \sqrt{\frac{2}{n}}, \quad \tau_W = \frac{2}{n}.$$

The estimation of a ranking

14. In a previous note (1942) I considered the problem of estimating the true ranking (or the ranking on which there was the greatest measure of agreement) for m-rankings of n individual exemplified by

Object	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
	4	2	1	7	6	3	5	8
	7	2	1	6	4	5	3	8
	7	4	2	6	5	3	1	8
Sum of ranks	18	8	4	19	15	11	9	24

It was shown that a reasonable estimate was to be obtained by ranking according to the sums of ranks, beginning with the lowest, e.g. in this example the ranking would be

$$A_3 \quad A_2 \quad A_7 \quad A_6 \quad A_5 \quad A_1 \quad A_4 \quad A_8.$$

This is the best in that it minimizes the sum of squares of deviations from what they would be if the m-rankings were identical; and it also maximizes the average ρ between the observed and the estimated rankings.

15. The above method may also be regarded in this way: if an object is ranked r, it is preferred to $n-r$ members but $r-1$ members are preferred to it. Allotting as usual $+1$ for the first type and -1 to the second we see that the individual scores $n+1-2r$ in its own ranking. Summing over the m-rankings we see that an individual ranked X_r, Y_r, Z_r, etc. has altogether a score of

$$m(n+1)-2\Sigma(X). \tag{14}$$

If then we rank the individuals according to their total scores, beginning with the highest, we arrive at exactly the same result as by ranking according to $\Sigma(X)$ beginning with the lowest. Thus our method arranges the objects in the order of numbers of preferences; a further argument in its favour. It is also easy to see that the method minimizes the sums of squares of deviations of preferences from what they would be if there were complete agreement. In fact, denote the estimated ranking by $X_1, \ldots, X_n$ and let the corresponding sums of preferences be $\xi_1, \xi_2, \ldots, \xi_n$, this being a permutation of $m(n+1)-mX_j$, $j = 1, \ldots, n$. If the actual preferences are given in sum by $S_1, \ldots, S_n$ we have to minimize

$$\sum_{j=1}^{n}(S_j-\xi_j)^2 = \Sigma S^2+\Sigma\xi^2-2\Sigma(S\xi). \tag{15}$$

The first two terms on the right are constants and we have therefore to maximize $\Sigma(S\xi)$. This is clearly done by multiplying the largest S by the largest ξ, that is to say the largest S by the smallest X, and so on. In other words, we allot X_1 to the largest S and so on in order.

16. To complete the story one would like to be able to prove that the method maximized the average τ between observed and estimated rankings.

Unfortunately the proposition fails, as is shown by the following example:

A_1	A_2	A_3	A_4
2	3	1	4
1	2	4	3
1	2	4	3
4	7	9	10

The estimated order here would be that running from left to right across the page as written, and the total score between that order and the three observed orders will be found to be $2+4+4 = 10$. But if we interchange the last two columns it becomes $0+6+6 = 12$. Such a situation, however, is of rare occurrence and can only occur when there is substantial disagreement between judges on the two objects interchanged, in which case no ranking is very reliable. I do not think, therefore, that the failure of the result in extreme cases is important.

17. Suppose now that some of the ranks are tied. Does the method of summing ranks apply unchanged to give a good estimate?

(*a*) In the first place, the method continues to give an answer which appears reasonable on the face of it. Moreover it may be regarded as an average result for all the ways of permuting the tied ranks when replaced by the appropriate integral ranks.

(*b*) If the question is regarded as one of ranking according to preferences, the replacement of a pair of integral ranks by a tie does not affect the preferences with other members and merely cancels a preference between the tied pair; and so for any set of ties. In consequence the method preserves the property of ranking according to the number of preferences.

(*c*) If we measure the average ρ with the estimated ranking in Woodbury's form of corrected ρ the method provides a maximum average ρ unless the estimated ranking itself contains ties, in which case the result might conceivably fail, though it is unlikely to do so.

In fact, let the estimated ranking be $X_1, \ldots, X_n$ and the rank of the jth object in the kth ranking be Y_{jk}. We shall maximize the average ρ by maximizing

$$\begin{aligned} V &= \sum_{k=1}^{m} \sum_{j=1}^{n} \{X_j - \tfrac{1}{2}(n+1)\}\{Y_{jk} - \tfrac{1}{2}(n+1)\} \\ &= \sum_{j=1}^{n} \{X_i - \tfrac{1}{2}(n+1)\}\{S_j - \tfrac{1}{2}m(n+1)\}, \end{aligned} \tag{16}$$

which reduces to maximizing $\Sigma(XS)$. If, however, there are ties in the estimated ranking our problem is to minimize something of the form

$$\frac{V}{\left\{\tfrac{1}{2}(n^2-1) - \dfrac{1}{n}\Sigma(T_X)\right\}^{\frac{1}{2}}}, \tag{17}$$

and variations in T_X may upset the result. This, however, is not likely to be serious unless there are many ties in the estimated ranking, in which case estimation of any kind is unreliable.

(*d*) If we measure the average ρ with the estimated ranking in 'Student's' form the result again may fail for (16) then becomes of the form

$$\Sigma\Sigma\{X_j - \tfrac{1}{2}(n+1)\} A_{jk}\{Y_{jk} - \tfrac{1}{2}(n+1)\}, \tag{18}$$

where the coefficients A_{jk} differ from our ranking to another because they depend on differing variances.

(*e*) Similar considerations apply to the proposition that the method minimizes the sums of squares of deviations from what the sums of ranks or preferences would be if all rankings were alike. Apart from complications due to ties in the estimated ranking, the minimal properties continue to obtain.

18. To sum up, therefore, the method of estimating the ranking according to sums of ranks appears to give satisfactory results when ties are involved.

Example 4. Consider the three rankings

	1	2	3	4½	4½	6	7½	7½	9	10
	1	2½	2½	4½	4½	6½	6½	8	9½	9½
	1	2	4½	4½	4½	4½	8	8	8	10
Sums of ranks	3	6½	10	13½	13½	17	22	23½	26½	29½
Estimated ranking	1	2	3	4½	4½	6	7	8	9	10

The sums of ranks give an estimated ranking as shown. There is one case here where the sums of ranks are equal and the individual ranks yielding those sums are also equal. There seems no better course than to tie them. Had the sums for these two been

4½	3½
4½	4½
4½	5½
13½	13½

we might perhaps have ranked the former as 4 and the latter as 5, on the ground that the variance of the former is less and the group therefore 'cluster' better than the other. This is a new principle deriving no support from the various minimal principles already introduced. It is the usual practice, I think, to regard an estimate as better when it is based on more closely grouped observations; but here the resulting estimate of the mean ranking is $4\frac{1}{2}$ so it can also be argued that the tie should remain. On the whole this seems to me the better course.*

TESTS OF SIGNIFICANCE FOR m RANKINGS

19. I proceed to consider what modifications, if any, are required in significance tests when ties can appear in the rankings. Babington Smith & I (1939) have discussed the case when the ranks are integral. The algebra required for the more extended discussion has been given by Pitman (1938) in considering a similar problem in the analysis of variance.

Consider an array of m rows

$$\begin{matrix} a_1 & a_2 & \dots & a_n \\ b_1 & b_2 & \dots & b_n \\ \dots & \dots & \dots & \dots \\ k_1 & k_2 & \dots & k_n \end{matrix} \tag{19}$$

If S is the sum of square of column totals about their mean and S' the sum of squares of all values about their mean, we define

$$W = \frac{S}{mS'} \tag{20}$$

as the coefficient of concordance. In the case when the a's, b's, etc. are permutations of the natural numbers 1 to n we have

$$W = \frac{S}{\frac{1}{12}m^2(n^3-n)}. \tag{21}$$

W can vary from 0 to 1, attaining the latter value only if all rankings are identical.

20. Let us in the first place consider what happens to formula (21) when some of the integral ranks are replaced by ties. If the T-numbers for the various rows are T_a, T_b, etc., the formula becomes

$$W = \frac{S}{\frac{1}{12}m^2(n^3-n)-m\Sigma(T)}. \tag{22}$$

This is as simple a form as we require.

In the data of Example 4, for instance, we find

$$S = 682{\cdot}9,$$

$$W = \frac{682{\cdot}9}{742{\cdot}5-30}$$

$$= 0{\cdot}958.$$

* In my note (1942), which dealt only with integral ranks, I suggested that, where the sums of ranks are equal, precedence should be given to the one with the smaller variance; but I was there considering only an estimated ranking which itself was integral. When ties are permitted I should, as stated above, use them where sums of ranks are equal.

21. Denoting by α_r the rth moment of the a-row in (19), and similarly for β_r, γ_r, etc., and by α'_r the rth k-statistic, we have for the moments of W, from Pitman's results,

$$\overline{W} = E(W) = \frac{1}{m}, \tag{23}$$

$$E(W-\overline{W})^2 = \frac{4}{m^2(n-1)}\frac{\Sigma\alpha_2\beta^2}{\Sigma^2\alpha_2}, \tag{24}$$

$$E(W-\overline{W})^3 = \frac{48}{m^3(n-1)}\frac{\Sigma\alpha_2\beta_2\gamma_2}{\Sigma^3\alpha_2}+\frac{8(n-1)(n-2)}{m^3n^4}\frac{\Sigma\alpha'_3\beta'_3}{\Sigma^3\alpha_2}, \tag{25}$$

$$E(W-\overline{W})^4 = \frac{48}{m^4(n-1)^2}\frac{\Sigma^2\alpha_2\beta_2}{\Sigma^4\alpha_2}-\frac{96}{m^4(n-1)^2(n+1)}\frac{\Sigma\alpha_2^2\beta_2^2}{\Sigma^4\alpha_2}+\frac{1152}{m^4(n-1)^3}\frac{\Sigma\alpha_2\beta_2\gamma_2\delta_2}{\Sigma^4\alpha_2}$$
$$+\frac{16(n-1)(n-2)(n-3)}{m^4n^5(n+1)}\frac{\Sigma\alpha'_4\beta'_4}{\Sigma^4\alpha_2}+\frac{252(n-2)}{m^4n^4}\frac{\Sigma\alpha'_3\beta'_3\gamma_2}{\Sigma^4\alpha_2}. \tag{26}$$

In the case when the numbers are permutations of the first n integers these expressions reduce to

$$\overline{W} = \frac{1}{m}, \tag{27}$$

$$E(W-\overline{W})^2 = \frac{2(m-1)}{m^3(n-1)}, \tag{28}$$

$$E(W-\overline{W})^3 = \frac{8(m-1)(m-2)}{m^5(n-1)^2}, \tag{29}$$

$$E(W-\overline{W})^4 = \frac{12(m-1)^2}{m^6(n-1)^2}+\frac{48(m-1)(m-2)(m-3)}{m^7(n-1)^3}-\frac{48(m-1)}{m^7(n-1)^2(n+1)}. \tag{30}$$

If m and n are moderately large, these expressions are approximately the same (exactly so for the first two) as the moments of

$$dF = \frac{1}{B(p,q)}W^{p-1}(1-W)^{q-1}dW, \tag{31}$$

where

$$\left.\begin{aligned} p &= \tfrac{1}{2}(n-1)-\frac{1}{m},\\ q &= (m-1)p.\end{aligned}\right\} \tag{32}$$

It follows that W can be tested in Fisher's z-distribution by writing

$$z = \tfrac{1}{2}\log\frac{(m-1)W}{1-W}, \tag{33}$$

$$\left.\begin{aligned} \nu_1 &= (n-1)-\frac{2}{m},\\ \nu_2 &= (m-1)\nu_1.\end{aligned}\right\} \tag{34}$$

How far does this require modification for tied ranks?

22. For the purposes of an accurate test we can, of course, work out the first four moments of W from (23) to (26) in individual cases and fit an *ad hoc* curve; but this is a tedious process and some approximation is necessary.

The first two moments of (31) are

$$\mu_1 \text{ (about zero)} = \frac{p}{p+q},$$
$$\mu_2 = \frac{pq}{(p+q)^2(p+q+1)},$$

and if we identify them with the first two moments of W, $\frac{1}{m}$ and $\mu_2(W)$, say, we find

$$\left.\begin{aligned} p &= -\frac{1}{m}+\frac{m-1}{m^3\mu_2(W)}, \\ q &= (m-1)p, \end{aligned}\right\} \tag{35}$$

so that approximately W can be tested in the z-distribution with

$$\left.\begin{aligned} \nu_1 &= -\frac{2}{m}+\frac{2(m-1)}{m^3\mu_2(W)}, \\ \nu_2 &= (m-1)\nu_1. \end{aligned}\right\} \tag{36}$$

23. We have, as in (24),

$$\mu_2(W) = \frac{4}{m^2(n-1)}\frac{\Sigma\alpha_2\beta_2}{\Sigma^2\alpha_2}.$$

Writing A for $\Sigma\alpha_2$ and B for $\Sigma\alpha_2^2$ we have

$$\mu_2(W) = \frac{2}{m^2(n-1)}\left\{1-\frac{B}{A^2}\right\},$$

so that the appropriate degrees of freedom are

$$\left.\begin{aligned} \nu_1 &= -\frac{2}{m}+\frac{(n-1)(m-1)}{m(1-B/A^2)}, \\ \nu_2 &= (m-1)\nu_1. \end{aligned}\right\} \tag{37}$$

If the T-numbers appropriate to the various rankings are small compared with $\frac{1}{12}(n^3-n)$ we can approximate further. In fact write N for $\frac{1}{12}(n^3-n)$. Then

$$A = mN-\Sigma(T),$$
$$B = mN^2-2N\Sigma(T)+\Sigma(T^2),$$

and to the first order in $\Sigma(T)$

$$\begin{aligned} \frac{B}{A^2} &= \frac{mN^2-2N\Sigma(T)}{m^2N^2-2mN\Sigma(T)} \\ &= \frac{1}{m}\left\{1-\frac{2\Sigma(T)}{mN}\right\}\left\{1-\frac{2\Sigma(T)}{mN}\right\}^{-1} \\ &= \frac{1}{m}. \end{aligned}$$

On substitution in (27)

$$\left.\begin{aligned} \nu_1 &= -\frac{2}{m}+(n-1), \\ \nu_2 &= (m-1)\nu_1. \end{aligned}\right\} \tag{38}$$

This, of course, is the same as (34) so that, if the number or extent of the ties is small, the test for untied ranks requires no modification (other than that necessary in the calculation of W itself).

24. It thus appears that we can apply the usual test unchanged unless the ties are numerous enough to render $\Sigma(T)$ not small compared with mN. If the ties are numerous we can work with the modified degrees of freedom given by (35), but in such a case it would probably be as well to verify by direct calculation that the third and fourth moments of W were in reasonable agreement with those given by the β-approximation implicit in the use of the z-test. If it happens that one or two rankings contribute the major part of $\Sigma(T)$ we may perhaps reject them on the grounds that the judges are very bad, but the rejection of observations has to be done with some care and only after we are satisfied that they really are exceptional and not merely outlying members of a continuous range.

Addition of extra members to a ranking

25. There is one class of case in which I have found the coefficient τ to have definite advantages over ρ. An example will illustrate the point best. Suppose I send out an inquiry to a number of firms asking for some information which they may or may not wish to disclose; and suppose that the information is of a type for which one would expect that the non-participants might differ from the participants. By a certain date a number of replies have been received and it is then necessary to close the inquiry and to summarize the results. How far can I assume that the replies to hand are representative of the population to which the inquiry was addressed? Is there any evidence to suggest that those who reply earlier to the inquiry differ from those who reply later?

26. To simplify the illustration suppose that I receive 15 replies in the form of a percentage figure which occur in the following order:

Order of receipt:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Percentage:	15	13	12	16	25	8	9	14	17	11	18	20	10	21	19
Rank of percentage:	8	6	5	9	15	1	2	7	10	4	11	13	3	14	12

I have chosen percentages which are all different so as not to complicate the example, but if ties are present there is no essential modification.

Now if there is some relation between the order of reply and the magnitude of the percentage, it ought to be shown up by the rank correlation between the order of reply and the order of magnitude of the percentage. The latter is shown in the last row of the example above and we find

$$\Sigma(d^2) = 392,$$

$$\rho = 0{\cdot}300.$$

This in fact, is barely significant, but I am not for the moment concerned with significance. Suppose that after we have completed this calculation two more replies arrive with percentages 7 and 23. We now have to calculate a revised value of ρ by re-numbering nearly all the replies and working *ab initio*. In practice the continual arrival of stragglers is quite common and to work out a new value of ρ each time is a great arithmetical nuisance. The point I wish to make is that τ is not subject to this disability, extra values being capable of addition as required.

In the above example for 15 members the value of $\Sigma(a_{ij}b_{ij})$ is easily seen to be

$$0+3+4+1-10+9+8+3+2+3+2+1+2-1 = 27,$$

so that

$$\tau = \tfrac{27}{105} = 0{\cdot}257.$$

If now we add a 16th member with the value 7, the contribution to $\Sigma(ab)$ is obtained by considering this new member in conjunction with the other fifteen, and is seen to be -15. Similarly, a further member valued 23 adds 13. The new value of $\Sigma(ab)$ is thus 25 and the new τ is given by

$$\tau = \tfrac{25}{136} = 0{\cdot}184.$$

In this way a kind of running value of τ can be ascertained without re-ranking at each stage as is necessary witn ρ. Thus τ has a decided advantage in this class of case, namely the calculation of ranking coefficients for time series which may be extended in length.

REFERENCES

DANIELS, H. E. (1944). The relation between measures of correlation in the universe of sample permutations. *Biometrika*, **33**, 129.

DUBOIS, P. (1939). Formulas and tables for rank correlation. *Psychol. Rec.* **3**, 45.

KENDALL, M. G. (1938). A new measure of rank correlation. *Biometrika*, **30**, 81.

KENDALL, M. G. (1942). Note on the estimation of a ranking. *J.R. Statist. Soc.* **105**, 119.

KENDALL, M. G. & BABINGTON SMITH, B. (1939). The problem of *m* rankings. *Ann. Math. Statist.* **10**, 275.

PITMAN, E. J. G. (1938). Significance tests which may be applied to samples from any population. III. The analysis of variance test. *Biometrika*, **29**, 322.

'STUDENT' (1921). An experimental determination of the probable error of Dr Spearman's correlation coefficient. *Biometrika*, **13**, 263.

WOODBURY, M. A. (1940). Rank correlation when there are equal variates. *Ann. Math. Statist.* **11**, 358.

16