OXFORD JOURNALS
OXFORD UNIVERSITY PRESS

Non-Null Ranking Models. I

Author(s): C. L. Mallows

Source: *Biometrika*, Jun., 1957, Vol. 44, No. 1/2 (Jun., 1957), pp. 114-130
Published by: Oxford University Press on behalf of Biometrika Trust
Stable URL: https://www.jstor.org/stable/2333244

REFERENCES
Linked references are available on JSTOR for this article:
https://www.jstor.org/stable/2333244?seq=1&cid=pdf-reference#references_tab_contents
You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms

and *Oxford University Press* are collaborating with JSTOR to digitize, preserve and extend access to *Biometrika*

JSTOR

This content downloaded from
210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms

NON-NULL RANKING MODELS. I

By C. L. MALLOWS
University College, London

1. Introduction and summary

Kendall (1950) has remarked that the major outstanding problem in ranking theory is the specification of a suitable population of ranks in non-null cases. Much attention has been concentrated on situations which Daniels (1950, § 5) calls of type (i):

'The sample is regarded as having been randomly chosen from a bivariate population of ranks'

the underlying population being either finite or infinite, e.g. bivariate Normal. Rather less work has been done on Daniels's type (ii):

'There is a fixed set of individuals being assessed by a population of judges, or by the same judge in repeated trials, on a particular attribute whose ranking is known *a priori*. The random element is uncertainty of preference, the correlation being the result of real differences between the individuals, and the population is one of rankings conditional on a given objective order.'

Daniels (1950), following Babington Smith, Thurstone and Mann, treats this problem as one of regression. The present approach is by way of paired-comparison theory. The judge is assumed to arrive at a ranking of n objects $U_1, U_2, \ldots, U_n$ by first making all the ${}^nC_2$ comparisons between pairs *independently*, but then only accepting the results if they are consistent with a ranking of the n objects.

Various non-null models are proposed. The general model (eqn. (1)) depends on ${}^nC_2$ parameters—this number is then reduced to $n-1$ by using the Bradley–Terry (1952) paired-comparison model (eqn. (4)). An alternative method of simplification is proposed which makes the probability of putting (in paired-comparisons) any two objects U_i and U_j in the correct order equal to

$$\tfrac{1}{2}+\tfrac{1}{2}\tanh(k\log\theta+\log\phi),$$

where θ and ϕ are parameters, and $k = j-i$ is the difference between the true ranks of the two objects. Thus this probability is a simple monotonic function of a quantity which is composed of a term increasing linearly with k, and a constant term. The null hypothesis corresponds to $\theta = \phi = 1$. It is found that θ is associated with Spearman's coefficient r_s, and ϕ with Kendall's t_k (eqn. (9)). Each of these parameters has a further interpretation. Thus θ may be regarded as assigning weights to the n objects, these weights being in geometric progression. The paired-comparisons are then made in such a way that the probability of ranking one object higher than another is a simple function of the weights assigned to these two objects. The parameter ϕ has the following further interpretation: having obtained a ranking of length $n-1$ and wishing to introduce a further object, ϕ specifies the probabilities of this object being ranked in the various possible positions; these probabilities being in geometric progression, decreasing away from the true position (eqn. (18)).

Putting $\phi = 1$ in the general model gives a special case of the Bradley–Terry model (eqn. (11)). It is shown, however, that asymptotically, when the joint distribution of r_s

This content downloaded from
210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms

and t_k tends to the bivariate Normal form, the two coefficients cannot distinguish between the two parameters; it is therefore proposed to put $\theta = 1$ and to use only the one parameter ϕ (eqn. (12)). This leads to exceptionally simple results, including an explicit form for the probability generating function (p.g.f.) of t_k and an invariance property of the probabilities p_{ij} of obtaining two objects U_i and U_j in the correct order in the ranking; p_{ij} is found to depend only on $j-i$ and ϕ, and not on i or n (§ 9).

Methods of estimating ϕ are given; tests are derived which may be used to decide whether two judges' rankings of the same objects are consistent with the same non-null hypothesis; more generally, given m judges, each of whom produces l rankings, we may test both for differences between judges and for inconsistencies within judges. The power of these tests follows immediately from standard theory.

An expression is derived giving the conditional expectation of r_s for given t_k.

2. THE GENERAL MODEL, $\mathscr{H}_\lambda$

In the method of paired-comparisons, n objects $U_1, U_2, \ldots, U_n$, whose *a priori* ranking is here assumed known, are ranked in pairs, there being ${}^nC_2$ such comparisons in all. Babington Smith (1950) suggested that a suitable non-null hypothesis for this situation would be to assign to each pair (i, j) $(i<j)$ the probability π_{ij} of ranking U_i lower than U_j; this we shall denote by

$$\pi_{ij} = P\{U_i \prec U_j\} \quad (1 \leqslant i < j \leqslant n),$$

and to assume that all the comparisons are independent.

The proposed general non-null ranking model, depending on ${}^nC_2$ parameters π_{ij}, is as follows:

(i) The above model is used to generate a set of ${}^nC_2$ comparisons.

(ii) If the resulting complex of comparisons is consistent, i.e. if there are no circular triads of comparisons such as $U_i \prec U_j \prec U_k \prec U_i$, then the complex is equivalent to a ranking; to each of the objects U_i may be assigned an integer u_i giving the position of that object in the ranking. Thus in the case $n = 4$, the comparisons

$$U_1 \succ U_2, \quad U_1 \succ U_3, \quad U_1 \succ U_4, \quad U_2 \prec U_3, \quad U_2 \prec U_4, \quad U_3 \succ U_4$$

give the ranking

$$U_2 \prec U_4 \prec U_3 \prec U_1,$$

which may be expressed by

$$u_1 = 4, \quad u_2 = 1, \quad u_3 = 3, \quad u_4 = 2.$$

We then have

$$u_2 < u_4 < u_3 < u_1.$$

(iii) If the complex of comparisons is inconsistent, no ranking is possible. The procedure (i) is then repeated as often as necessary until a consistent complex is obtained; and the corresponding ranking is accepted.

The above is a possible way of generating rankings. It should be emphasized that it is not suggested that any actual experiment would be performed in this fashion; if the paired-comparisons *can* be made independently (and recorded), there is little point in trying to force the preferences into a ranking. What is here attempted is the production of a model for situations where the observed data is a ranking, the difficulty being that the comparisons are no longer independent (thus if $U_i \prec U_j$ and $U_j \prec U_k$, we must have $U_i \prec U_k$). We are introducing this dependence by considering a conditional distribution of paired-comparisons; we admit only a subset (containing $n!$ members) of the total $2^{{}^nC_2}$ possible outcomes.

This content downloaded from
210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms

The probability of obtaining any given ranking (u) from the above construction will be proportional to the probability under the Babington Smith model of the corresponding complex of comparisons. Thus the probability of the ranking

$$(u) \equiv (u_1, u_2, u_3, u_4) = (4, 1, 3, 2)$$

above is proportional to

$$(1-\pi_{12})(1-\pi_{13})(1-\pi_{14})\pi_{23}\pi_{24}(1-\pi_{34});$$

putting

$$\lambda_{ij}^2 = \pi_{ij}/(1-\pi_{ij}),$$

we have that the probability of the general ranking $(u) = (u_1, u_2, \ldots, u_n)$ under this hypothesis is

$$\mathscr{H}_\lambda:\quad P\{(u)\} = K_n \prod_{1 \leqslant i < j \leqslant n} \lambda_{ij}^{\operatorname{sgn}(u_j - u_i)}, \tag{1}$$

where K_n is so chosen that the probabilities sum to 1; i.e.

$$K_n^{-1} = \sum_{(u)} \prod_{1 \leqslant i<j \leqslant n} \lambda_{ij}^{\operatorname{sgn}(u_j-u_i)} = P\{\text{comparisons are consistent}\} \prod_{1 \leqslant i<j \leqslant n} [\pi_{ij}(1-\pi_{ij})]^{-\frac{1}{2}}.$$

3. Specializations

The above general model is rather cumbersome; we now consider possible ways of specializing it so as to reduce the number of parameters. We shall attempt to find models which simplify

(i) the probabilities $P\{(u)\}$ of the various rankings (u);
(ii) the distribution problems connected with the two coefficients r_s and t_k;
(iii) various other interesting and important quantities, such as

$$p_{ij} = P\{u_i < u_j \mid \text{all comparisons consistent}\}, \tag{2}$$

which are complicated in the general case.

4. The Bradley-Terry model

In the paired-comparison case, Bradley & Terry (1952) reduce the number of parameters from ${}^nC_2$ to (effectively) $n-1$ by assuming that

$$\pi_{ij} = P\{U_i \prec U_j\} = \pi_j/(\pi_i + \pi_j) \quad (1 \leqslant i < j \leqslant n)$$

for some non-negative numbers $\pi_1, \pi_2, \ldots, \pi_n$. This gives

$$\lambda_{ij}^2 = \pi_j/\pi_i, \tag{3}$$

and for the proposed ranking model, we have

$$\mathscr{H}_B:\quad P\{(u)\} \propto \prod_{1 \leqslant i<j \leqslant n} (\pi_j/\pi_i)^{\frac{1}{2}\operatorname{sgn}(u_j-u_i)} = \prod_{1 \leqslant i \leqslant n} \pi_i^{\frac{1}{2}\sum_j \operatorname{sgn}(u_i-u_j)} \propto \prod_{1 \leqslant i \leqslant n} \pi_i^{u_i}. \tag{4}$$

This is a term in the expansion of the permanent

$$\overset{+}{\underset{}{\left|\begin{matrix} \pi_1 & \pi_1^2 & \cdots & \pi_1^n \\ \pi_2 & \pi_2^2 & \cdots & \pi_2^n \\ \vdots & \vdots & & \vdots \\ \pi_n & \pi_n^2 & \cdots & \pi_n^n \end{matrix}\right|}}^{+} = \overset{+}{|}\,\pi_i^j\,\overset{+}{|}_{i,j=1,2,\ldots,n}. \tag{5}$$

We note that this model is closely associated with what we may call a 'generalized matching coefficient', this being taken to mean a coefficient

$$A(u) = \sum_{1 \leqslant i \leqslant n} \alpha_{iu_i},$$

This content downloaded from 210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms

where the $\{\alpha_{ij}\}$ can take any values whatsoever. Examples of measures which are essentially of this form are:

(i) Number of matches; put $\alpha_{ij} = \delta_i^j$ (Kronecker δ).

(ii) Spearman's r_s; put $\alpha_{ij} = (i-j)^2$.

(iii) Spearman's 'footrule'; put $\alpha_{ij} = |i-j|$.

Others which may be considered are:

(iv) Number of matches + near misses; put $\alpha_{ij} = \delta_i^{j-1} + \delta_i^j + \delta_i^{j+1}$.

(v) Weighted form of (iv); put $\alpha_{ij} = \delta_i^{j-1} + K\delta_i^j + \delta_i^{j+1}$, where, for example, $K = 2$.

In the above cases, the various measures, $C(u)$ say, are equal either to the corresponding $A(u)$ directly, or to a simple function of $A(u)$; we have

$$C(u) = F(A(u)) \quad \text{for all } (u), \tag{6}$$

where also

$$A(u) = F^{-1}(C(u)) \quad \text{for all } (u). \tag{7}$$

The score $S(u)$ for Kendall's t_k (see eqn. (8)) can be expressed in the form (6), but for $n \geqslant 4$ the inverse function as in (7) does not exist (see Appendix I).

The p.g.f. of $A(u)$ for the above model is simply

$$\mathscr{H}_B: \quad \sum_{(u)} P\{(u)\}\, z^{A(u)} = \overset{+}{|} z^{\alpha_{ij}} \pi_i^j \overset{+}{|} / \overset{+}{|} \pi_i^j \overset{+}{|}.$$

However, in the absence of simple methods of manipulating permanents, we shall not investigate this model further at present.

Barton & David (1956) have proposed an alternative hypothesis based on distorting the null matching distribution.

5. AN ALTERNATIVE METHOD

We now consider an alternative method of simplification of the model $\mathscr{H}_\lambda$. The most general non-null model possible would specify the probability of each ranking separately, needing $n!-1$ parameters. We have reduced this number to ${}^nC_2$ by assuming a special structure for the model; it can only be further reduced by even more restrictive assumptions. Consider the following additional assumption:

Assumption A. Rankings (u), giving (when compared with the standard ranking $(1, 2, \ldots, n)$) the same values for both r_s and t_k, have the same probability.

This assumption is equivalent to

Assumption A'. The probability of any ranking (u) is a function only of

$$\left.\begin{aligned} R(u) &= \sum_{1\leqslant i<j\leqslant n} (j-i)\,\mathrm{sgn}\,(u_j-u_i) \\ &= 2\sum_{1\leqslant i\leqslant n} iu_i - \tfrac{1}{2}n(n+1)^2 \\ &= \tfrac{1}{6}n(n^2-1)\,r_s \\ \text{and} \qquad S(u) &= \sum_{1\leqslant i<j\leqslant n} \mathrm{sgn}\,(u_j-u_i) \\ &= \tfrac{1}{2}n(n-1)\,t_k. \end{aligned}\right\} \tag{8}$$

Consider also the following assumption:

Assumption B. Pairs of rankings which are inverses of one another have the same probability.

This content downloaded from
210fff:ffff:ffff:ffff:ffff on Thu, 01 Jan 1976 12:34:56 UTC
All use subject to https://about.jstor.org/terms

Two rankings (u), (v) are said to be inverses (sometimes 'conjugates') if when $u_i = j$, then $v_j = i$. Thus the rankings $(u) = (3, 5, 2, 4, 1)$ and $(v) = (5, 3, 1, 4, 2)$ are inverses since

$$u_1 = 3,\ v_3 = 1; \quad u_2 = 5,\ v_5 = 2; \quad \text{etc.}$$

It is well known (see, for example, Kendall, 1955, p. 6 (*not* p. 11)) that for such a pair of rankings

$$R(u) = R(v), \quad S(u) = S(v),$$

and so Assumption B is contained in Assumption A.

From Assumption B and eqn. (1) we may obtain certain relations between the parameters λ_{ij}; thus for the pair of rankings above we obtain

$$P\{(u)\} = K_5 \lambda_{12}\lambda_{13}^{-1}\lambda_{14}\lambda_{15}^{-1}\lambda_{23}^{-1}\lambda_{24}^{-1}\lambda_{25}^{-1}\lambda_{34}\lambda_{35}^{-1}\lambda_{45}^{-1},$$
$$P\{(v)\} = K_5 \lambda_{12}^{-1}\lambda_{13}^{-1}\lambda_{14}^{-1}\lambda_{15}^{-1}\lambda_{23}^{-1}\lambda_{24}\lambda_{25}^{-1}\lambda_{34}\lambda_{35}\lambda_{45}^{-1}$$

and for these to be equal we must have

$$\lambda_{12}^2\lambda_{14}^2 = \lambda_{24}^2\lambda_{35}^2.$$

Proceeding thus for all pairs of inverses we obtain the following relations (for $n = 5$):

$$\lambda_{12} = \lambda_{23} = \lambda_{34} = \lambda_{45} \quad = \lambda_1, \quad \text{say}$$
$$\lambda_{13} = \lambda_{24} = \lambda_{35} \quad = \lambda_2, \quad \text{say}$$
$$\lambda_{14} = \lambda_{25} \quad = \lambda_3, \quad \text{say}$$

and we put

$$\lambda_{15} \quad = \lambda_4.$$

Further,

$$\lambda_1\lambda_3 = \lambda_2^2, \quad \lambda_2\lambda_4 = \lambda_3^2.$$

Similar relations (in fact subsets of the above) are obtained for $n = 3, 4$. These relations are generalized in Appendix II and show that Assumption B contains

Assumption C. For any (fixed) n, we have

$$\text{(i)} \qquad \lambda_{ij} = \lambda_{j-i} \quad (1 \leqslant i < j \leqslant n);$$
$$\text{(ii)} \quad \lambda_{k-1}\lambda_{k+1} = \lambda_k^2 \quad (2 \leqslant k \leqslant n-1).$$

Assumption C (ii) implies that the $\{\lambda_k\}$ are in geometrical progression, i.e. for some θ, ϕ we have

$$\lambda_k = \theta^k\phi,$$

i.e.

$$P\{U_i \prec U_{i+k}\} = \frac{\theta^k\phi}{\theta^k\phi + \theta^{-k}\phi^{-1}} = \tfrac{1}{2} + \tfrac{1}{2}\tanh(k\log\theta + \log\phi).$$

Hence

$$\mathscr{H}(\theta,\phi)\colon \quad P\{(u)\} = K_n \prod_{1 \leqslant i < j \leqslant n} (\theta^{j-i}\phi)^{\operatorname{sgn}(u_j - u_i)}$$
$$= K_n \theta^{R(u)}\phi^{S(u)}. \tag{9}$$

Assumption C thus contains Assumption A. We deduce that all three assumptions are equivalent. We shall take (9) to be our basic non-null model, $\mathscr{H}(\theta, \phi)$.

6. Properties of the model $\mathscr{H}(\theta, \phi)$

Certain properties of this model are immediate.

(*a*) Let the null-hypothesis probability of a pair of values (R, S) be

$$P_0(R, S) = (n!)^{-1} F(R, S),$$

This content downloaded from
210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms

so that there are just $F(R, S)$ different rankings (u) having $R(u) = R$, $S(u) = S$. Then for the present model we have

$$\mathscr{H}(\theta, \phi): \quad P_{\theta,\phi}(R, S) \propto \theta^R \phi^S P_0(R, S).$$

If the p.g.f. under the null hypothesis is

$$\mathscr{H}_0: \quad M_0(x, y) = \sum_{R,S} P_0(R, S)\, x^R y^S,$$

then the p.g.f. for the present model is

$$\mathscr{H}(\theta, \phi): \quad M_{\theta,\phi}(x, y) = M_0(x\theta, y\phi)/M_0(\theta, \phi). \tag{10}$$

(*b*) The present model will be a special case of the one considered in § 4 if (from eqn. (3))

$$\pi_{i+k} = \theta^{2k}\phi^2\pi_i \quad (1 \leqslant i < i+k \leqslant n);$$

this requires $\phi = 1$, and then we may take without loss of generality

$$\pi_i = \theta^{2i} \quad (1 \leqslant i \leqslant n).$$

Then
$$P\{U_i \prec U_j\} = \theta^k/(\theta^k + \theta^{-k})$$

and
$$\mathscr{H}(\theta): \quad P\{(u)\} = \theta^{2\Sigma i u_i}/\overset{+}{|}\theta^{2ij}\overset{+}{|} \propto \theta^{R(u)}. \tag{11}$$

As with $\mathscr{H}_B$, this model is unmanageable because of the permanents involved.

(*c*) It is known (Daniels, 1944; Hoeffding, 1948) that under the null hypothesis, the joint distribution of R and S tends to the bivariate Normal form; i.e. that for any suitable† set $\mathscr{A}$ of points (R, S) containing α such points we have

$$\sum_{(R,S)\in\mathscr{A}} P_0(R, S) \simeq \iint_{\mathscr{A}^*} f(R, S)\, dR\, dS,$$

where $\mathscr{A}^*$ is a convex region of area 4α containing $\mathscr{A}$, but no points of $\bar{\mathscr{A}}$; the factor 4 enters since $\Delta R = \Delta S = 2$. $f(R, S)$ is the bivariate Normal distribution function with

$$\begin{aligned} \mu_R &= 0, \quad \sigma_R^2 = \tfrac{1}{36}n^2(n-1)(n+1)^2, \\ \mu_S &= 0, \quad \sigma_S^2 = \tfrac{1}{18}n(n-1)(2n+5), \end{aligned} \qquad \rho = \frac{2(n+1)}{\sqrt{\{2n(2n+5)\}}} \simeq 1 - \frac{1}{4n}.$$

For the model $\mathscr{H}(\theta, \phi)$ we have approximately (for $\mathscr{A}$ 'small')

$$\sum_{(R,S)\in\mathscr{A}} P_{\theta,\phi}(R, S) \simeq \theta^R \phi^S \iint_{\mathscr{A}^*} f(R, S)\, dR\, dS,$$

whence asymptotically, under suitable approaches to the limit for θ and ϕ, R and S are bivariate Normal with

$$\begin{aligned} \mathscr{E}(R) &\simeq \sigma_R(\sigma_R \log\theta + \rho\sigma_S \log\phi), \quad \operatorname{var} R \simeq \sigma_R^2, \\ \mathscr{E}(S) &\simeq \sigma_S(\rho\sigma_R \log\theta + \sigma_S \log\phi), \quad \operatorname{var} S \simeq \sigma_S^2, \end{aligned} \qquad \rho(R, S) \simeq \rho.$$

† The usual proofs demonstrate only that the *cumulative* distribution tends to the Normal form; what is required here is that the *ordinates*, averaged over α points, tend to the Normal ordinates. I believe that the result for $\alpha = 1$ has not been proved; Haden (1947) has proved the result for the marginal distribution of S, but in view of the well-known erratic behaviour of R the bivariate result is more difficult. I conjecture it to be true; however, for the present we only need to consider sets with

$$\alpha = O(n^{-1}\sigma_R\sigma_S) = O(n^3).$$

This content downloaded from
210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms

Conditions under which this is the correct limiting form are as follows. The asymptotic result in the null case may be written

$$\lim_{n\to\infty}\left\{\sum_{\substack{R\leqslant R_0\\ S\leqslant S_0}} P_0(R,S)\Big/\int_{-\infty}^{R_0}\int_{-\infty}^{S_0} f(R,S)\,dR\,dS\right\} = 1,$$

where, as $n\to\infty$,

$$R_0/\sigma_R = O(1), \quad S_0/\sigma_S = O(1),$$
$$R_0/\sigma_R - S_0/\sigma_S = O(n^{-1}).$$

This last condition follows from the approach of ρ to 1. (It is possible that weaker conditions (e.g. $R_0/\sigma_R = O(n^{\frac{1}{2}})$) may suffice; however, this is not known.) For the non-null asymptotic result to hold

$$\mathscr{E}(R)/\sigma_R = O(1), \quad \mathscr{E}(S)/\sigma_S = O(1),$$
$$\mathscr{E}(R)/\sigma_R - \mathscr{E}(S)/\sigma_S = O(n^{-1}),$$

whence $$\sigma_R \log\theta = O(1), \quad \sigma_S\log\phi = O(1),$$

i.e. $$\log\theta = O(n^{-\frac{3}{2}}), \quad \log\phi = O(n^{-\frac{3}{2}}).$$

Putting $$\lim\sigma_R\log\theta = \mu', \quad \lim\sigma_S\log\phi = \mu'',$$

we have $$\lim\mathscr{E}(R)/\sigma_R = \mu'+\mu'' = \lim\mathscr{E}(S)/\sigma_S.$$

Thus the two parameters are asymptotically indistinguishable; each of them merely shifts the bivariate distribution in the direction $R/\sigma_R = S/\sigma_S$. We deduce further that R and S are asymptotically equivalent in the model $\mathscr{H}(\theta,\phi)$, i.e. they have asymptotically equal power for detecting a change in $(\sigma_R\log\theta+\sigma_S\log\phi)$, at least whenever the above is the correct limiting form.

(*d*) From the form of eqn. (9), $R(u)$ and $S(u)$ are respectively (and jointly) sufficient for θ and ϕ; they will therefore provide most-efficient estimators whenever their asymptotic joint distribution is Normal. This brings out a distinction between the present model (which is a pure ranking model) and other models which have been suggested. Thus if a sample of n is drawn from a bivariate Normal population with correlation ρ, and the variate values are replaced by ranks, then Pitman's asymptotic relative efficiency (A.R.E.) of R (or S) in estimating ρ (relative to the product-moment correlation) is known to be $9\pi^{-2}$ (Hotelling & Pabst, 1936; Moran, 1951). Lehmann (1953) considers a certain non-parametric system of alternatives to independence, and shows that R gives the optimum test for these alternatives, for all sample sizes. For the Thurstone–Babington Smith model, where a judge arrives at a ranking of n objects by first making an estimate of the 'value' of each object, these estimates being independently and Normally distributed about a regression depending on the true ranking of the objects (for simplicity the means for the different objects are usually taken to be in arithmetic progression), Stuart (1954) has shown that the A.R.E. of R or S (relative to the product-moment regression coefficient) is $3\pi^{-1}$. It is hoped in a further communication to investigate the relations between the present models and these alternatives.

7. The model $\mathscr{H}(\phi)$

We shall now concentrate on the parameter ϕ. Putting $\theta = 1$ in the basic model (9) gives the 'ϕ-model':

$$\mathscr{H}(\phi):\quad P\{(u)\} = K_n\phi^{S(u)} \tag{12}$$

and $$P\{U_i \prec U_j\} = \pi = \phi/(\phi+\phi^{-1}) \quad (1\leqslant i<j\leqslant n).$$

This content downloaded from
210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms

Writing now (u_n) for (u) to denote the dependence on n, the p.g.f. of $S_n = S(u_n)$ under the null hypothesis is known to be (see, for example, Moran, 1950)

$$M_n(z) = \sum_{(u_n)} P_0\{(u_n)\} z^{S(u_n)} = \prod_{1 \leqslant k \leqslant n} \frac{1}{k}(z^{k-1}+z^{k-3}+\ldots+z^{1-k}), \tag{13}$$

whence the random variable S_n can be expressed as the sum of n independent random variables s_k with p.g.f.'s

$$m_k(z) = \frac{1}{k}\frac{z^k-z^{-k}}{z-z^{-1}} = \frac{\sin kx}{k \sin x},$$

where $z = e^{ix}$. Thus s_k has a uniform distribution on the k points $s_k = k-1, k-3, \ldots, 1-k$. Hence Moran (1950) obtains the cumulant generating function (c.g.f.) of s_k as

$$\log m_k(e^{ix}) = \sum_{m=1}^{\infty} \frac{(ix)^{2m}}{2m!}(-1)^m \frac{B_m 2^{2m-1}}{m}(1-k^{2m}). \tag{14}$$

For the present model we have

$$\begin{aligned} M_{n,\phi}(z) &= M_n(\phi z)/M_n(\phi) \\ &= \prod_{1 \leqslant k \leqslant n} \frac{\phi^k z^k - \phi^{-k} z^{-k}}{\phi z - \phi^{-1} z^{-1}} \frac{\phi - \phi^{-1}}{\phi^k - \phi^{-k}}, \end{aligned} \tag{15}$$

whence S_n can again be expressed as the sum of n independent random variables s_k, with p.g.f.'s

$$m_{k,\phi}(z) = m_k(\phi z)/m_k(\phi).$$

Thus s_k now has a distribution on the k points $k-1, k-3, \ldots, 1-k$ with probabilities in geometric progression, proportional to ϕ^{s_k}. Also

$$K_n^{-1} = M_n(\phi). \tag{16}$$

Putting $\log \phi = \delta$ we have for the c.g.f. of s_k

$$\log m_{k,\phi}(e^{ix}) = \log \frac{\sinh(kix+k\delta)\sinh\delta}{\sinh(ix+\delta)\sinh k\delta};$$

hence, writing

$$\beta_k = \coth k\delta = (\phi^k+\phi^{-k})(\phi^k-\phi^{-k})^{-1},$$

we have

$$\begin{aligned} \kappa_{1,k} &= k\beta_k - \beta_1, \\ \kappa_{2,k} &= (\beta_1^2-1) - k^2(\beta_k^2-1), \\ \kappa_{3,k} &= 2k^3\beta_k(\beta_k^2-1) - 2\beta_1(\beta_1^2-1), \\ \kappa_{4,k} &= 2(\beta_1^2-1)(3\beta_1^2-1) - 2k^4(\beta_k^2-1)(3\beta_k^2-1). \end{aligned}$$

These reduce to Moran's values when $\delta \to 0$. An extreme case is obtained when $k \to \infty$ with ϕ constant $\neq 1$; $\beta_k \sim 1 + 2\phi^{-2k}$ and we find

$$\begin{aligned} \kappa_{1,k} &\sim k - \beta_1 + 2k\phi^{-2k}, \\ \kappa_{2,k} &\sim (\beta_1^2-1) - 4k^2\phi^{-2k}, \\ \gamma_1 &\to -2\beta_1(\beta_1^2-1)^{-\frac{1}{2}} = -(\phi+\phi^{-1}), \\ \gamma_2 &\to 2(3\beta_1^2-1)(\beta_1^2-1)^{-1} = 2+\gamma_1^2. \end{aligned}$$

The corresponding limiting distribution of s_k is geometric;

$$P\{s_k = k-1-2j\} = (1-\phi^{-2})\phi^{-2j} \quad (j = 0, 1, 2, \ldots). \tag{17}$$

This content downloaded from
210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms

Thus even in this case, the central limit theorem will apply; asymptotically,

$$(S_n - \tfrac{1}{2}n(n+1) + n\beta_1)\,(n(\beta_1^2-1))^{-\frac{1}{2}}$$

is a unit Normal variable. We notice that this is essentially a different limit from that considered in § 6; it is not known whether the corresponding limits for $R(u_n)$ and for the bivariate distribution are Normal, or what happens to the correlation between R_n and S_n. In intermediate cases, when $\log\phi \to 0$ slower than $n^{-\frac{3}{2}}$, we would expect the limiting distribution of S_n to be Normal, with

$$\mathscr{E}(S_n) = o(n^2), \quad \operatorname{var} S_n = o(n).$$

8. Decomposition of s_k

In the null case, Feller (1945) has shown that the variables s_k may be taken as

$$s_k = \sum_{1 \leqslant i \leqslant k-1} \operatorname{sgn}(u_k - u_i);$$

this representation may be carried over to the ϕ-model case. (This contradicts a remark of Moran's (1950).) Let us go back to the original paired-comparison approach. Suppose the judge has decided the ${}^{n-1}C_2$ preferences among $U_1, U_2, \ldots, U_{n-1}$ and has as yet no inconsistencies. Then we have an $(n-1)$-ranking (u_{n-1}) say; and the p.g.f. of $S(u_{n-1})$ is $M_{n-1,\phi}(z)$. Now the judge makes the final set of $n-1$ comparisons, $U_i \prec$ or $\succ U_n$ for $i = 1, 2, \ldots, n-1$. These comparisons are independent amongst themselves and of the preferences already decided:

$$P\{U_i \prec U_n\} = \pi \quad (1 \leqslant i \leqslant n-1).$$

Of the 2^{n-1} possible sets of outcomes for these comparisons, only these sets will be consistent which insert U_n into one of the n intervals between the $n-1$ U's already ranked; the probability that U_n is ranked higher than the lower $j-1$ U's and lower than the higher $n-j$ U's is

$$\pi^{j-1}(1-\pi)^{n-j} \quad (1 \leqslant j \leqslant n)$$

independently of the ranking (u_{n-1}).

This set of comparisons gives a ranking (u_n) with

$$S(u_n) = S(u_{n-1}) + 2j - 1 - n = S(u_{n-1}) + s_n,$$

say. The probability that the comparisons are consistent is thus

$$\sum_{1 \leqslant j \leqslant n} \pi^{j-1}(1-\pi)^{n-j} = (\pi^n - (1-\pi)^n)\,(2\pi-1)^{-1} = C_n \quad \text{say;}$$

and we have as before that s_n is a random variable independent of $S(u_{n-1})$ (and of (u_{n-1})) with

$$P\{s_n\} = C_n^{-1}\pi^{\frac{1}{2}(n-1+s_n)}(1-\pi)^{\frac{1}{2}(n-1-s_n)} = \frac{\phi - \phi^{-1}}{\phi^n - \phi^{-n}}\phi^{s_n} \quad (s_n = 2j-1-n;\ 1 \leqslant j \leqslant n). \tag{18}$$

9. An invariance property

We now prove for the ϕ-model an important invariance property, namely, that

$$p_{ij}^{(n)} = P\{u_i < u_j \mid n, \phi, \text{comparisons are consistent}\}$$

depends only on $j-i$ and ϕ, and not on i or n. We prove that

$$p_{i\,i+m-1}^{(n)} = p_{1m}^{(m)} \quad \text{for} \quad 1 \leqslant i < i+m-1 \leqslant n.$$

This content downloaded from
210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms

This will be proved by induction on n. It is true by definition for $n = m$; assume that it has been proved for $n = m, m+1, \ldots, a$. Fix i, with $1 \leqslant i \leqslant a-m+1$. Consider the rankings (u_a) with $S(u_a) = S$ and having the property (property L) that $u_i < u_{i+m-1}$; we put $\sum_{(u_a)|L,S}$ to denote summation over these rankings. We shall write $((u_a)j)$ for a ranking (u_{a+1}) which has $u_{a+1} = j$. We have by the inductive hypothesis

$$p^{(m)}_{1m} = p^{(a)}_{i\,i+m-1} = \sum_S \sum_{(u_a)|L,S} P\{(u_a)\}.$$

Also
$$p^{(a+1)}_{i\,i+m-1} = \sum_S \sum_{(u_{a+1})|L,S} P\{(u_{a+1})\} = \sum_S \sum_{1\leqslant j \leqslant a+1} \sum_{\substack{(u_{a+1})|L,S \\ u_{a+1}=j}} P\{((u_a)j)\}.$$

Now $S((u_a)j) = S(u_a) + 2j - 2 - a$, and $((u_a)j)$ has the property L if and only if (u_a) has; hence

$$\begin{aligned} p^{(a+1)}_{i\,i+m-1} &= \sum_S \sum_{1\leqslant j\leqslant a+1} \sum_{(u_a)|L,S} P\{((u_a)j)\} \\ &= K_{a+1} \sum_{1\leqslant j \leqslant a+1} \phi^{2j-2-a} \sum_S \sum_{(u_a)|L,S} \phi^S \\ &= K_{a+1} \frac{\phi^{a+1} - \phi^{-a-1}}{\phi - \phi^{-1}} K_a^{-1} p^{(m)}_{1m} = p^{(m)}_{1m}. \end{aligned}$$

This proves the result for all i except $i = (a+1) - m + 1$; this case can be proved similarly by considering rankings $(u_{a+1}) = (j(u_a))$. Hence by induction on a we have finally $p^{(n)}_{i\,i+k} = p_k$ $(1 \leqslant i < i+k \leqslant n)$, where the $\{p_k\}$ depend only on ϕ.

We may obtain the $\{p_k\}$ explicitly as follows: we write $(i(u_{k-2})j)$ for a ranking (u_k) which has $u_1 = i$, $u_k = j$. Then

$$S(i(u_{k-2})j) = S(u_{k-2}) + 2(j-i) + \operatorname{sgn}(i-j)$$

and
$$\begin{aligned} p_{k-1} = p^{(k)}_{1k} &= \sum_{1\leqslant i<j\leqslant k} \sum_{(u_{k-2})} P\{(i(u_{k-2})j)\} \\ &= K_k \sum_{1\leqslant i<j\leqslant k} \phi^{2(j-i)-1} \sum_{(u_{k-2})} \phi^{S(u_{k-2})} \\ &= K_k(\phi^{2k-1} - k\phi + (k-1)\,\phi^{-1})\,(\phi-\phi^{-1})^{-2} K^{-1}_{k-2}, \end{aligned}$$

whence for any i $(1 \leqslant i < i+k \leqslant n)$

$$\begin{aligned} \gamma_k = \mathscr{E}(\operatorname{sgn}(u_{i+k} - u_i)) &= 2p_k - 1 \\ &= (k+1)\frac{\phi^{k+1} + \phi^{-k-1}}{\phi^{k+1} - \phi^{-k-1}} - k\frac{\phi^k + \phi^{-k}}{\phi^k - \phi^{-k}} = (k+1)\,\beta_{k+1} - k\beta_k \end{aligned} \tag{19}$$

in the notation of § 7.

From this result we have

$$\mathscr{E}(S_n) = \mathscr{E} \sum_{1\leqslant i<j\leqslant n} \operatorname{sgn}(u_j - u_i) = \sum_{1\leqslant i<j\leqslant n} \gamma_{j-i} = \sum_{1\leqslant k\leqslant n-1} (n-k)\,\gamma_k,$$
$$\mathscr{E}(S_n) - \mathscr{E}(S_{n-1}) = \sum_{1\leqslant k\leqslant n-1} \gamma_k = k\beta_k - \beta_1,$$

in agreement with § 7. Similarly we may obtain $\mathscr{E}(R_n)$:

$$\begin{aligned} \mathscr{E}(R_n) = \mathscr{E} \sum_{1\leqslant i<j\leqslant n} (j-i)\operatorname{sgn}(u_j - u_i) &= \sum_{1\leqslant i<j\leqslant n} (j-i)\,\gamma_{j-i} \\ &= \sum_{1\leqslant k\leqslant n-1} k(n-k)\,\gamma_k = \sum_{1\leqslant k\leqslant n} k(2k-n-1)\,\beta_k \\ &= \sum_{1\leqslant k\leqslant n} (2k-n-1)\,(k\beta_k - \beta_1). \end{aligned} \tag{20}$$

This content downloaded from
210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms

10. Interpretation of the parameters

Having now considered some of the consequences of the Assumptions of § 5, we can give interpretations of the two parameters θ and ϕ. Each of them corresponds to a certain type of departure from the null hypothesis, that the judge cannot detect the real differences between the objects:

(*a*) The model $\mathscr{H}(\theta)$, by its association with the Bradley–Terry model, may be considered to be assigning a weight $w(i) = \theta^{2i}$ to each of the objects U_i to be ranked; this weight remains constant throughout the process of making the paired-comparisons and arriving at the ranking. The paired-comparisons are made according to the rule

$$P\{U_i \prec U_j\} = \frac{w(j)}{w(i)+w(j)} = \frac{\theta^{2j}}{\theta^{2i}+\theta^{2j}} = \frac{\theta^k}{\theta^k+\theta^{-k}},$$

where $k = j-i$ for $1 \leqslant i < j \leqslant n$. Thus this probability depends on the difference between the *a priori* ranks of the objects, but not on their absolute positions. While this seems an attractive hypothesis, it must be remembered that the rejection of inconsistent sets of comparisons introduces a distortion, the effect of which it is difficult to assess.

Table 1. $\mathscr{H}(\phi)$: *Values of* $p_k = P\{u_i < u_{i+k}\}$

k \ ϕ^2 / π	1·2 $0{\cdot}\dot{5}\dot{4}$	1·5 0·6	2·0 $0{\cdot}\dot{6}$
1	0·5455	0·6000	0·6667
2	0·5754	0·6632	0·7619
3	0·6049	0·7215	0·8386
4	0·6337	0·7737	0·8946
5	0·6617	0·8191	0·9339
6	0·6887	0·8577	0·9599
7	0·7145	0·8897	0·9763
8	0·7392	0·9155	0·9862
9	0·7626	0·9361	0·9922

(*b*) The model $\mathscr{H}(\phi)$ puts for the paired-comparisons

$$P\{U_i \prec U_j\} = \pi = \frac{\phi}{\phi+\phi^{-1}} \quad (1 \leqslant i < j \leqslant n), \tag{21}$$

this probability being independent of both i and j. This may not seem as reasonable an assumption as that for the θ-model above; however, the quantities we are really interested in are the p_{ij} of eqn. (2).

Table 1 gives the values of

$$p_k = P\{u_i < u_{i+k}\} \quad (1 \leqslant i < i+k \leqslant n)$$

(from (19)) for three values of ϕ^2, namely, 1·2, 1·5 and 2·0 corresponding to

$$\pi = \tfrac{6}{11} = 0{\cdot}\dot{5}\dot{4}, 0{\cdot}6, \tfrac{2}{3}.$$

It was proved in § 9 that p_k is independent of i and n.

This content downloaded from
210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms

It will be seen that these values progress in a very reasonable manner—the further apart in the true ranking two objects are, the more probably will the judge put them in the correct order.

From eqn. (18) we have a further interpretation of the parameter ϕ. Given that the objects $U_1, U_2, \ldots, U_{n-1}$ have been compared and a ranking obtained, the judge now compares U_n with each of these; each comparison being independent of the rest, and according to (21). Now the restriction that only consistent comparisons are accepted means (eqn. (18)) that U_n takes the jth position in the ranking with probability proportional to ϕ^{2j}, i.e. (if $\phi > 1$) the probability is largest that U_n take its correct position ($j = n$), and falls off geometrically as j decreases to 1.

(c) The general model $\mathscr{H}(\theta, \phi)$ (eqn. (9)) puts

$$P\{U_i \prec U_j\} = \frac{\theta^k \phi}{\theta^k \phi + \theta^{-k}\phi^{-1}} = \tfrac{1}{2} + \tfrac{1}{2}\tanh(k\log\theta + \log\phi)$$

for $k = j-1$, $1 \leqslant i < j \leqslant n$. This is a simple monotonic function of $k\log\theta + \log\phi$; thus the effects of θ and ϕ are in a sense additive. The term $\log\phi$ gives a departure from the null hypothesis which is the same for all pairs (i, j); the other term adds to this a departure depending linearly on $k = j - i$. This model would appear to be an attractive approximation to the general model $\mathscr{H}_\lambda$ (eqn. (1)).

11. Applications of the ϕ-model

From the form of the model $\mathscr{H}(\phi)$ (eqn. (12)) we have that $S(u)$ is sufficient for ϕ; problems of estimation and of hypothesis-testing will therefore be most naturally expressed in terms of S. However, since R and S are asymptotically equivalent, any procedure using S may if desired be replaced (approximately) by one using R.

We shall use the following notation:

$$\delta = \log\phi, \quad \mathscr{E}_\phi = \mathscr{E}(S \mid \mathscr{H}(\phi)), \quad V_\phi = \operatorname{var}(S \mid \mathscr{H}(\phi)),$$
$$\Delta = \mathscr{E}_\phi V_\phi^{-\frac{1}{2}}, \quad V_0 = \operatorname{var}(S \mid \mathscr{H}_0) = \tfrac{1}{18}n(n-1)(2n+5),$$

which though not entirely consistent, will not lead to confusion.

Approximations. From Moran's expression (14) we have an expansion of $\log M_n(\phi)$ (eqn. (13)) in powers of δ:

$$\log M_n(\phi) = \frac{\delta^2}{6}\sum_{k=2}^{n}(k^2-1) - \frac{\delta^4}{180}\sum_{k=2}^{n}(k^4-1) + \ldots$$
$$= \tfrac{1}{2}V_0\delta^2 - \frac{n(n-1)}{5400}(6n^3+21n^2+31n+31)\delta^4 + \ldots$$

whence from (15)

$$\mathscr{E}_\phi = \frac{\partial}{\partial z}\frac{M_n(\phi z)}{M_n(\phi)}\bigg|_{z=1} = \phi\frac{d}{d\phi}\log M_n(\phi) = V_0\delta - \frac{n(n-1)}{1350}(6n^3+21n^2+31n+31)\delta^3 + \ldots,$$

$$V_\phi = \phi\frac{d}{d\phi}\log M_n(\phi) + \phi^2\frac{d^2}{d\phi^2}\log M_n(\phi) = V_0\left(1 - \frac{\delta^2}{25}\,\frac{6n^3+21n^2+31n+31}{2n+5} + \ldots\right).$$

We notice the curious result

$$V_\phi = \phi\frac{d}{d\phi}\mathscr{E}_\phi.$$

This content downloaded from
210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms

We obtain further
$$V_\phi = V_0\left(1 - \frac{27}{25}\frac{\Delta^2}{n} + \left(\frac{27}{50} + \frac{9936}{30625}\Delta^2\right)\frac{\Delta^2}{n^2} + \ldots\right).$$

Hence
$$V_\phi \simeq V'_\phi = V_0 - \frac{27}{25}\frac{\mathscr{E}_\phi^2}{n}.$$

Table 2 gives values of the percentage error of the approximation V'_ϕ for various values of n and ϕ, with the corresponding values of Δ. It will be seen that the approximation is remarkably good even for small n, provided Δ is not too large; except for very small n, the approximation improves with increasing n if Δ is held constant (compare the three entries with $\Delta = 1{\cdot}72$ or $1{\cdot}73$).

Table 2. $\mathscr{H}(\phi)$: *Percentage error of the approximation* V'_ϕ, *with corresponding values of* $\Delta = \mathscr{E}_\phi V_\phi^{-\frac{1}{2}}$

% error = $100(V'_\phi - V_\phi)\, V_\phi^{-1}$

n	$\phi^2 = 1{\cdot}1$		$\phi^2 = 1{\cdot}2$		$\phi^2 = 1{\cdot}5$		$\phi^2 = 2{\cdot}0$	
	% error	Δ	% error	Δ	% error	Δ	% error	Δ
2	+0·11	0·05	+0·38	0·09	+1·92	0·20	+ 5·75	0·35
4	+0·08	0·14	+0·31	0·27	+1·31	0·61	+ 3·15	1·07
6	+0·05	0·25	+0·17	0·49	+0·08	1·12	− 4·09	2·01
8	+0·02	0·39	−0·02	0·75	−2·35	1·73	−19·28	3·17
10	−0·02	0·54	−0·30	1·04	−6·63	2·44		
12	−0·07	0·70	−0·71	1·36				
14	−0·14	0·88	−1·32	1·72				
16	−0·23	1·07	−2·16	2·11				
18	−0·34	1·28						
20	−0·49	1·50						
22	−0·67	1·73						
24	−0·90	1·97						
26	−1·18	2·23						

Under $\mathscr{H}(\phi)$, S is approximately Normally distributed with mean $\mathscr{E}_\phi$ and variance V_ϕ; for small departures from the null hypothesis, the variance is approximately V'_ϕ or more crudely still, V_0.

In the following it should be remembered that we are investigating the judges, and not the underlying ranking. This was given *a priori*; each judge makes repeated attempts to reproduce it; we are concerned with

(*a*) estimating the degree to which a judge can reproduce the ranking, i.e. (assuming he operates according to the ϕ-model) estimating his ϕ;

(*b*) comparing two attempts at ranking;

(*c*) testing for differences between the attempts of several judges.

Estimation. Suppose a judge has produced l different rankings of the same n objects; it

This content downloaded from
210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms

is desired to estimate his ϕ. We shall do this by the method of maximum likelihood. We have from eqns. (12) and (16)

$$\log P\{(u)_1, (u)_2, \ldots, (u)_l\} = \log \phi \sum_{i=1}^{l} S(u)_i - l \log M_n(\phi),$$

whence the estimation equation is

$$\bar{S} = \phi \frac{d}{d\phi} \log M_n(\phi) \mid_{\phi=\hat{\phi}} = \mathscr{E}_{\hat{\phi}} = \mathscr{E}(S \mid \hat{\phi});$$

thus the model is fitted by its first moment. For small departures from $\mathscr{H}_0$ we have approximately

$$\bar{S} \simeq V_0 \log \hat{\phi} = V_0 \hat{\delta}.$$

We may give an approximate confidence interval for ϕ; since under $\mathscr{H}(\phi)$, $\bar{S}$ is approximately Normal with mean $\mathscr{E}_\phi$ and variance V'_ϕ/l, we have approximately

$$P\{\bar{S} - \lambda_\alpha \surd(V'_\phi/l) < \mathscr{E}_\phi < \bar{S} + \lambda_\alpha \surd(V'_\phi/l) \mid \phi\} \simeq 1 - 2\alpha,$$

where λ_α is the 100α % point of the unit Normal distribution; more crudely

$$P\{\bar{S} - \lambda_\alpha \surd(\hat{V}'_\phi/l) < V_0 \log \phi < \bar{S} + \lambda_\alpha \surd(\hat{V}'_\phi/l) \mid \phi\} \simeq 1 - 2\alpha,$$

where

$$\hat{V}'_\phi = V_0 - \frac{27}{25} \frac{\bar{S}^2}{n}. \tag{22}$$

Test for consistency between two judges. Suppose two judges have produced rankings (u_1) and (u_2), according to the ϕ-model, with respective parameters ϕ_1 and ϕ_2. Since $S(u)_1$ and $S(u)_2$ are approximately Normally distributed with means $V_0 \log \phi_1$ and $V_0 \log \phi_2$, and with variances V_0, we may test whether $\phi_1 = \phi_2$ by referring

$$(S(u)_1 - S(u)_2)/\surd(2V_0)$$

to Normal tables. An improved test is obtained by replacing V_0 by $\hat{V}'_\phi$ from (22); i.e. we refer

$$(S(u)_1 - S(u)_2) \Big/ \sqrt{\left\{2V_0 - \frac{27}{50n}(S(u)_1 + S(u)_2)^2\right\}}$$

to Normal tables.

This test is asymptotically equivalent to the likelihood ratio test.

Test for consistency between several judges. Suppose m judges have each produced l rankings with scores S_{ij} ($j = 1, 2, \ldots, l$; $i = 1, 2, \ldots, m$). We associate with each judge a value ϕ_i of the parameter ϕ; then under the hypothesis $\phi_1 = \phi_2 = \ldots = \phi_m$, the 'between judges sum of squares'

$$l \sum_{i=1}^{m} (\bar{S}_{i.} - \bar{S}_{..})^2$$

(with the usual notation for means) is approximately distributed as $V_0 \chi^2_{m-1}$, and the 'within judges sum of squares'

$$\sum_{i=1}^{m} \sum_{j=1}^{l} (S_{ij} - \bar{S}_{i.})^2$$

is approximately distributed as $V_0 \chi^2_{m(l-1)}$. Thus we may test for differences between judges and for inconsistency within judges. More accurate tests would replace V_0 by the estimated $\hat{V}'_\phi$ as before.

This content downloaded from
210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms

The approximate powers of the above tests are immediately obtainable from standard Normal theory. Thus, for example, the power of the test to detect a difference between ϕ_1 and ϕ_2 when a two-tailed test is used is approximately

$$1-F(\lambda_\alpha-\mu)+F(-\lambda_\alpha-\mu),$$

where λ_α is as before the 100α % Normal point, and

$$\mu = \sqrt{(\tfrac{1}{2}V_0)}\,(\log\phi_1-\log\phi_2).$$

12. The conditional expectation of R

We now obtain an expression giving the conditional expectation of R, given S; this is the same for the ϕ-model as for the null case. Writing as in § 6 $M_0(x,y)$ for the null-hypothesis p.g.f. of R and S,

$$M_0(x,y) = \sum_{R,S} x^R y^S P_0(R,S),$$

the p.g.f. for the ϕ-model is $M_0(x,y\phi)/M_0(1,\phi)$. Thus

$$\mathscr{E}_\phi(R) = \frac{\partial}{\partial x}\frac{M_0(x,\phi)}{M_0(1,\phi)}\bigg|_{x=1}. \tag{23}$$

However, we have for the null hypothesis

$$\frac{\partial}{\partial x}M_0(x,y)\,|_{x=1} = \sum_S y^S P_0(S)\,\mathscr{E}_0(R\mid S),$$

and so from (20) and (23)

$$\begin{aligned}\sum_S \phi^S P_0(S)\,\mathscr{E}_0(R\mid S) &= \mathscr{E}_\phi(R)\,M_0(1,\phi)\\ &= (n!)^{-1}\sum_{2\leqslant k\leqslant n}(2k-n-1)\left\{k\frac{\phi^k+\phi^{-k}}{\phi^k-\phi^{-k}}-\frac{\phi+\phi^{-1}}{\phi-\phi^{-1}}\right\}\prod_{1\leqslant i\leqslant n}\frac{\phi^i-\phi^{-i}}{\phi-\phi^{-1}}\\ &= (n!)^{-1}\sum_{2\leqslant k\leqslant n}(2k-n-1)\,\{(k-1)\,\phi^{k-1}+(k-3)\,\phi^{k-3}+\ldots+(1-k)\,\phi^{1-k}\}\\ &\qquad\times\prod_{\substack{2\leqslant i\leqslant n\\ i\neq k}}\{\phi^{i-1}+\phi^{i-3}+\ldots+\phi^{1-i}\}.\end{aligned}$$

13. Conclusion

The ϕ-model exhibits several features of simplicity; however, difficulty has been encountered in attempting to evaluate such quantities as $P\{u_i=j\}$ and $P\{u_i=j,\,u_k=l\}$; knowledge of the latter quantity would enable the variance of R and covariance of R and S to be obtained; hence also the conditional variance of R, given S.

It is hoped in a further communication to present:

(a) an investigation of the result of § 12;

(b) a comparison of the present distributions of R and S with those obtained from other models, e.g. sampling from a bivariate Normal distribution;

(c) an extension of the method to cover those cases where the true ranking is not known *a priori*.

This content downloaded from 210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms

APPENDIX I

Demonstration that Kendall's t_k is not essentially a 'generalized matching coefficient'

To satisfy eqn. (6) with $C(u) = S(u)$ (the score for Kendall's t_k), we could take (for $n \leqslant 9$)

$$\alpha_{ij} = j\,.\,10^i$$

when $A(u)$ is simply the decimal number with digits u_i; e.g. $A(3, 5, 2, 4, 1) = 35{,}241$. We then define the function $F(x)$ suitably at the $n!$ different values of x (e.g. between 12,345 and 54,321) and have (6). However, no matter what $\{\alpha_{ij}\}$ are, we cannot satisfy the inverse relation (7) for all rankings (u) without having some of the values $F^{-1}(S) = g_S$ equal for different S. Thus with $n = 4$, consider the rankings with $u_1 = 1$:

(u)	$S(u)$	$F^{-1}(S(u))$		$A(u)$
(1, 2, 3, 4)	6	g_6	=	$\alpha_{11}+\alpha_{22}+\alpha_{33}+\alpha_{44}$
(1, 2, 4, 3)	4	g_4	=	$\alpha_{11}+\alpha_{22}+\alpha_{34}+\alpha_{43}$
(1, 3, 2, 4)	4	g_4	=	$\alpha_{11}+\alpha_{23}+\alpha_{32}+\alpha_{44}$
(1, 3, 4, 2)	2	g_2	=	$\alpha_{11}+\alpha_{23}+\alpha_{34}+\alpha_{42}$
(1, 4, 2, 3)	2	g_2	=	$\alpha_{11}+\alpha_{24}+\alpha_{32}+\alpha_{43}$
(1, 4, 3, 2)	0	g_0	=	$\alpha_{11}+\alpha_{24}+\alpha_{33}+\alpha_{42}$

Hence for any $\{\alpha_{ij}\}$ we must have

$$g_6 - 2g_4 + 2g_2 - g_0 = 0.$$

There are three similar equations obtainable from the rankings with $u_1 = 2, 3, 4$ in turn; together, they imply that $g_6 = g_{-6}$. *A fortiori*, (7) is impossible for $n > 4$.

APPENDIX II

To prove that Assumption B contains Assumption C we shall exhibit various pairs of inverse rankings. We shall use the 'cycle' notation; thus with $n = 6$, the cycle (1, 5, 6, 2) (or (5, 6, 2, 1) or (6, 2, 1, 5) or (2, 1, 5, 6)) will represent the ranking

$$u_1 = 5,\ u_5 = 6,\ u_6 = 2,\ u_2 = 1;\quad u_3 = 3,\ u_4 = 4;$$

i.e. the ranking (5, 1, 3, 4, 6, 2). The inverse ranking (2, 6, 3, 4, 1, 5) is represented by the cycle (2, 6, 5, 1), which is the reflexion of the first cycle. Thus any cycle whose reflexion is not equivalent to itself will specify a pair of inverse rankings. From the following cycles we now obtain as in § 5 the required relations between the λ_{ij}'s

Cycle	Relation		
To prove C (i):			
(1, 2, 3)	$\lambda_{12} = \lambda_{23}$		
(2, 3, 4)	$\lambda_{23} = \lambda_{34}$		
etc.		Hence	$\lambda_{i\,i+1} = \lambda_1 \quad (1 \leqslant i \leqslant n-1)$
(1, 2, 3, 4)	$\lambda_{12}\lambda_{13} = \lambda_{24}\lambda_{34}$	i.e.	$\lambda_{13} = \lambda_{24}$
(2, 3, 4, 5)	$\lambda_{23}\lambda_{24} = \lambda_{35}\lambda_{45}$	i.e.	$\lambda_{24} = \lambda_{35}$
etc.		Hence	$\lambda_{i\,i+2} = \lambda_2 \quad (1 \leqslant i \leqslant n-2)$
(1, 2, 3, 4, 5)	$\lambda_{12}\lambda_{13}\lambda_{14} = \lambda_{25}\lambda_{35}\lambda_{45}$	i.e.	$\lambda_{14} = \lambda_{25}$
etc.		Hence	$\lambda_{i\,i+3} = \lambda_3 \quad (1 \leqslant i \leqslant n-3)$
and so on.		Hence	$\lambda_{i\,i+k} = \lambda_k \quad (1 \leqslant i < i+k \leqslant n)$
To prove C (ii):			
(1, 3, 4, 2)	$\lambda_{13}\lambda_{23}\lambda_{24} = \lambda_{12}\lambda_{14}\lambda_{34}$	i.e.	$\lambda_1\lambda_3 = \lambda_2^2$
(1, 4, 5, 2)	$\lambda_2\lambda_4 = \lambda_3^2$		
(1, 5, 6, 2)	$\lambda_3\lambda_5 = \lambda_4^2$		
etc.		Hence	$\lambda_{k-1}\lambda_{k+1} = \lambda_k^2 \quad (2 \leqslant k \leqslant n-1)$

This content downloaded from
210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms

REFERENCES

Barton, D. E. & David, F. N. (1956). 'Spearman's Rho' and the matching problem. *Brit. J. Statist. Psychol.* **9**, 69–73.

Bradley, R. A. & Terry, M. A. (1952). Rank analysis of incomplete block designs. I. *Biometrika*, **39**, 324–45.

Daniels, H. E. (1944). The relation between measures of correlation in the universe of sample permutations. *Biometrika*, **33**, 129–35.

Daniels, H. E. (1950). Rank correlation and population models. *J. R. Statist. Soc.* B, **12**, 171–81.

Feller, W. (1945). The fundamental limit theorems in probability. *Bull. Amer. Math. Soc.* **51**, 800–32.

Haden, H. G. (1947). A note on the distribution of the different orderings of n objects. *Proc. Camb. Phil. Soc.* **43**, 1–9.

Hoeffding, W. (1948). A class of statistics with asymptotically Normal distribution. *Ann. Math. Statist.* **19**, 293–325.

Hotelling, H. & Pabst, M. R. (1936). Rank correlation and tests of significance involving no assumption of Normality. *Ann. Math. Statist.* **7**, 29–43.

Kendall, M. G. (1950). Discussion on symposium on ranking methods. *J. R. Statist. Soc.* B, **12**, 189.

Kendall, M. G. (1955). *Rank Correlation Methods*. London: Charles Griffin and Co. Ltd.

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

Moran, P. A. P. (1950). Recent developments in ranking theory. *J. R. Statist. Soc.* B, **12**, 153–62.

Moran, P. A. P. (1951). Partial and multiple rank correlation. *Biometrika*, **38**, 26–32.

Smith, B. Babington (1950). Discussion on Professor Ross's paper. *J. R. Statist. Soc.* B, **12**, 54.

Stuart, A. (1954). The correlation between variate-values and ranks in samples from a continuous distribution. *Brit. J. Statist. Psychol.* **7**, 37–44.

This content downloaded from
210.36.46.141 on Fri, 03 Sep 2021 08:04:29 UTC
All use subject to https://about.jstor.org/terms