THE ASYMPTOTIC DISTRIBUTIONS OF STATISTICS ARISING IN CERTAIN NON-PARAMETRIC TESTS

by SAMUEL D. SILVEY

(Received 12th October, 1953)

§ 1. *Introduction.* In a previous paper [5] the equivalence of randomisation and normal theory distributions of linear combinations was discussed. In the present paper we discuss the asymptotic randomisation distributions of statistics used in analysis of variance and in a closely related problem which includes, in particular, the " problem of m-rankings ". Kruskal [4] has studied the first of these questions in the case where observations are replaced by ranks.

The most general sufficient conditions yet given for asymptotic normality of the randomisation distribution of a linear combination are contained in the paper [5] and in a paper by Hoeffding [2]. The latter paper gives the simplest form of these conditions and we begin by using Hoeffding's criterion in a discussion of the joint distribution of several linear combinations.

§ 2. *Joint randomisation distributions of linear combinations.* For a sequence $\{Z_n\}$ of ordered n-tuples $Z_n=(z_{n1}, z_{n2}, \dots, z_{nn})$ in which the variance of $z_{n1}, z_{n2}, \dots, z_{nn}$ is non-zero for each n, we write

$$z_{n\cdot}=\frac{1}{n}\sum_{\tau=1}^{n} z_{n\tau}, \quad z'_{n\tau}=\frac{z_{n\tau}-z_{n\cdot}}{\sqrt{\sum_{\tau=1}^{n}(z_{n\tau}-z_{n\cdot})^2}}, \quad \text{for } \tau=1, 2, \dots, n,$$

and $M_n(z)=\max_{1\leqslant\tau\leqslant n}|z'_{n\tau}|$.

Let $[A_n]$ and $[Y_{in}]$, $i=1, 2, \dots, p$, be sequences of ordered n-tuples, the elements of each n-tuple having non-zero variance. For each n let $x_{n1}, x_{n2}, \dots, x_{nn}$ be a set of n random variables whose joint distribution is defined by $P\{x_{n\tau}=a_{n\rho_\tau}, \tau=1, 2, \dots, n\}=\frac{1}{n!}$ for each permutation $(\rho_1, \rho_2, \dots, \rho_n)$ of $(1, 2, \dots, n)$ and let $r_{in}=\sqrt{n}\sum_{\tau=1}^{n} y'_{in\tau}x'_{n\tau}$, $i=1, 2, \dots, p$. Then Hoeffding has shown that r_{in} is asymptotically normally distributed with zero mean and unit variance provided $\sqrt{n}M_n(y_i)M_n(a)\to 0$ as $n\to\infty$. We use this result to prove the following theorem.

2.1. THEOREM. *If $\mathbf{V}_n$ is the covariance matrix of the random variables r_{in}, $i=1, 2, \dots, p$, then the asymptotic joint distribution of these random variables is normal provided*

(i) $\sqrt{n}M_n(a)\max_{1\leqslant i\leqslant p}M_n(y_i)\to 0$ *and*

(ii) $\mathbf{V}_n$ *tends to a non-singular matrix* $\mathbf{V}$ *as* $n\to\infty$.

We show that any linear combination of $r_{in}, r_{2n}, \dots, r_{pn}$ is asymptotically normal and the result follows from a theorem of Cramer [1], by the same argument as is used by Wald and Wolfowitz [6].

Let d_i, $i=1, 2, \dots, p$, be any p real numbers not all zero and let $\mathbf{d}$ denote the column vector with components $d_1, d_2, \dots, d_p$.

https://doi.org/10.1017/S2040618500032998 Published online by Cambridge University Press

We consider the random variable $R_n = \sum_{i=1}^{p} d_i r_{in} = \sqrt{n} \sum_{\tau=1}^{n} z_{n\tau} x'_{n\tau}$, say, where

$$z_{n\tau} = \sum_{i=1}^{p} d_i y'_{in\tau}, \quad \tau = 1, 2, \ldots, n.$$

Then $z'_{n\tau} = \dfrac{\sum_{i=1}^{p} d_i y'_{in\tau}}{\sqrt{(\mathbf{d}'\mathbf{V}_n\mathbf{d})}}$, where $\mathbf{d}'$ denotes the transpose of $\mathbf{d}$.

Since $\mathbf{V}_n \to \mathbf{V}$ where $\mathbf{V}$ is non-singular, $\dfrac{1}{\mathbf{d}'\mathbf{V}_n\mathbf{d}} = O(1)$ for all non-zero $\mathbf{d}$, and so, for such $\mathbf{d}$ $M_n(z) = O\left[\max_{1 \leqslant i \leqslant p} M_n(y_i)\right]$, and $\sqrt{n} M_n(z) M_n(a) \to 0$ as $n \to \infty$. Hence every linear combination of $r_{1n}, r_{2n}, \ldots, r_{pn}$ is asymptotically normally distributed, and normality of the asymptotic joint distribution of these random variables follows.

§ 3. *Analysis of Variance.* In this section we consider the application of Theorem 2.1. to analysis of variance—firstly in the case of one-way classification. The set of numbers $a_{n1}, a_{n2}, \ldots, a_{nn}$ is split up into p classes containing respectively $n_1, n_2, \ldots, n_p$ of the numbers where $n_1 + n_2 + \ldots + n_p = n$. Hence it will be convenient to have the alternative notation $a_{n11}, a_{n12}, \ldots, a_{n1n_1}$; $a_{n21}, \ldots, a_{n2n_2}$; $\ldots$; $\ldots$, $a_{np1}, \ldots, a_{npn_p}$ for the set $a_{n1}, a_{n2}, \ldots, a_{nn}$, and similarly for the set of random variables $x_{n1}, x_{n2}, \ldots, x_{nn}$.

$a_{ni\tau}$, $\tau = 1, 2, \ldots, n_i$, will be regarded as independent observations of a random variable ξ_i, say, $i = 1, 2, \ldots, p$. Interest lies in testing the hypothesis H_0 that the ξ's are independent random variables with the same distribution against the alternative H_1 that $E(\xi_i) \neq E(\xi_j)$ for some i, j. Instead of making this a parametric test by assuming normality of the ξ's, we discuss the test based on the statistic $s^2 = \sum_{i=1}^{p} n_i x_{ni\cdot}^2 - n x_{n\cdot}^2$, where $n_i x_{ni\cdot} = \sum_{\tau=1}^{n_i} x_{ni\tau}$, and show that under very wide conditions such a test is equivalent to the usual test based on the assumption of normality.

If we let

$$\sqrt{\left\{\frac{1}{n} \sum_{\tau=1}^{n} (x_{n\tau} - x_{n\cdot})^2\right\}}\, r_{in} = c_{in} \left\{ \sum_{k=1}^{i} n_k x_{nk\cdot} - x_{n\,i+1\cdot} \sum_{k=1}^{i} n_k \right\}, \quad i = 1, 2, \ldots, p-1,$$

$$= \sqrt{n} x_{n\cdot}, \qquad i = p,$$

where

$$1 = c_{in}^2 \left[\sum_{k=1}^{i} n_k + \frac{1}{n_{i+1}} \left(\sum_{k=1}^{i} n_k \right)^2 \right] \qquad i = 1, 2, \ldots, p-1,$$

then $\dfrac{n s^2}{\sum_{\tau=1}^{n} (x_{n\tau} - x_{n\cdot})^2}$ reduces to the form $r_{1n}^2 + r_{2n}^2 + \ldots + r_{p-1\,n}^2$.

r_{in}, $i = 1, 2, \ldots, p-1$, are orthogonal linear combinations, *i.e.*, their correlation matrix is the unit matrix. Further, if r_{in} is written in the form $n^{\frac{1}{2}} \sum_{\tau=1}^{n} y'_{ni\tau} x'_{n\tau}$, then

$$y'_{ni\tau} = c_{in} \quad \tau = 1, 2, \ldots, n_1 + n_2 + \ldots + n_i,$$

$$= -c_{in} \frac{n_1 + n_2 + \ldots + n_i}{n_{i+1}}, \quad \tau = n_1 + n_2 + \ldots + n_i + 1, \ldots, n_1 + n_2 + \ldots + n_{i+1},$$

$$= 0, \quad \text{otherwise},$$

https://doi.org/10.1017/S2040618500032998 Published online by Cambridge University Press

and it is easily shown that

$$\max_{1\leqslant i\leqslant p-1}\{M_n(y)\}\leqslant[\min_{1\leqslant i\leqslant p} n_i]^{-\frac{1}{2}}.$$

Hence r_{in}, $i=1, 2, \ldots, p-1$, are asymptotically jointly normal and $\dfrac{ns^2}{\sum_{\tau=1}^{n}(x_{n\tau}-x_{n\cdot})^2}$ is asymptotically distributed in the χ^2-form with $(p-1)$ degrees of freedom, provided

$$\sqrt{n}\,[\min_{1\leqslant i\leqslant p} n_i]^{-\frac{1}{2}} M_n(a)\to 0 \text{ as } n\to\infty.$$

In particular if $\dfrac{n}{\min_{1\leqslant i\leqslant p} n_i}$ is bounded then the above statistic has an asymptotic χ^2-distribution provided $M_n(a)\to 0$ as $n\to\infty$. Again if $A_n=1, 2, \ldots, n$, *i.e.*, if the observations are ranks, then $M_n(a)=O(n^{-\frac{1}{2}})$ and the result holds if $\min_{1\leqslant i\leqslant p} n_i\to\infty$ as $n\to\infty$.

It is easy to deduce that, in these circumstances, the randomisation test of the hypothesis H_0 against H_1 is asymptotically equivalent to the usual normal theory test. The practical implications of this result are that in dealing with a one-way classification analysis of variance the normal assumption can be dropped, provided the number of observations in each class is reasonably large and the set of observations does not have one small subset containing observations differing widely from the remaining observations. Furthermore, this type of argument can clearly be extended to higher classifications.

§ 4. *The problem of m-rankings.* The problem considered in this section is as follows. $\{A_{in}\}$, $i=1, 2, \ldots, p$, are p sequences of ordered n-tuples $A_{in}=a_{in\tau}$, $\tau=1, 2, \ldots, n$. A_{in}, $i=1, 2, \ldots, p$ are regarded as observations on pn independent random variables $\xi_{in\tau}$, say, $i=1, 2, \ldots, p$, $\tau=1, 2, \ldots, n$. A null-hypothesis H_0 states that the distribution of $\xi_{in\tau}$ depends on i but not on τ, while an alternative H_1 states that $\xi_{in\tau}$ is of the form $\xi_{in\tau}=\xi'_{in}+\alpha_\tau$ where ξ'_{in} is a random variable whose distribution depends on i, but not on τ, and α_τ, $\tau=1, 2, \ldots, n$, is a set of constants with non-zero variance. Under H_0, for each given i, all permutations of the set (A_{in}) are equally likely. Hence we define p mutually independent sets (X_{in}), $i=1, 2, \ldots, p$, of n random variables where $X_{in}=(x_{in1}, x_{in2}, \ldots, x_{inn})$ and where $P\{x_{in\tau}=a_{in\rho_\tau}, \tau=1, 2, \ldots, n\}=\dfrac{1}{n!}$ for each permutation $(\rho_1, \rho_2, \ldots, \rho_n)$ of $(1, 2, \ldots, n)$. Then we base a test of H_0 against H_1 on the statistic $s^2=\sum_{\tau=1}^{n} x'^2_{\cdot n\tau}$ where $px'_{\cdot n\tau}=\sum_{i=1}^{p} x'_{in\tau}$. This test is, of course, a well-known non-parametric test. Here we discuss the asymptotic distribution of s^2.

If $W=\dfrac{s^2-E(s^2)}{\sqrt{\{\text{var }(s^2)\}}}$, then W can be expressed in the form $W=\dfrac{\sqrt{2}}{\sqrt{\{m(m-1)\}}}\sum_{\substack{i,j=1\\ i<j}}^{p} r_{ijn}$, where $r_{ijn}=\sqrt{n}\sum_{\tau=1}^{n} x'_{in\tau}x'_{jn\tau}$. We now require the following result.

4.1. THEOREM. *r_{ijn}, $i, j=1, 2, \ldots, p$, $i<j$, are asymptotically independent normal random variables if $n^{\frac{1}{2}}\max_{1\leqslant i\leqslant p}[M_n(a_i)]\to 0$ as $n\to\infty$.*

We prove the theorem for $p=3$. A more general proof follows the same lines.

If $(\alpha_i, \beta_i, \gamma_i)$, $i=1, 2, \ldots, s$, are s sets of three non-negative integers such that (i) for no i

https://doi.org/10.1017/S2040618500032998 Published online by Cambridge University Press

is the set $(\alpha_i, \beta_i, \gamma_i)$ (0, 0, 0), (ii) the non-zero α's form a partition of an integer u, the non-zero β's a partition of an integer v and the non-zero γ's a partition of an integer w, (iii) $\alpha_1 \leqslant \alpha_2 \leqslant \ldots \leqslant \alpha_s$, β's corresponding to equal α's are arranged in ascending order of magnitude and γ's corresponding to equal (α, β)'s are arranged in ascending order of magnitude, then $(\alpha_i, \beta_i, \gamma_i)$, $i = 1, 2, \ldots, s$, will be called a joint partition of u, v and w of order s and written $(\alpha, \beta, \gamma)_s$.

We now consider $m_{uvw} = E\{r_{12n}^u r_{23n}^v r_{31n}^w\}$. This can be expressed as a sum of terms of the form

$$C[(\alpha, \beta, \gamma)_s] n^{[s]} n^{\frac{1}{2}(u+v+w)} \prod_{i=1}^{3} \frac{1}{n^{[h_i]}} S_n(a_i', \epsilon_i),$$

where

(i) $(\alpha, \beta, \gamma)_s$ is a joint partition of u, v, w, of order s,

(ii) $C[(\alpha, \beta, \gamma)_s]$ is a constant independent of n,

(iii) $\epsilon_1 = (\epsilon_{11}, \epsilon_{12}, \ldots, \epsilon_{1h_1})$ is a partition of $(u+w)$ formed by the h_1 non-zero sums $\epsilon_{1i} = (\alpha_i + \gamma_i)$ of the joint partition $(\alpha, \beta, \gamma)_s$ and similarly for ϵ_2 and ϵ_3,

(iv) $S_n(a_i', \epsilon_i)$ denotes the symmetric polynomial in the $a_{in\tau}'$'s introduced in the paper (**5**),

(v) the number of terms is independent of n.

We will require the following inequalities :

(*a*) If every $\epsilon_{ij} \geqslant 2$ then $h_1 + h_2 + h_3 \leqslant u + v + w$, and if, in addition, some $\epsilon_{ij} > 2$ then $h_1 + h_2 + h_3 < u + v + w$.

(*b*) $h_1 + h_2 + h_3 \geqslant 2s$. For if there are p zero α's, q zero β's and r zero pairs (α, β), then we have

$$(s-q) + (s-p) \geqslant s - r,$$
$$h_1 = s - r,$$
$$h_2 \geqslant r + (s-q),$$
$$h_3 \geqslant r + s - p,$$

and the result follows.

(*c*) If every $\epsilon_{ij} = 2$ and at least one $\alpha = 1$, then $h_1 + h_2 + h_3 > 2s$. For if $\alpha_i = 1$, $i = p+1$, $p+2, \ldots, p+k$, then for the same range of values of i, $\beta_i = \gamma_i = 1$. For the remaining values of i in the range $1 \leqslant i \leqslant s$, $(\alpha_i, \beta_i, \gamma_i)$ is either (2, 0, 0), (0, 0, 2) or (0, 2, 0). Hence

$$h_1 = k + \tfrac{1}{2}(u-k) + \tfrac{1}{2}(w-k),$$

and similar equalities hold for h_2 and h_3 so that $h_1 + h_2 + h_3 = u + v + w$, while

$$s = k + \tfrac{1}{2}(u-k) + \tfrac{1}{2}(v-k) + \tfrac{1}{2}(w-k) = \frac{u+v+w}{2} - \frac{k}{2}.$$

Now $S_n(a_i', \epsilon_i)$ can be expressed as a sum of terms of the form $C(\epsilon_i, \epsilon_i') \prod_{j=1}^{h_i'} \left\{ \sum_{\tau=1}^{n} a'^{\epsilon_{ij}'}_{in\tau} \right\}$ where $\epsilon_i' = (\epsilon_{i1}', \epsilon_{i2}', \ldots, \epsilon_{ih_i}')$ is a partition in which each ϵ_{ij}' is either an ϵ_{ij} or a sum of ϵ_{ij}'s, $\epsilon_{ij}' \geqslant 2$, $j = 1, 2, \ldots, h_i'$, the number of terms is independent of n and $C(\epsilon_i, \epsilon_i')$ is a constant independent of n. Hence m_{uvw} can be expressed as a sum of terms of the form

$$C[(\alpha, \beta, \gamma)_s] n^{[s]} n^{\frac{u+v+w}{2}} \prod_{i=1}^{3} \left[\frac{1}{n^{[h_i]}} C(\epsilon_i, \epsilon_i) \prod_{j=1}^{h_i'} \left\{ \sum_{\tau=1}^{n} a'^{\epsilon_{ij}'}_{in\tau} \right\} \right].$$

https://doi.org/10.1017/S2040618500032998 Published online by Cambridge University Press

For $k \geqslant 2$, $|\sum_{\tau=1}^{n} a_{i\pi\tau}'^{k}| \leqslant [M_n(a)]^{k-2}$, where $M_n(a) = \max_{1 \leqslant i \leqslant 3} [M_n(a_i)]$. Hence since $h_1' + h_2' + h_3' \leqslant h_1 + h_2 + h_3$ the above term is $O[f(n)]$, where

$$f(n) = [n^{\frac{1}{4}} M_n(a)]^{2(u+v+w-h_1'-h_2'-h_3')} n^{s-\frac{1}{2}(h_1+h_2+h_3)}.$$

It follows from the inequalities (a), (b) and (c) and the fact that (a) applies to ϵ''s and h''s that $f(n) = o(1)$ unless every $(\alpha_i, \beta_i, \gamma_i)$ is one of the forms (2, 0, 0), (0, 2, 0) or (0, 0, 2). The remainder of the proof is similar to that of Theorem 3.6 in [**5**].

From this theorem it follows that the distribution of W is asymptotically normal with zero mean and unit variance provided $n^{\frac{1}{4}} M_n(a) \to 0$ as $n \to \infty$. This condition is satisfied in particular when $A_{in} = (1, 2, \ldots, n), i = 1, 2, \ldots, p$. Asymptotic normality of Kendall's coefficient of concordance [**3**] in the problem of m-rankings can easily be deduced from this.

REFERENCES

(**1**) Cramer, H., "Random variables and probability distributions," *Cambridge Tracts in Math.* **36**, (1937), p. 101.

(**2**) Hoeffding, W., "A combinatorial central limit theorem," *Ann. Math. Stats.*, **22**, 4 (1951).

(**3**) Kendall, M. G., *Advanced theory of statistics* (3rd ed., London, 1947), p. 411.

(**4**) Kruskal, W. H., "A non-parametric test for the several sample problem," *Ann. Math. Stats.*, **23**, 4 (1952.).

(**5**) Silvey, S. D., "The equivalence of asymptotic distributions under randomisation and normal theories," *Proc. Glasgow Math. Ass.*, **1**, 3 (1953).

(**6**) Wald, A., and Wolfowitz, J., "Statistical tests based on permutations of the observations," *Ann. Math. Stats.*, **15**, 4, (1944).

UNIVERSITY OF GLASGOW

https://doi.org/10.1017/S2040618500032998 Published online by Cambridge University Press