Multiple Runs

Author(s): D. E. Barton and F. N. David

Source: *Biometrika*, Jun., 1957, Vol. 44, No. 1/2 (Jun., 1957), pp. 168-178
Published by: Oxford University Press on behalf of Biometrika Trust
Stable URL: https://www.jstor.org/stable/2333249

MULTIPLE RUNS

By D. E. BARTON and F. N. DAVID
University College, London

1. The number of ways in which r_1 elements of one kind and r_2 of another can be arranged in a line to form a sequence of $2t$ or $2t+1$ groups was solved by Whitworth (1886, Problems 193 and 194), and it was probably not new to him. The solution was revived by Stevens (1939), Wald & Wolfowitz (1940) and Mood (1940). Sequences in which there are k different kinds of elements have also been treated. Whitworth takes three kinds of elements with the same number of each while Mood obtains a general solution for the distribution of the number of runs of one kind, given k types of element. Using a simple generalization of Whitworth's method we show here how the distributions of the total number of runs can be built up for the multiple case from that for two alternatives. The method of the characteristic random variable is used to obtain reasonably compact expressions for the moments of the total number of multiple runs. The assumption of normality for this total number is not discussed by us at length, since it obviously follows from the work of Wald & Wolfowitz and of Mood. We show, however, an alternative limit which is Poisson.

2. Starting with the two alternatives case it is assumed that there are r_1 white balls and r_2 red balls. If they are arranged in a line randomly, then T, the total number of groups (or runs) will be made up of t white groups $+\,t$ red groups, or $(t+1)$ white and t red, or t white and $(t+1)$ red. The number in the fundamental probability set is

$$^{r}C_{r_1},$$

where
$$r = r_1 + r_2.$$

The probability distribution of T is

$$P\{T = 2t\} = 2\,{}^{r_1-1}C_{t-1}\,{}^{r_2-1}C_{t-1}/{}^{r}C_{r_1}$$

and
$$P\{T = 2t+1\} = P\{T = 2t\}\left(\frac{r-2t}{2t}\right).$$

Now let there be r_1 white, r_2 red and r_3 black balls, and suppose it is required to find the probability distribution of
$$T = t_1 + t_2 + t_3,$$

where t_1 is the number of white groups, t_2 the number of red and t_3 the number of black groups, respectively. The number of ways in which r_i $(i = 1, 2, 3)$ can be split into t_i groups is

$$^{r_i-1}C_{t_i-1},$$

so the total number of ways is
$$\prod_{i=1}^{3} {}^{r_i-1}C_{t_i-1}.$$

We now consider the t_i as units and look for the number of ways in which they can be arranged along a line so that no two like colours are together. Following Whitworth we take the t_2 and t_3 groups and put them down in any order. Suppose that there is a total of x contacts of the same colour (RR or BB). There will be t_2+t_3-1 total contacts so that the number of RB or BR contacts will be t_2+t_3-1-x. Now take up the t_1 white groups. x of them will have to go between the RR and BB contacts leaving $t_1 - x$ to be placed between the BR and RB contacts or at either end of the line. This can be done in

$$^{t_2+t_3+1-x}C_{t_1-x}$$

ways. The number of ways in which t_2 and t_3 can be arranged to produce t_2+t_3-1-x contacts is the number of ways in which t_2 and t_3 can be arranged to produce $t_2+t_3-x = G$ (say) groups. The number of ways for this, depending on whether G is even or odd, is quoted above. The total number of ways of producing

$$T = t_1+t_2+t_3$$

will be then
$$\sum \prod_{i=1}^{3} {}^{r_i-1}C_{t_i-1} \sum_{x} {}^{t_2+t_3+1-x}C_{t_1-x}[P\{G = t_2+t_3-x\}\,{}^{t_2+t_3}C_{t_2}],$$

where the inner sum is taken over all possible values of x and the outer sum is over all possible 3-partitions of T. There are restrictions on both x and T in order that the combinatorial coefficient shall have sense, but these are obvious in any calculation. The number in the fundamental probability set is

$$\frac{r!}{\prod_{i=1}^{3} r_i!},$$

where r is the sum of the r_i. The probability of obtaining a given number of runs is just the ratio of the two expressions.

3. The distributions for four and more runs can be built up successively from the distribution for three. Let there be $r_i\,(i = 1, 2, \ldots, k)$ with r the sum of the r_i. Let there be $t_i\,(i = 1, 2, \ldots, k)$ groups respectively and suppose that

$$p(T) = p\left(\sum_{i=1}^{k} t_i\right)$$

is required. The number of ways of forming T groups is

$$\prod_{i=1}^{k} {}^{r_i-1}C_{t_i-1}.$$

Arrange
$$T' = \sum_{i=1}^{k-1} t_i = T - t_k$$

in any random order and let there be x total contacts of self-colours, i.e. x is the total number of white-white, red-red, black-black, ..., contacts. Since there are $T'-1$ contacts in all there will be $T'-1-x$ contacts of different colours. Take the t_k groups which so far have not been arranged. Put x of these in the x self-colour contacts and arrange the remaining $t_k - x$ in the $T'-1-x$ different-colour contacts, or at the ends of the line. The number of ways of doing this will be
$$^{T'+1-x}C_{t_k-x}.$$

The number of ways in which T' elements can be arranged to form $T'-x$ groups is obtained from the distribution of runs of $k-1$ colours and we have

$$P\{T = T_0\} = \frac{\prod_{i=1}^{k} r_i!}{r!} \sum_{T_0} \prod_{i=1}^{k} {}^{r_i-1}C_{t_i-1} \sum_{x} {}^{T'+1-x}C_{t_k-x}\left[P\left\{T' = \sum_{1}^{k-1} t_i - x\right\} \frac{\left(\sum_{j=1}^{k-1} t_j\right)!}{\prod_{j=1}^{k-1} t_j!}\right],$$

where the first sum is over all possible k-partitions to T_0 and the second sum over all possible x. The restrictions on both x and the partitions of T_0 become obvious in calculation and the distributions can be calculated surprisingly quickly. Using the formulae of this and of the preceding section the distributions for $r = 3, \ldots, 12$ and $k = 3, 4$ were found and are given in Tables 1 and 2. More extensive calculations did not appear profitable in the light of approximations discussed later.

4. The moments of T, the total number of runs of all colours, follow directly from first principles. Define a characteristic random variable α_t which has the property that it equals unity when the balls on either side of the tth gap are the same colour and zero when they are of different colours. Define S by

$$S = r - T = \sum_{t=1}^{r-1} \alpha_t.$$

The moments of S can now, in principle, be written down. Thus

$$\mathscr{E}(\alpha_t) = \sum_{i=1}^{k} \frac{r_i(r_i-1)}{r(r-1)}$$

and
$$\mathscr{E}(S) = \frac{1}{r} \sum_{i=1}^{k} r_i(r_i-1).$$

We shall write generally
$$F_\omega = \sum_{i=1}^{k} r_i^{(\omega)},$$

so that in this notation
$$r\,.\,\mathscr{E}(S) = F_2.$$

Again
$$\mathscr{E}[S(S-1)] = \mathscr{E}\left[\left(\sum_{t=1}^{r-1} \alpha_t\right)^2 - \sum_{t=1}^{r-1} \alpha_t\right].$$

Because α_t is a characteristic random variable

$$\mathscr{E}(\alpha_t^m) = \mathscr{E}(\alpha_t),$$

but it will be necessary to divide the double sum into two parts:

$$\mathscr{E} \sum_{\substack{t=1 \\ t\neq\omega}}^{r-1} \sum_{\omega=1}^{r-1} \alpha_t \alpha_\omega = \mathscr{E} \sum_{t=1}^{r-1} \alpha_t \alpha_{t+1} + \mathscr{E} \sum_{\substack{t=1 \\ |t-\omega| \geqslant 2}}^{r-1} \sum_{\omega=1}^{r-1} \alpha_t \alpha_\omega.$$

The expected values of the products are

$$\mathscr{E}(\alpha_t \alpha_{t+1}) = \sum_{i=1}^{k} \frac{r_i(r_i-1)(r_i-2)}{r(r-1)(r-2)}$$

and
$$\mathscr{E}(\alpha_t \alpha_\omega) = \sum_{i \neq i} \frac{r_i(r_i-1)\,r_j(r_j-1)}{r(r-1)(r-2)(r-3)} + \sum_{i=1}^{k} \frac{r_i(r_i-1)(r_i-2)(r_i-3)}{r(r-1)(r-2)(r-3)},$$

there being $2(r-2)$ products in which the suffices of the α's differ by unity and $(r-2)(r-3)$ products in which the suffices differ by two or more. From David & Kendall's tables of symmetric functions the double sums in the expectations can be eliminated and we have finally

$$r^{(2)} \mu_{(2)}(S) = F_2(F_2-2) - 2F_3.$$

By a similar process we obtain the third and fourth factorial moments of S which are

$$r^{(3)} \mu_{(3)}(S) = F_2(F_2-2)(F_2-4) - 6F_3(F_2-3) + 10F_4 + 10F_3,$$

$$\begin{aligned} r^{(4)} \mu_{(4)}(S) = {} & F_2(F_2-2)(F_2-4)(F_2-6) - 12F_3(F_2-3)(F_2-5) + 40F_4(F_2-4) \\ & + 40F_3(F_2-3) + 12F_3(F_3-3) - 84F_5 - 296F_4 - 120F_3. \end{aligned}$$

5. Two approaches to the limit are possible. First fix k, the number of colours, and let r, the total number of balls, increase without limit. Then we have that

$$\frac{S - F_2/r}{\sigma_S}$$

is in the limit a unit normal variable where

$$\sigma_S^2 = \frac{F_2(r-3)}{r(r-1)} + \frac{F_2^2}{r^2(r-1)} - \frac{2F_3}{r(r-1)}.$$

This result follows by generalizing Mood's procedure for runs of one colour. As an alternative we can fix the possible numbers of each colour, i.e. we can put an upper bound on r_i, and let k and therefore r tend to infinity. Under these conditions we have approximately

$$\mu_{(w)}(S) = \mathscr{E}\Big(\sum_{i_1 \neq i_2 \neq \ldots \neq i_w} \alpha_{i_1}\alpha_{i_2}\ldots\alpha_{i_w}\Big) \doteqdot \sum \frac{r_{i_1}^{(i_1+1)} \ldots r_{i_w}^{(i_w+1)}}{r^{\left(\sum_{j=1}^{w} i_j + w\right)}},$$

since the other sums in which some or all of the subscripts are equal will be of lower order in r. Accordingly if we put

$$\lambda = F_2/r,$$

then

$$\mu_{(w)}(S) \rightarrow \lambda^w,$$

which is the wth Poisson factorial moment. This indicates that S tends in the limit to be distributed as a Poisson variable with parameter λ. A rigorous proof of this can be given following the lines already set out elsewhere for moments of a similar structure (Barton, 1957). We have therefore that as k (and r) increase without limit the distribution of the total number of runs tends to that of Poisson's binomial limit with parameter F_2/r. In practice this limit will not be reached until k is large.

6. Following Aitken (1939) we define the factorial cumulants $\kappa_{(t)}$ by the relation

$$\kappa_t = \sum_{i=1}^{t} \frac{\Delta^i (0)^t}{i!} \kappa_{(i)},$$

where κ_t is the ordinary moment cumulant. The first four of these are

$$\kappa_1 = \kappa_{(1)}, \quad \kappa_2 = \kappa_{(2)} + \kappa_{(1)}, \quad \kappa_3 = \kappa_{(3)} + 3\kappa_{(2)} + \kappa_{(1)}, \quad \kappa_4 = \kappa_{(4)} + 6\kappa_{(3)} + 7\kappa_{(2)} + \kappa_{(1)}.$$

From these relations we may find the factorial cumulants of S:

$$\kappa_1(S) = \frac{F_2}{r}, \quad \kappa_{(2)} = \frac{F_2^2 - 2rF_2 - 2rF_3}{r^2(r-1)},$$

$$\kappa_{(3)}(S) = \frac{1}{r^3(r-1)(r-2)}[4F_2^3 - 12rF_2^2 + 8r^2F_2 - 12rF_3F_2 + 28r^2F_3 + 10r^2F_4].$$

The expression for $\kappa_{(4)}(S)$ is lengthy and we do not reproduce it here. Write

$$\lambda = F_2/r, \quad \lambda_t = F_{2+t}/r,$$

whence

$$\kappa_1(S) = \lambda, \quad \kappa_{(2)}(S) = \frac{\lambda^2 - 2\lambda - 2\lambda_1}{r-1}, \quad \kappa_{(3)}(S) = \frac{1}{(r-1)(r-2)}[4\lambda^{(3)} - 6\lambda_1(2\lambda - 3) + 10(\lambda_1 + \lambda_2)].$$

These cumulants simplify considerably if we let

$$r_i = R \quad (i = 1, 2, \ldots, k),$$

for in this case

$$\lambda_t = \lambda^{(t+1)}.$$

Thus for equal numbers of each colour we have

$$\kappa_1 = \lambda, \quad \kappa_{(2)} = -\frac{\lambda^2}{r-1}, \quad \kappa_{(3)} = \frac{2\lambda^2(\lambda-1)}{(r-1)(r-2)}, \quad \kappa_{(4)} = -\frac{6\lambda^2}{(r-1)(r-2)(r-3)}\left[\lambda^2 - 4\lambda + 2 + \frac{\lambda^2}{r-1}\right].$$

If we consider a simple binomial expression, say

$$(q+p)^r,$$

where

$$\lambda = rp,$$

the factorial moments are

$$\kappa_1 = \lambda, \quad \kappa_{(2)} = -\frac{\lambda^2}{r}, \quad \kappa_{(3)} = \frac{2\lambda^3}{r^2}, \quad \kappa_{(4)} = -\frac{6\lambda^4}{r^3}.$$

It will be noted that the leading term in the factorial cumulant of S is the same as the factorial cumulants of the binomial in the four cases given.

7. The similarity between the first four cumulants of the distribution of S, when the numbers of each colour are the same, and those of the simple binomial, coupled with the fact that two possible distributions under conditions analogous to those of proceeding to binomial limits are the normal distribution and Poisson, suggests that a binomial distribution may be a useful approximation to the distribution of S. The case $r = 12$ with four (k) colours, each three (R) in number was considered. S can take values $0, 1, 2, \ldots, 8$, and

$$\kappa_1(S) = \frac{F_2}{r} = R-1, \quad \kappa_2(S) = \frac{F_2^2}{r^2(r-1)} + \frac{F_2(r-3)}{r(r-1)} - \frac{2F_3}{r(r-1)} = \frac{(R-1)(r-R)}{r-1}.$$

Three approximations to the distribution of S are now considered. First we let the binomial index equal 8, and fit the binomial

$$(q+p)^n = (\tfrac{3}{4} + \tfrac{1}{4})^8,$$

the 'p' of the binomial being found from

$$8p = \kappa_1 = 2.$$

This is approximation I of Table 3. Secondly, we equate the first two moments of S, i.e. $R-1$ and $(R-1)(r-R)/(r-1)$, to the first two binomial moments, which is equivalent to taking

$$p = \frac{R-1}{r-1} = \frac{2}{11}, \quad n = \frac{\kappa_1}{p} = r-1 = 11.$$

This is approximation II. Thirdly, we may calculate areas corresponding to a normal curve with the correct mean and variance of S. This is approximation III. The approximate and exact distributions are given in Table 3, the normal curve area in the first group representing the whole left-hand tail. Approximation II—the simple binomial with the correct first two moments of S—is clearly the best, but no mistake is likely in a test of significance whichever approximation is used.

7. Although it is possible to show similarity between the factorial binomial cumulants and those of the distribution of S in a simple way when there are equal numbers of each colour, the approximations to the distribution of S are so good for this particular case that one might expect them to be reasonable for the case when the numbers of each colour are not the same. We found that this was so. It is supposed there are 6 balls of one colour, 4 of another and 1 of each of two other colours. Thus if we consider approximation I and assume

$$n = 8, \quad \kappa_1 = \frac{1}{r}\sum_i r_i(r_i - 1) = 3{\cdot}5, \quad p = \tfrac{7}{16},$$

and the binomial is

$$(\tfrac{9}{16} + \tfrac{7}{16})^8.$$

Table 3. *Comparison of true distribution of S with three approximations* ($r = 12$. $[3^4]$)

S	0	1	2	3	4	5	6	7	8
Approximation I	0·1001	0·2670	0·3115−	0·2076	0·0865+	0·0231	0·0038	0·0004	0·0000
Approximation II	0·1100	0·2689	0·2987	0·1992	0·0885+	0·0275+	0·0061	0·0010	0·0001
Approximation III	0·1205	0·2274	0·3042	0·2274	0·0952	0·0222	0·0029	0·0002	—
True	0·1118	0·2670	0·2966	0·2003	0·0903	0·0275+	0·0057	0·0008	0·0001

Table 4. *Comparison of true distribution of S with three approximations* ($r = 12$. $[641^2]$)

S	0	1	2	3	4	5	6	7	8
Approximation I	0·0100	0·0624	0·1698	0·2641	0·2567	0·1598	0·0621	0·0138	0·0013
Approximation II	0·0083	0·0563	0·1654	0·2714	0·2697	0·1632	0·0564	0·0091	0·0002
Approximation III	0·0126	0·0552	0·1599	0·2723	0·2723	0·1599	0·0552	0·0112	0·0013
True	0·0054	0·0574	0·1688	0·2749	0·2673	0·1582	0·0567	0·0104	0·0009

For approximation II

$$\kappa_2 = \frac{(r-3)}{r(r-1)}\Sigma r_i(r_i-1) + \left(\frac{r-3}{r(r-1)}\right)^2 (\Sigma r_i(r_i-1))^2 - \frac{2}{r(r-1)}\Sigma r_i(r_i-1)(r_i-2)$$
$$= \tfrac{79}{44},$$

giving the binomial

$$(\tfrac{79}{154} + \tfrac{75}{154})^{539/75} = (0{\cdot}51299 + 0{\cdot}48701)^{7{\cdot}186}.$$

The fractional value of the index was used to calculate the probabilities. In the third approximation it was assumed that S is normally distributed with mean 3·5 and variance 79/44, the tail areas being put together in each of the two end groups. The agreement is surprisingly good considering that r is small and that there is disparity between the numbers of each colour. The results would suggest that for $r > 12$, whatever the composition of the numbers of colours, the normal approximation (III) will be adequate for tests of significance.

8. An extension of the theory of runs can be used in what we might call a test for the persistence of type. Assuming the r events of k possible types, the null hypothesis will be that these r events are a sample from a multinomial, the probability of the ith type of which is p_i. This is to say that under H_0 we assume

$$p(r_1, r_2, \ldots, r_k) = \frac{r!}{\prod_{i=1}^{k} r_i!} \prod_{i=1}^{k} p_i^{r_i},$$

with each of the

$$\frac{r!}{\prod_{i=1}^{k} r_i!}$$

sequences being equally likely. The alternate hypothesis, H_1, might be that we have a simple Markoff chain in which the probability of getting an event of the ith type at any given drawing is greater if the immediately preceding drawing also is of the ith type and less if otherwise. We may write these transition probabilities

$$p_{ii} = p_i^2(1+\theta),$$

$$p_{ij} = p_i p_j(1-\theta W) \quad (i \neq j),$$

where

$$W = \Sigma p_i^2/(1-\Sigma p_i^2).$$

Assuming $\theta = 0$ for the first drawing of the sample, the probability of any given sequence in which there are T groups is

$$(1+\theta)^{r-T-1}(1-\theta W)^T \prod_{i=1}^{k} p_i^{r_i}.$$

If we now consider the conditional distribution under H_1 for a sequence of given composition of numbers of each type, this is equal to

$$\left(\frac{1-\theta W}{1+\theta}\right)^T K,$$

where K is a factor of proportionality and depends only on $r_1, \ldots, r_k$ and

$$\frac{1-\theta W}{1+\theta} = \phi \quad \text{(say)}.$$

It follows that T is sufficient for ϕ and that the likelihood ratio test for $\theta = 0$, or equivalently $\phi = 0$, is a function of T. Thus the use of T is equivalent to the likelihood ratio test in this case.

9. We have been interested in the multiple runs problem for its own sake but, apart from the simple stochastic problem of the preceding section, it is perhaps worth while to describe one possible analysis of variance application. The use of runs in statistical theory is obvious, but there appears room for them in the category which Tukey has aptly described elsewhere as 'quick and dirty'. By quick and dirty tests we would understand tests which are easy to apply but probably of low power, so that any effect which is significant using them will certainly be more significant with a more sensitive test. On the other hand, if the looked-for effect is not significant with the 'quick and dirty' test this does not necessarily mean that it cannot be picked out by more refined methods. We illustrate the working on an example of

the breaking strength of cement-mortar briquettes (Table 5). We mark the observations in order of magnitude, ties being decided by tossing a coin. The ranking is simply in order to enable the runs to be counted more easily. Each group will correspond to a different colour. The number of runs is 22 and

$$S = 25-22 = 3, \quad \mathscr{E}(S) = 4, \quad \text{var}(S) = \tfrac{10}{3}.$$

Assuming S to be normally distributed we see that significance is not achieved, a result which, in this case, is confirmed by the orthodox F-test.

Table 5. *Breaking strength of cement-mortar briquettes*

Group	1	2	3	4	5
Tension in lb.	518 (5) 560 (20) 538 (11) 510 (4) 544 (6)	508 (3) 574 (23) 528 (6) 534 (8) 538 (13)	538 (12) 544 (15) 554 (18) 579 (24) 598 (25)	535 (9) 540 (14) 550 (17) 555 (19) 567 (21)	492 (1) 506 (2) 528 (7) 536 (10) 572 (22)

10. Although the analysis of variance runs test has little power this will not be the case for all applications. We owe the following elegant illustration to E. S. Pearson. Consider the falls in the price of shares on the London Stock Exchange during the period 6 November to 8 December 1956, both dates inclusive. Five types of industrial activity, A Insurance, B Breweries and Distilleries, C Electrical Equipment and Radio, D Motor and Aircraft, E Oil were chosen for study. The closing prices as given in *The Times* for eighteen businesses of each type were taken, and Table 6 shows, for each day, the type of industrial activity for which the greatest number of the eighteen showed a fall in price from the previous day. In the few cases where there were equal numbers for two types, that type which also showed the fewer rises in price was taken:

Table 6. *Type of industrial activity showing greatest number of falls in price of shares*

6 Nov.	A	13 Nov.	B	20 Nov.	E	27 Nov.	B	4 Dec.	C
7 Nov.	A	14 Nov.	C	21 Nov.	C	28 Nov.	E	5 Dec.	C
8 Nov.	D	15 Nov.	C	22 Nov.	E	29 Nov.	A	6 Dec.	D
9 Nov.	D	16 Nov.	C	23 Nov.	E	30 Nov.	E	7 Dec.	C
10 Nov.	A	17 Nov.	E	24 Nov.	E	1 Dec.	E	8 Dec.	B

We have $r = 25$ with $r_A = 4$, $r_B = r_D = 3$, $r_C = 7$, $r_E = 8$, $T = 16$ and $S = r-T = 9$. From formulae we find

$$F_2 = \sum_{i=1}^{5} r_i(r_i-1) = 122, \quad F_3 = \sum_{i=1}^{5} r_i(r_i-1)(r_i-2) = 582,$$

$$\mu_2(S) = \frac{F_2(r-3)}{r(r-1)} + \frac{F_2^2}{r^2(r-1)} - \frac{2F_3}{r(r-1)} = 1{\cdot}54106,$$

$$\sigma(S) = 1{\cdot}241 \quad \text{and} \quad \mathscr{E}(S) = 4{\cdot}88,$$

so that

$$\frac{S-\mathscr{E}(S)}{\sigma_S} = \frac{4{\cdot}12}{1{\cdot}241} > 3.$$

This normal deviate is significant, and we conclude therefore that the test is picking out the fact that during the Middle East crisis there was some persistence from day to day in the way in which different classes of shares were affected.

11. In nearly all statistical applications of the theory of runs the test of significance will be one-tailed. For example, in Wald & Wolfowitz's application of the two-colour distribution to the two-sample means test, a small number of runs could be held to indicate a possible separation of the population means. The critical region in their case will be the tail where T is small or S large. This is also true for the applications given above; for instance, in the analysis of variance application a complete separation of the twenty-five observations into five groups would indicate the possibility of the five population means being different instead of equal under the null hypothesis. We choose therefore the lower tail of T or the upper tail of S. The result when there is too great an alternation of colour and T is significant at the upper tail has not so obvious a statistical interpretation.

REFERENCES

Aitken, A. C. (1939). *Statistical Mathematics*. Edinburgh: Oliver & Boyd.
Barton, D. E. (1957). *J. R. Statist. Soc.* B (to appear).
Mood, A. M. (1940). *Ann. Math. Statist.* **11**, 367.
Stevens, W. L. (1939). *Ann. Eugen., Lond.*, **9**, 10.
Wald, A. & Wolfowitz, J. (1940). *Ann. Math. Statist.* **11**, 112.
Whitworth, W. A. (1886). *Choice and Chance*. Cambridge: Deighton Bell and Co.

Table 1. *Three-colour runs* ($S = r - T$)

(The probabilities are obtained by dividing the number tabled by the corresponding value of $\frac{r!}{\Pi r_i!}$)

r	r_1	r_2	r_3	$\frac{r!}{\Pi r_i!}$	Values of T									
					3	4	5	6	7	8	9	10	11	12
6	2	2	2	90	6	18	36	30						
	3	2	1	60	6	18	26	10						
	4	1	1	30	6	18	6	—						
7	3	2	2	210	6	24	62	80	38					
	3	3	1	140	6	24	52	40	18					
	4	2	1	105	6	24	42	30	3					
	5	1	1	42	6	24	12	—	—					
8	3	3	2	560	6	30	100	180	170	74				
	4	2	2	420	6	30	90	150	120	24				
	4	3	1	280	6	30	80	90	60	14				
	5	2	1	168	6	30	60	60	12	—				
	6	1	1	56	6	30	20	—	—	—				
9	3	3	3	1680	6	36	150	360	510	444	174			
	4	3	2	1260	6	36	140	310	405	284	79			
	5	2	2	756	6	36	120	240	252	96	6			
	4	4	1	630	6	36	120	180	180	84	24			
	5	3	1	504	6	36	110	160	132	56	4			
	6	2	1	252	6	36	80	100	30	—	—			
	7	1	1	72	6	36	30	—	—	—	—			
10	4	3	3	4200	6	42	202	580	1050	1234	838	248		
	4	4	2	3150	6	42	192	510	870	894	498	138		
	5	3	2	2520	6	42	182	470	752	692	332	44		
	5	4	1	1260	6	42	162	300	372	252	108	18		
	6	2	2	1260	6	42	152	350	440	240	30	—		
	6	3	1	840	6	42	142	250	240	140	20	—		
	7	2	1	360	6	42	102	150	60	—	—	—		
	8	1	1	90	6	42	42	—	—	—	—	—		
11	4	4	3	11550	6	48	266	900	2010	3064	3012	1764	480	
	5	3	3	9240	6	48	256	840	1802	2568	2340	1168	212	
	5	4	2	6930	6	48	246	750	1527	1968	1548	702	135	
	6	3	2	4620	6	48	226	660	1220	1360	870	220	10	
	5	5	1	2772	6	48	216	480	744	672	432	144	30	
	6	4	1	2310	6	48	206	450	645	560	300	90	5	
	7	2	2	1980	6	48	186	480	690	480	90	—	—	
	7	3	1	1320	6	48	176	360	390	280	60	—	—	
	8	2	1	495	6	48	126	210	105	—	—	—	—	
	9	1	1	110	6	48	56	—	—	—	—	—	—	
12	4	4	4	34650	6	54	342	1350	3618	6894	9036	7938	4320	1092
	5	4	3	27720	6	54	332	1270	3300	5974	7388	5982	2826	588
	6	3	3	18480	6	54	312	1140	2778	4570	5060	3360	1100	100
	5	5	2	16632	6	54	312	1080	2592	4104	4272	2880	1110	222
	6	4	2	13860	6	54	302	1030	2388	3620	3550	2130	710	70
	7	3	2	7920	6	54	272	880	1818	2350	1820	660	60	—
	6	5	1	5544	6	54	272	700	1320	1400	1120	540	110	22
	7	4	1	3960	6	54	252	630	1008	1050	660	270	30	—
	8	2	2	2970	6	54	222	630	1008	840	210	—	—	—
	8	3	1	1980	6	54	212	490	588	490	140	—	—	—
	9	2	1	660	6	54	152	280	168	—	—	—	—	—
	10	1	1	132	6	54	72	—	—	—	—	—	—	—

Table 2. *Four-colour runs* ($S = r - T$)

(The probabilities are obtained by dividing the number tabled by the corresponding value of $\frac{r!}{\Pi r_i!}$)

r	r_1	r_2	r_3	r_4	$\frac{r!}{\Pi r_i!}$	Values of T								
						4	5	6	7	8	9	10	11	12
6	2	2	1	1	180	24	72	84						
	3	1	1	1	120	24	72	24						
7	2	2	2	1	630	24	108	252	246					
	3	2	1	1	420	24	108	192	96					
	4	1	1	1	210	24	108	72	6					
8	2	2	2	2	2520	24	144	504	984	864				
	3	2	2	1	1680	24	144	444	684	384				
	3	3	1	1	1120	24	144	384	384	184				
	4	2	1	1	840	24	144	324	294	54				
9	3	2	2	2	7560	24	180	780	2010	2880	1686			
	3	3	2	1	5040	24	180	720	1560	1720	836			
	4	2	2	1	3780	24	180	660	1320	1260	336			
	4	3	1	1	2520	24	180	600	870	660	186			
	5	2	1	1	1512	24	180	480	600	216	12			
10	3	3	2	2	25200	24	216	1140	3720	7480	8416	4204		
	3	3	3	1	16800	24	216	1080	3120	5160	5016	2184		
	4	2	2	2	18900	24	216	1080	3330	6210	6066	1974		
	4	3	2	1	12600	24	216	1020	2730	4170	3366	1074		
	5	2	2	1	7560	24	216	900	2160	2736	1368	156		
	4	4	1	1	6300	24	216	900	1740	1980	1116	324		
	5	3	1	1	5040	24	216	840	1560	1536	768	96		
11	3	3	3	2	92400	24	252	1584	6360	16680	27756	27408	12336	
	4	3	2	2	69300	24	252	1524	5820	14400	22056	18708	6516	
	4	3	3	1	46200	24	252	1464	5070	10720	14256	10848	3566	
	5	2	2	2	41580	24	252	1404	4950	11016	14184	8364	1386	
	4	4	2	1	34650	24	252	1404	4350	9000	10656	6768	2016	
	5	3	2	1	27720	24	252	1344	4200	7896	8484	4704	816	
	5	4	1	1	13860	24	252	1224	2910	4176	3384	1584	306	
	6	2	2	1	13860	24	252	1164	3210	4920	3480	780	30	
	6	3	1	1	9240	24	252	1104	2460	2920	1980	480	20	
12	3	3	3	3	369600	24	288	2112	10176	33360	74016	109632	98688	41304
	4	3	3	2	277200	24	288	2052	9486	29590	61916	83952	66638	23254
	4	4	2	2	207900	24	288	1992	8796	26100	51216	62892	43128	13464
	5	3	2	2	166320	24	288	1932	8316	23856	44520	50484	30468	6432
	4	4	3	1	138600	24	288	1932	7896	20580	35616	39312	25428	7524
	5	3	3	1	110880	24	288	1872	7416	18616	30320	31104	17528	3712
	5	4	2	1	83160	24	288	1812	6726	15966	23820	21384	10818	2322
	6	2	2	2	83160	24	288	1752	6876	17460	27120	22080	7020	540
	6	3	2	1	55440	24	288	1692	5976	13060	17120	12780	4160	340
	5	5	1	1	33264	24	288	1632	4656	8352	9024	6336	2448	504
	6	4	1	1	27720	24	288	1572	4386	7410	7620	4680	1590	150