Tables for Three-Sample Test

By Yoshihiko HIRAGA, Hidenori MORIMURA and Hisao WATANABE

(Received July 25, 1953)

The total number of runs for elements of two kinds was proposed as the criterion for a two-sample test by Wald and Wolfowitz [4], and was also used by Stevens [2] for testing independence or randomness in ordered sequences. Detailed tables have been provided for use in these tests by Swed and Eisenhart [3]. The distribution theory of runs being developed by Mood [1], these tests can be readily extended to the case of three or more kinds of elements, as pointed out by Wilks [6]. Whitney [5] has recently proposed a certain kind of three-sample test and, as an illustration, has given tables available only for a particular case.

Using the total number of runs of three kinds as a test criterion, the following tables have been prepared for testing the hypothesis that three samples are from the same population, giving 5% critical values of the criterion adopted. Consider random arrangements of n elements of three kinds, for example, n_1 elements of one kind, n_2 elements of a second kind and n_3 elements of a third kind with $n_1+n_2+n_3=n$, which may be regarded as three samples of size n_1, n_2, n_3, and let r_i denote the number of runs of elements of the i-th kind, and put $r=r_1+r_2+r_3$, that is, the total number of runs. According to Mood [1], the tabulation of the exact distribution of r reduces to that of the function which represents the number of different arrangements of r_1 objects of one kind, r_2 objects of a second kind and r_3 objects of a third kind so that no two adjacent objects are of the same kind, and in [1] this function is expressed as $F(r_1, r_2, r_3)$. Table 1 gives the largest integer, r_0, for which $\Pr(r \leq r_0) \leq 0.05$ for $1 \leq n_1 \leq n_2 \leq n_3 \leq 10$. This table was obtained from the exact distribution of r. Table 2 gives values of r_0 for particular values of $n_1 \leq n_2 \leq n_3$ from 10 to 30 inclusive, and Tables 3 gives values of r_0 for all values of $n_1=n_2=n_3$ from 11 to 30 together with 1% critical values of r. These two tables were obtained by using the normal approximation given in Mood's paper together with a correction for mean and variance using exact values computed from formulas (4.6) and (4.8) in [1]. It seems that this corrected approximation is quite satisfactory for $n_1=n_2=n_3 \geq 10$. Since detailed tabulation for many different values of the level of significance appears unnecessary, tables have been constructed only for the 5% level except in Table 3.

Many thanks are due to Miss Mieko Iwazuru and Miss Eiko Ozaki for most of the computations, and most of the checking of these tables was done by H. Morimura.

Table 1

5 % critical values of r for $1 \leq n_1 \leq n_2 \leq n_3 \leq 10$

(The largest integer, r_0, for which $\Pr(r \leq r_0) \leq 0.05$)

$n_1=1$

n_3 \ n_2	1	2	3	4	5	6	7	8	9	10
1										
2										
3			3							
4			3	3						
5		3	3	4	4					
6		3	3	4	4	5				
7		3	4	4	5	5	6			
8		3	4	4	5	5	6	6		
9		3	4	5	5	6	6	7	7	
10	3	3	4	5	5	6	7	7	8	8

$n_1=2$

n_3 \ n_2	2	3	4	5	6	7	8	9	10
2									
3	3	3							
4	3	4	4						
5	4	4	5	5					
6	4	4	5	5	6				
7	4	5	5	6	6	7			
8	4	5	6	6	7	7	8		
9	4	5	6	6	7	7	8	8	
10	4	5	6	7	7	8	8	9	9

Table 1 (continued)

$n_1=3$

n_3 \ n_2	3	4	5	6	7	8	9	10
3	4							
4	4	5						
5	5	5	6					
6	5	6	6	7				
7	5	6	7	7	8			
8	6	6	7	8	8	9		
9	6	7	7	8	9	9	9	
10	6	7	8	8	9	9	10	10

$n_1=4$

n_3 \ n_2	4	5	6	7	8	9	10
4	6						
5	6	7					
6	6	7	8				
7	7	7	8	9			
8	7	8	8	9	9		
9	8	8	9	9	10	10	
10	8	9	9	10	10	11	11

$n_1=5$

n_3 \ n_2	5	6	7	8	9	10
5	7					
6	8	8				
7	8	9	9			
8	9	9	10	10		
9	9	10	10	11	11	
10	9	10	11	11	12	12

Table 1 (continued)

$n_1=6$

n_3 \ n_2	6	7	8	9	10
6	9				
7	9	10			
8	10	10	11		
9	10	11	12	12	
10	11	11	12	13	13

$n_1=8$

n_3 \ n_2	8	9	10
8	12		
9	13	13	
10	13	14	15

$n_1=9$

n_3 \ n_2	9	10
9	14	
10	15	15

$n_1=7$

n_3 \ n_2	7	8	9	10
7	11			
8	11	12		
9	12	12	13	
10	12	13	13	14

$n_1=10$

n_3 \ n_2	10
10	16

Table 2

5 % critical values of r *for* $10 \leq n_1 \leq n_2 \leq n_3 \leq 30$

$n_1=10$

n_3 \ n_2	10	12	14	16	18	20	25	30
12	18	19						
14	19	20	21					
16	19	21	22	23				
18	20	22	23	24	25			
20	21	22	24	25	26	27		
25	23	24	26	27	28	30	32	
30	24	26	27	29	30	31	34	37

Table 2 (continued)

$n_1=12$

n_3 \ n_2	12	14	16	18	20	25	30
14	21	22					
16	22	23	25				
18	23	24	26	27			
20	24	25	27	28	29		
25	26	27	29	30	31	34	
30	27	29	31	32	33	36	39

$n_1=14$

n_3 \ n_2	14	16	18	20	25	30
16	25	26				
18	26	27	28			
20	27	28	29	30		
25	29	30	32	33	36	
30	31	32	34	35	38	41

$n_1=16$

n_3 \ n_2	16	18	20	25	30
18	28	30			
20	30	31	32		
25	32	33	35	37	
30	34	35	37	40	43

$n_1=20$

n_3 \ n_2	20	25	30
25	37	41	
30	40	43	46

$n_1=18$

n_3 \ n_2	18	20	25	30
20	32	33		
25	35	36	39	
30	37	38	42	45

$n_1=25$

n_3 \ n_2	25	30
30	47	50

Table 3

1 % and 5 % critical values of r for $11 \leq n_1 = n_2 = n_3 \leq 30$

n_i	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1 %	16	18	20	22	23	25	27	29	30	32	34	36	38	39	41	43	45	47	48	50
5 %	18	20	22	24	25	27	29	31	33	35	36	38	40	42	44	46	48	49	51	53

INSTITUTE OF STATITICAL MATHEMATICS

REFERENCES

[1] A.M. Mood, The distribution theory of runs, *Ann. Math. Stat.*, **11** (1940), pp. 367–392.
[2] W.L. Stevens, Distribution of groups in a sequence of alternatives, *Ann. of Eugenics*, **9** (1939), pp. 10–17.
[3] F.S. Swed and C. Eisenhart, Tables for testing randomness of grouping in a sequence of alternatives, *Ann. Math. Stat.*, **14** (1943), pp. 66–87.
[4] A. Wald and J. Wolfowitz, On a test whether two samples are from the same population, *Ann. Math. Stat.*, **11** (1940), pp. 147–162.
[5] D.R. Whitney, A bivariate extension of the U statistic, *Ann. Math. Stat.*, **22** (1951), pp. 274–282.
[6] S.S. Wilks, Order statistics, *Bull. Amer. Math. Soc.*, **54** (1948), pp. 6–50.