ON THE APPLICATIONS OF THE STATISTICS W'S AND T'S FOR TESTING TWO SAMPLES

By B. N. Singh

Indian Standards Institution, New Delhi

1. Introduction

1.1. Iyer and Singh (1955) have recently considered a number of new *statistics* W_r, W_r', T_r and T_r' for a sequence of n observations taken from a continuous or discontinuous population. The *statistic* W_r is obtained by treating the sequence as $(n-r+1)$ blocks (or sub-sequences) each consisting of r consecutive observations and considering the differences (taken in the same sense) between all possible pairs of observations in each block. The score $+1$ or -1 is assigned to each of the differences according as it is positive or negative and W_r is the total positive (or negative) score for the $(n-r+1)$ blocks. The *statistic* W_r' represents the total score for all the blocks as obtained by assigning the score $+1$ to each of the differences whether positive or negative. The *statistic* T_r represents the number of positive (or negative) scores between pairs of observations separated by $(r-2)$ observations or less, subject to the condition that each difference is taken once only. If both positive and negative scores are considered, we get the *statistic* T_r'. It may be noted that T_r and T_r' are based on simply the sum of the scores for the differences between adjacent, alternate and other observations separated by 2, 3, $\cdots$ $(r-2)$ observations, each difference occurring once only. In W_r and W_r', the overlapping of the blocks results in the differences being repeated more than once. When $r=2$, W_r and T_r reduce to Moore-Wallis difference-sign test (1943). Further, when $r=n$ and all the observations in the sequence are different, Kendall's τ (1945) is given in terms of W_n (or T_n) by $\tau = \{4\, W_n/n\,(n-1) - 1\}$.

1.2. In case the sequence consists of n elements of two kinds, A and B, the *statistics* W's and T's are obtained by considering (or assigning scores to) the joins of different kinds, AB and/or BA, that arise from the blocks or spacings of the corresponding sizes. For example, the *statistic* W_r will represent the total number of AB joins (or positive scores) arising from $(n-r+1)$ moving blocks of r consecutive elements. The other *statistics* are also similarly defined. For the sequences obtained by ordering two samples A and B by pooling them together, it may be noted that for $r = 2$, $W_2' + 1$ (or $T_2' + 1$) gives the total number of runs of A's and B's in the combined sequence. When $r = n$, W_n (or T_n) represents the Mann-Whitney U *statistic* (1947)

1.3. It has been shown that all these *statistics* are consistent and asymptotically normal. The efficiency of these *statistics* has been examined (1955) for large samples ($n = 100$ and 200) by calculating their powers for different hypotheses and alternatives by making use of normal approximations and also by finding the squares of their coefficients of variation. It has been found that, under certain conditions, some of these *statistics* (based on both the tails) are more powerful and asymptotically more efficient than the Mann-Whitney U *statistic*. Though the Mann-Whitney *statistic* [or the Wilcoxon's (1945) test] is almost as efficient as the t-test for normal translation alternatives, its power is expected to be less for the alternatives of $F \neq G$ (F and G being the cumulative distribution functions). For the normal distribution both the mean (μ) and standard deviation (σ) may vary together. In view of these findings, these *statistics* should prove to be more useful in testing two samples when no information is available regarding the form of the distributions.

1.4. The purpose of this paper is: (i) to make a comparative study of the powers of some of the *statistics* empirically by drawing random samples from a normal population, (ii) to provide tables giving (*a*) the exact probability distributions for small samples and (*b*) means and standard deviations for large samples, so as to enable and facilitate the use of these *statistics* in practice and (iii) to discuss their applications to varying situations which may involve the testing of (*a*) the randomness of a sequence of observations *or* (*b*) the homogeneity of two samples, large or small, *or* (*c*) the difference between two sets of observations being greater (or less) than a specified value.

2. Empirical Study of Comparative Powers

2.1. Fifty sets of random samples, every set consisting of two samples of five observations each, were drawn from a normal population with zero mean and unit standard deviation with the help of the tables given by Mahalanobis *et al* (1934). The two samples in each set were designated A and B respectively. New sets of samples were formed from each of these sets by increasing A's by 0·5 and 1 and keeping B's unchanged. Thus three types of situations were examined: (i) the samples A and B in each set were drawn from a normal population with zero mean and unit standard deviation, $N(0, 1)$, (ii) A was drawn from

N (0·5, 1) and B from N(0, 1) and (iii) A was drawn from N (1, 1) and B from N (0, 1). In each of the three situations, the statistics W_2, W_3, W_4, T_3, T_4 and U were employed, as indicated in Section 3 to test whether the two samples in a set were drawn from the same or different populations. The power of the statistics W's and T's for two levels of significance, $\alpha = 0 \cdot 05$ and $0 \cdot 1$, were determined by finding the values of the statistics for the above sets of samples and noting the number of sets for which the observed values lay in the critical region (w) defined by $\alpha = 0 \cdot 05$ and $0 \cdot 1$. Thus, if for any statistic, m denotes the number of sets falling in the critical region and M the total number of sets considered, the proportion m/M would give an estimate of the power of the statistic for the alternative under consideration. More specifically, this amounted to the empirical calculation of the powers of statistics for the alternative hypotheses that the difference (δ) between the means of the two samples in a set was 0, 0·5 and 1. It may be observed here that in the determination of the critical regions the values of the statistics were chosen by interpolation from their respective probability distributions so as to make the probabilities of those values correspond exactly to 0·05 and 0·1, the selected levels of significance. These values were then made use of in ascertaining the difference between the two samples in the various sets corresponding to $\delta = 0$, 0·5 and 1. It may be noted that in the first case ($\delta = 0$) the value of m/M would give an estimate of the selected level of significance.

2·1·1. The results obtained in the above manner for both the two-tailed and one-tailed statistics are presented in Tables I A and I B respectively.

2·1·2. It would be seen from these tables that the estimated levels of significance for the statistics are different from the nominal values, *i.e.*, 0·05 and 0·1 and this is due to sampling fluctuations. Hence in the comparison of the powers of the statistics, it would be desirable to make allowance for this difference which is positive in certain cases and negative in others. The power values after these adjustments indicate that the statistic T_3 is likely to be more powerful than U for both the tails, but U appears to be the most powerful of all when a single tail alone is used. Some of the other statistics like T_4 and W_3 (two-tailed) appear to be good competitors. Also, as expected, the powers of the statistics appear to be much more when only one tail is used than when both the tails are taken into account. However, because of the fact that this investigation is based on a moderate number of samples only, its findings should be regarded merely as an indication of the

TABLE I A

Empirical power for two-tailed statistics

Statistic \ δ	0		0·5		1	
α	0·05	0·1	0·05	0·1	0·05	0·1
W_2	0·07 (0·05)	0·12 (0·10)	0·11 (0·09)	0·17 (0·15)	0·24 (0·22)	0·35 (0·33)
W_3	0·04 (0·05)	0·09 (0·10)	0·17 (0·18)	0·25 (0·26)	0·40 (0·41)	0·52 (0·53)
W_4	0·07 (0·05)	0·12 (0·10)	0·15 (0·13)	0·20 (0·18)	0·40 (0·38)	0·51 (0·49)
T_3	0·04 (0·05)	0·07 (0·10)	0·20 (0·21)	0·31 (0·34)	0·41 (0·42)	0·54 (0·57)
T_4	0·07 (0·05)	0·11 (0·10)	0·17 (0·15)	0·31 (0·30)	0·39 (0·37)	0·53 (0·52)
U	0·07 (0·05)	0·10 (0·10)	0·20 (0·18)	0·31 (0·31)	0·40 (0·38)	0·56 (0·56)

TABLE I B

Empirical power for one-tailed statistics

Statistic \ δ	0		0·5		1	
α	0·05	0·1	0·05	0·1	0·05	0·1
W_2	0·08 (0·05)	0·14 (0·10)	0·16 (0·13)	0·27 (0·23)	0·35 (0·32)	0·59 (0·55)
W_3	0·08 (0·05)	0·16 (0·10)	0·22 (0·19)	0·31 (0·25)	0·52 (0·49)	0·60 (0·54)
W_4	0·08 (0·05)	0·16 (0·10)	0·20 (0·17)	0·31 (0·25)	0·51 (0·48)	0·59 (0·53)

TABLE I B—*Contd.*

Statistic \ δ / α	0		0·5		1	
	0·05	0·1	0·05	0·1	0·05	0·1
T_3	0·07 (0·05)	0·15 (0·10)	0·29 (0·27)	0·38 (0·33)	0·54 (0·52)	0·67 (0·62)
T_4	0·08 (0·05)	0·10 (0·10)	0·31 (0·28)	0·38 (0·38)	0·53 (0·50)	0·75 (0·75)
U	0·08 (0·05)	0·12 (0·10)	0·32 (0·29)	0·41 (0·39)	0·56 (0·53)	0·79 (0·77)

Note.—The figures in the brackets denote the values obtained after adjusting for the observed discrepancy between the observed and expected values of α when $\delta = 0$.

possible trend and is subject to further confirmation. It is proposed to carry on the investigation further by drawing a much larger number of samples and studying the comparative behaviour of the statistics.

2.2. At the suggestion of Savage, the power of the different statistics was also examined by using empirical percentage frequencies tabulated by Teichroew (1955) for 7,000 sets of samples of sizes 4 and 3 drawn respectively from the normal populations $N(0, 1)$ and $N(-\delta, 1)$ for varying values of δ. The empirical powers of the statistics W_2, W_3, W_4, T_3, T_4 and U for $\delta = 0, 0{\cdot}25, 0{\cdot}50, 0{\cdot}75$ and $1{\cdot}00$ are presented, in per cent., in Tables II A and II B.

It may be remarked here that the power values as obtained for the statistics T_3 and U are exact values whereas for other statistics some interpolation had to be done. Further, in Table II B the power for the statistics W_3 and W_4 and that for T_4 and U came out to be the same because the rankings to be selected in these cases for the chosen level of significance happened to be the same.

Tables II A and II B indicate that (i) for both the tails, the statistic T_3 is again likely to be more powerful than U or any other statistic, (ii) for a single tail, U is the most powerful of all and (iii) the statistic T_4 appears to be a good competitor, thus leading to the same conclusions as reached earlier.

TABLE II A

Empirical power for two-tailed statistics based on 7,000 samples

Statistic \ δ	0	0·25	0·50	0·75	1·00
W_2	11·20	12·93	15·93	19·77	25·36
W_3	11·33	13·10	16·32	20·86	26·93
W_4	11·45	12·41	15·79	20·30	26·58
T_3	11·23	13·48	18·49	24·44	32·65
T_4	11·49	13·00	18·16	24·07	32·32
U	11·82	12·46	17·60	23·74	31·97

TABLE II B

Empirical power for one-tailed statistics based on 7,000 samples

Statistic \ δ	0	0·25	0·50	0·75	1·00
W_2	11·39	14·94	19·96	25·24	31·99
W_3	10·89	15·90	23·67	31·02	39·15
W_4	10·89	15·90	23·67	31·02	39·15
T_3	11·27	17·03	25·20	33·63	43·35
T_4	11·42	18·04	27·38	36·70	48·51
U	11·42	18·04	27·38	36·70	48·51

Note.—The power values are all in percentage and the level of significance chosen is 4/35, *i.e.*, 11·43%.

3. APPLICATIONS

3·0. As already stated in 1.4, the statistics W's and T's can be employed for the testing of one or the other of the following hypotheses:

(i) whether a given sequence of observations is random;

(ii) whether two samples come from the same population; and

(iii) whether two sets of observations differ from each other by a specified amount or percentage.

We shall now illustrate the applications of some of these statistics in this section for testing the above hypotheses.

3.1. With a view to facilitating their applications, tables have been given in Appendices A and B for the statistics W_3, W_4, T_3 and T_4 which appear to be quite powerful and are at the same time comparatively easier to apply in practice.

3.1.1. Tables A-I to A-IV give the probability distributions of the statistics W_3, W_4, T_3 and T_4 for a few small values of n_1 and n_2, where n_1 and n_2 denote the sizes of the two samples. These tables would be useful in getting the probabilities associated with the observed values of the statistics under the null hypothesis.

3.1.2. When the sample sizes are large, the distributions of the statistics can be taken to be approximately normal and hence their standardized deviates can be used as test criteria for the purpose. The applications in such cases would be considerably simplified if the tables giving their means and standard deviations are available. Tables B-I and B-II give the values of the means and standard deviations of the statistics W_3, W_4, T_3 and T_4 for some values of n_1, n_2 and n, where n_1 and n_2 denote the sizes of the two samples and n that of the combined sample.

3.1.3. It may be noted that these tables can also be used for testing the randomness of a sequence of observations involving elements of two kinds.

3.2. *Randomness of a sequence.*—Suppose a sample of 50 cycle hub tubes has been collected as the tubes emerge from the lathe (*i.e.*, the machine) and 10 are found to be defective when their external diameters have been checked by a go and no-go gauge. It is required to know whether the defective ones occur at random or not. Let the sequence of the sample be as follows:—

GGGGDGGDGGGGGGGDDGGGGDGGGGGGGDDGGGGGDGGGD
GGGDGG

G denotes a good tube; *D* denotes a defective tube.

The application of the statistics *W* and *T* to the sequence would require the enumeration of the number of times that *G* comes before

D in the blocks and spacings. W' and T' would require the totalling of the scores from both G occurring before D and D before G. The observed and the expected values, the standard deviations and the corresponding standardized deviates for the statistics W_3, W_4, T_3, T_4 and W_3', W_4', T_3' and T_4' are given in Table III.

On comparing the values of the standardized deviates with 1·96, the value of a normal deviate at 5 per cent level, we find them all to be less than 1·96 indicating thereby that the validity of the null hypothesis about the random distribution of the defectives is not to be questioned.

TABLE III

Results of analysis for the G-D *sequence*

Statistic	Observed value	Expected value	Standard deviation	Standardized deviate $\left(\frac{o-e}{S.D.}\right)$
W_3	26	23·51	2·62	0·95
W_4	53	46·04	4·59	1·52
T_3	18	15.84	1.68	1.28
T_4	27	23·51	2·18	1·60
W_3'	50	47·02	4·97	0·60
W_4'	103	92·08	8·55	1·28
T_3'	36	31·67	3·10	1·39
T_4'	53	47·02	3·80	1·57

Hence the above statistics show that the given sequence can be considered to be a random one.

3·3. *Homogeneity of Two Samples*

3·3·1. *Small Samples*—In an experiment conducted to test the difference in impact strength on two types of insulating material, the following test results in foot-pounds were obtained:—

Type A—92, 92, 93, 86 and 91

Type B—88, 90, 95, 83 and 87.

It is required to determine whether the material of Type A possesses greater impact strength than that of Type B. If so, then only A may be preferred to B and not otherwise.

The hypothesis under test here is that the mean impact strength for the population of Type *A* material is less than or equal to the mean strength for the population of Type *B* material. The alternative hypothesis is that the population mean for *A* is greater than the population mean for *B*.

Combining the two types of test results in increasing order and replacing each observation by *A* or *B* depending on the type to which it belongs, we get the following sequence:—

BABBBAAAAB

Since the alternative here corresponds to the greater impact strength in Type *A* than in Type *B*, we shall consider the number of times that *A* precedes *B* in blocks or spacings of corresponding sizes for W_3, W_4, T_3 and T_4. The observed values of the statistics for the sequence and the probabilities associated with values as small as the observed values are given in Table IV.

TABLE IV

Probabilities for data on impact strength

Statistic	Observed Value	Probability of obtaining a value less than or equal to the observed value
W_3	5	0·214
W_4	8	0·123
T_3	4	0·377
T_4	6	0·397

It will be seen from Table IV that for all these four statistics the probability of obtaining a value less than or equal to the observed value works out to be more than 0·05 which may be chosen as our level of significance. This indicates that there is no significant difference in the impact strength on the two types of materials.

It would not be out of place to mention here that the applications of the one-tailed *t*-test and the *U*-statistic also lead to the same conclusion. The probability of obtaining a value less than or equal to the observed value (− 0·948) of *t* lies between 0·15 and 0·20 whereas the probability associated with the observed value (8) of *U* comes out to be equal to 0·210.

3.3.2. *Large Samples*: (i) *Same mean and same standard deviation.*—Table V shows two random samples of size 25 each drawn from a normal population with mean equal to 25 and standard deviation 10.

It is required to test whether the two samples can be regarded as having been drawn from the same population.

TABLE V

Two samples from the same population N (25, 10)

Sample *A* ..	17·743	31·849	21·628	18·255	33·655
Sample *B* ..	37·702	11·278	27·378	21·522	51·521
A	25·351	16·035	8·928	16·147	17·379
B	25·677	25·527	8·742	18·536	20·013
A	33·345	10·316	21·840	19·618	31·849
B	9·195	26·282	31·651	25·803	29·649
A	6·192	23·566	32·588	17·244	28·505
B	16·690	36·217	38·106	35·985	33·816
A	25·251	14·419	35·040	24·549	28·826
B	6·616	22·093	20·573	29·789	14·806

Pooling the two samples together, arranging the fifty observations in ascending order and replacing each of the observations by *A* or *B* according as it belongs to the sample *A* or *B** we obtain the following sequence:—

ABBABABABAABAAAABABBBAABAAAABBBBBAA
BBBAAAAABABBBBB

Now, for the application of the statistics W_3, W_4, T_3 and T_4 we shall consider the number of *BA* joins here rather than the *AB* joins, for, we get a smaller contribution in the former case. In general, whenever the alternative does not specify any direction and it is to be decided whether to take *AB* or *BA* joins for the applications of the statistics *W* and *T*, that type of join (*AB* or *BA*) shall be taken for which the contribution is smaller (or larger) so that the difference between the observed and expected values is maximum.

The results of examining the above sequence for homogeneity (or otherwise) of the two samples are given in Table VI.

* The ordering can be conveniently done in many cases by plotting the observations of the two samples as *A* or *B* on a line.

TABLE VI

Results of analysis for the samples from N (25, 10)

Statistic	Observed value	Expected value	Standard deviation	Standardized deviate
W_3	32	36·73	3·92	−1·21
W_4	65	71·94	6·48	−1·07
T_3	21	25·46	2·54	−1·76
T_4	33	36·73	3·17	−1·18

The absolute values of the standardized deviates are all less than 1·96, the value of the normal deviate at 5 per cent level of significance and hence they are all not significant. That is to say, all these statistics show, as expected, that there is no difference between the two samples and they can be regarded as having been drawn from the same population.

It may be noted that the applications of the t-test and the U-test also give the same kind of inference, the calculated values of t and standardized deviate for U being respectively equal to − 0·96 and − 1·00.

(ii) *Different means but same standard deviation.*—Table VII gives two random samples A and B of size 30 each drawn from two normal populations $N(0, 1)$ and $N(1, 1)$ respectively. The means are different in the two cases but the standard deviation is the same. It is required to examine whether or not the statistics W and T bring out the difference between the two samples. The hypothesis under test is that the two samples belong to the same population; the alternative is that they are from different populations and that B is stochastically larger than A.

Combining the two samples together and identifying each observation as A or B, we get the following sequence:—

AAAAAABAABABAAAAAAAABABBBAAAABBBBB
ABAABBBBBBAABBBBBABBABBBAB

For the application of the statistics W and T, we consider the BA joins here since B happens to be stochastically larger than A in the

TABLE VII

Two random samples from N (0, 1) *and* N (1, 1)

Sample *A*	1·678	—0·150	0·598	—0·899	—1·163
Sample *B*	0·463	—0·941	1·489	0·757	1·531
A	—0·261	—0·357	1·827	0·535	—2·056
B	0·556	1·658	0·115	0·372	1·402
A	—2·008	1·180	—1·141	0·358	—0·230
B	3·455	0·469	0·366	2·279	1·046
A	0·208	0·272	0·606	—0·307	—2·098
B	0·475	1·007	0·838	—0·618	1·378
A	0·079	—1·658	—0·344	—0·521	2·990
B	0·943	2·356	0·082	1·012	0·089
A	1·278	—0·144	—0·886	0·193	—0·199
B	2·237	—0·384	0·041	1·731	1·717

alternative, *i.e.*, the *B* observations tend to be larger than the *A*-observations. The number of *BA* joins is therefore expected to be smaller in this case than the *AB* joins.

The results of the examination of the above sequence for the difference between the two samples are presented in Table VIII.

TABLE VIII

Results of analysis for samples from N (0, 1) *and* N (1, 1)

Statistic	Observed value	Expected value	Standard deviation	Standardized deviate
W_3	33	44·24	4·30	—2·61
W_4	53	86·95	7·12	—4·77
T_3	22	29·75	2·77	—2·80
T_4	33	44·24	3·45	—3·26

The values of the standardized deviates are all much less than $-1 \cdot 96$, the value of the one-tailed normal deviate at $2 \cdot 5\%$ level, and hence they are all highly significant. This implies that the two samples should be regarded as having been taken from different populations and all the four statistics bring out this fact.

The calculated value of t comes out to be equal to $-3 \cdot 65$ which also demonstrates a significant difference between the two sample means and therefore between the two populations. The U-statistic too gives a similar verdict as its standardized deviate works out to be equal to $-3 \cdot 47$.

(iii) *Same mean but different standard deviations.*—Table IX gives two samples A and B of size 30 each drawn at random from two normal populations $N(0, 1)$ and $N(0, 2)$ respectively. The means are the same here but the standard deviations are different. The hypothesis under test is that the two samples can be considered to have been taken from the same population. The alternative hypothesis is that they have been drawn from different populations and that the spread of the observations is more for B than for A.

TABLE IX

Two random samples from N (0, 1) *and* N (0, 2)

Sample *A*	0·464	0·060	1·486	1·022	1·394	0·906
Sample *B*	1·536	0·750	−1·026	0·584	2·052	−2·668
A	1·179	−1·501	−0·690	1·372	−0·482	−1·376
B	−0·574	0·322	−2·692	2·500	1·260	0·750
A	−1·010	−0·005	1·393	−1·787	−0·105	−1·339
B	−2·840	−0·302	−0·618	0·848	1·186	1·724
A	1·041	0·279	−1·805	−1·186	0·658	−0·439
B	0·470	−3·706	0·274	−5·052	−0·708	−0·944
A	−1·399	0·199	0·159	2·273	0·041	−1·132
B	−1·110	−1·026	−2·110	−0·976	1·512	0·450

Arranging the two samples together in ascending order and denoting each observation by A or B depending on whether it falls in sample A or B we get the following sequence:—

BBBBBBAAAAAAAABBBABBBABBAABAAAAAABAB
BABBABBBAAAABBAAAABBBBAB

and $Y_1, Y_2, \cdots, Y_n$ differ from each other by (i) a given amount, say Q units, or (ii) a given percentage, say P per cent. We shall consider only the case where the X-observations are higher than the Y-observations, the other case being treated similarly.

3·4·1. *A given value.*—To test whether the X-observations exceed the Y-observations by Q units, each of the Y-observations would be increased by Q and then the homogeneity of the two sets $X_1, X_2, \cdots, X_n$ and $Y_1 + Q, Y_2 + Q, \cdots, Y_n + Q$ would be tested as usual with the help of the statistics W and T.

As an example, let us consider whether the B-observations in Table VII are higher than the A-observations by unity. A new set, say C, is constructed by increasing each of the A-observations by 1. Combining the B- and C-observations in ascending order, we get the sequence:—

$$CCCBCBBCCBBBCCBBBBBBCBCCCCBCCBCCBBBB$$

$$CCCCCBBBBCCCBBBCBCBBCCBC$$

Taking BC joins for the applications of the statistics W_3, W_4, T_3 and T_4 we get the values of the standardized deviates as 0·83, 1·55, 1·00 and 1·66 respectively, which are all less than 1·96, the value of the normal deviate at 5 per cent level and hence all of them are not significant. Accordingly, the two sets B and C are to be considered as homogeneous. This implies that, as expected, the B-observations are higher than the A-observations by unity.

The applications of the t-test and U-test also lead to the same inference, the calculated values of t and the standardized deviate for U being respectively equal to 0·065 and 0·059.

3·4·2. *A given percentage.*—To test whether the X-observations exceed the Y-observations by P per cent, the homogeneity of the sets $X_1, X_2, \cdots, X_n$ and $Y_1(1 + P/100), Y_2(1+P/100) \cdots, Y_n(1 + P/100)$ would be tested and appropriate inference drawn.

4. Summary

This paper makes a comparative study of the empirical powers of some of the new *statistics* W's and T's, recently developed by Iyer and Singh, by drawing random samples from a normal population. To facilitate the applications of the *statistics*, it provides tables giving their exact probability distributions for small samples and their means and standard deviations for large samples. It also discusses the applications of the *statistics* to varying situations which may require the testing

of the randomness of a sequence *or* the homogeneity of two samples *or* the difference of a given magnitude between two sets of observations.

The results show that the powers of some of the new *statistics* compare very favourably with that of the Mann-Whitney *statistic* (or Wilcoxon's test) as well as the *t*-test. In view of the fact that the new tests can admit general type of alternatives, they should prove on the whole more useful in practice than the other tests.

My sincere thanks are due to Dr. P. V. Krishna Iyer for having gone through the paper and making helpful suggestions.

REFERENCES

1. Bennet, C. A. and Franklin, N. L. — *Statistical Analysis in Chemistry and the Chemical Industry*, John Wiley & Sons, New York, 1954, 177.
2. Iyer, P. V. K. and Singh, B. N. — "On certain probability distributions arising from a sequence of observations and their applications," *J. Ind. Soc. Agric. Stat.*, 1955, **7**, 127.
3. Kendall, M. G. .. *The Advanced Theory of Statistics*, Charles Griffin & Co., 1952, **1**, 391.
4. Mahalanobis, P. C., Bose, S. S., Roy, P. R. and Banerji, S. K. — "Tables of random samples from a normal population," *Sankhyā*, 1934, **1**, 289.
5. Mann, H. B. and Whitney, D. R. — "On a test whether one of two random variables is stochastically larger than other," *Ann. Math. Stat.*, 1947, **18**, 50.
6. Moore, G. H. and Wallis, W. A. — "Time series significance tests based on signs of differences," *J. Am. Stat. Assocn.*, 1943, **38**, 153.
7. Teichroew, D. — "Empirical power functions for nonparametric two-sample tests for small samples", *Ann. Math. Stat.*, 1955, **26**, 340.
8. Wilcoxon, F. — "Individual comparisons by ranking methods," *Biometrics Bulletin*, 1945, **1**, 80.

APPENDIX A

Probability Tables for Two Small Samples

Tables A-I to A-IV give the cumulative probability distributions of the statistics W_3, W_4, T_3 and T_4 for a few small values of n_1 and n_2, the sizes of the two samples (*see* also 3.1.1).

Table A-I

Probabilities of obtaining values less than or equal to observed values of W_3

$n_1 = 3$

W_3 \ n_2	1	2	3
0	0·250	0·100	0·050
1	0·500	0·100	0·050
2	0·750	0·400	0·200
3	1·000	0·700	0·300
4		1·000	0·850
5			0·950
6			1·000

$n_1 = 4$

W_3 \ n_2	1	2	3	4
0	0·200	0·067	0·029	0·014
1	0·400	0·067	0·029	0·014
2	0·600	0·267	0·114	0·057
3	1·000	0·533	0·200	0·114
4		0·867	0·543	0·329
5		1·000	0·829	0·500
6			0·971	0·857
7			1·000	0·971
8				1·000

$n_2=5$

W_3 \ n_2	1	2	3	4	5	6
0	0·167	0·048	0·018	0·008	0·004	0·002
1	0·333	0·048	0·018	0·008	0·004	0·002
2	0·500	0·190	0·071	0·032	0·016	0·009
3	1·000	0·429	0·143	0·071	0·040	0·024
4		0·762	0·393	0·214	0·127	0·076
5		0·952	0·661	0·349	0·214	0·139
6		1·000	0·875	0·619	0·413	0·284
7			1·000	0·881	0·635	0·476
8				0·984	0·901	0·719
9				1·000	0·980	0·913
10					1·000	0·989
11						1·000

TABLE A-II

Probabilities of obtaining values less than or equal to observed values of W_4

$n_1 = 3$

W_4 \ n_2	1	2	3
0	0·250	0·100	0·050
1	0·500	0·100	0·050
2	0·750	0·300	0·050
3	1·000	0·500	0·200
4		0·700	0·450
5		0·800	0·500
6		0·900	0·500
7		1·000	0·900
8			0·950
9			0·950
10			1·000

$n_1 = 4$

W_4 \ n_2	1	2	3	4
0	0·200	0·067	0·029	0·014
1	0·400	0·067	0·029	0·014
2	0·400	0·200	0·029	0·014
3	0·800	0·267	0·114	0·057
4	0·800	0·333	0·143	0·114
5	1·000	0·667	0·286	0·157
6		0·800	0·371	0·300
7		0·867	0·514	0·414
8		0·933	0·743	0·571
9		1·000	0·886	0·786
10			1·000	0·943
11				0·986
12				1·000

$n_1 = 5$

W_4 \ n_2	1	2	3	4	5	6
0	0·167	0·048	0·018	0·008	0·004	0·002
1	0·333	0·048	0·018	0·008	0·004	0·002
2	0·333	0·143	0·018	0·008	0·004	0·002
3	0·667	0·190	0·071	0·032	0·016	0·009
4	0·667	0·238	0·089	0·032	0·016	0·009
5	0·833	0·381	0·189	0·063	0·032	0·017
6	1·000	0·619	0·214	0·095	0·052	0·030
7		0·762	0·339	0·143	0·063	0·037
8		0·857	0·464	0·230	0·123	0·069
9		1·000	0·714	0·349	0·167	0·100
10			0·893	0·524	0·282	0·171
11			0·946	0·706	0·409	0·240
12			1·000	0·865	0·611	0·370
13				0·952	0·746	0·511
14				0·984	0·845	0·654
15				1·000	0·972	0·799
16					0·992	0·920
17					0·996	0·972
18					1·000	0·996
19						1·000

TABLE A-III

Probabilities of obtaining values less than or equal to observed values of T_3

$n_1 = 3$

T_3 \ n_2	1	2	3
0	0·250	0·100	0·050
1	0·500	0·200	0·100
2	1·000	0·600	0·350
3		1·000	0·800
4			1·000

$n_1 = 4$

T_3 \ n_2	1	2	3	4
0	0·200	0·067	0·029	0·014
1	0·400	0·133	0·057	0·029
2	1·000	0·467	0·229	0·129
3		0·933	0·600	0·386
4		1·000	0·943	0·743
5			1·000	0·986
6				1·000

$n_1 = 5$

T_3 \ n_2	1	2	3	4	5
0	0·167	0·048	0·018	0·008	0·004
1	0·333	0·095	0·036	0·016	0·008
2	1·000	0·381	0·161	0·079	0·044
3		0·857	0·464	0·262	0·159
4		1·000	0·839	0·563	0·377
5			1·000	0·905	0·722
6				1·000	0·964
7					1·000

TABLE A-IV

Probabilities of obtaining values less than or equal to observed values of T_4

$n_1 = 3$

T_4 \ n_2	1	2	3
0	0·250	0·100	0·050
1	0·500	0·200	0·100
2	0·750	0·400	0·200
3	1·000	0·800	0·450
4		1·000	0·650
5			0·950
6			1·000

$n_1 = 4$

T_4 \ n_2	1	2	3	4
0	0·200	0·067	0·029	0·014
1	0·400	0·133	0·057	0·029
2	0·600	0·267	0·114	0·057
3	1·000	0·533	0·286	0·157
4		0·800	0·457	0·271
5		1·000	0·771	0·514
6			1·000	0·829
7				0·986
8				1·000

$n_1 = 5$

T_4 \ n_2	1	2	3	4	5
0	0·167	0·048	0·018	0·008	0·004
1	0·333	0·095	0·036	0·016	0·008
2	0·500	0·190	0·071	0·032	0·016
3	1·000	0·429	0·196	0·095	0·052
4		0·667	0·321	0·167	0·091
5		1·000	0·625	0·349	0·206
6			0·911	0·619	0·397
7			1·000	0·897	0·671
8				0·984	0·905
9				1·000	0·984
10					1·000

APPENDIX B

Mean and Standard Deviation Tables for Large Samples

Tables B-I and B-II give the means and standard deviations of the statistics W_3, W_4, T_3 and T_4 for a few large values of n_1, n_2 and n, the sizes of the two samples and the combined sample (*see* also 3·1·2).

Table B-I

Means and standard deviations of W_3 *and* W_4

(*a*) For W_3, $\mu_1' = 3\,(n-2)\,\dfrac{n_1 n_2}{n\,(n-1)}$

$$\sigma^2 = (9n-22)\frac{n_1 n_2}{n\,(n-1)} + (9n^2 - 67n + 128) \times \frac{n_1(n_1-1)\,n_2(n_2-1)}{n\,(n-1)\,(n-2)\,(n-3)} - \left\{3\,(n-2)\frac{n_1 n_2}{n\,(n-1)}\right\}^2$$

(*b*) For W_4, $\mu_1' = 6\,(n-3)\,\dfrac{n_1 n_2}{n\,(n-1)}$

$$\sigma^2 = (36n-136)\frac{n_1 n_2}{n\,(n-1)} + (36n^2 - 346n + 884) \times \frac{n_1(n_1-1)\,n_2(n_2-1)}{n\,(n-1)\,(n-2)\,(n-3)} - \left\{6\,(n-3)\frac{n_1 n_2}{n\,(n-1)}\right\}^2$$

n	n_1	n_2	W_3 Mean μ_1'	W_3 Standard deviation σ	W_4 Mean μ_1'	W_4 Standard deviation σ
20	10	10	14·21	2·45	26·84	3·98
	12	8	13·64	2·37	25·77	3·90
	15	5	10·66	1·96	20·13	3·48
24	12	12	17·22	2·69	32·87	4·39
	16	8	15·30	2·44	29·22	4·11
	18	6	12·91	2·13	24·65	3·75
30	15	15	21·72	3·02	41·90	4·95
	20	10	19·31	2·73	37·24	4·59
	24	6	13·90	2·07	26·81	3·73
	25	5	12·07	1·85	23·28	3·43
40	20	20	29·23	3·50	56·92	5·77
	30	10	21·92	2·79	42·69	4·68
	32	8	18·71	2·36	36·43	4·18
	35	5	12·79	1·71	24·90	3·24

TABLE B-I (*Contd.*)

n	n_1	n_2	W_3		W_4	
			Mean μ_1'	Standard deviation σ	Mean μ_1'	Standard deviation σ
	25	25	36·73	3·92	71·94	6·48
50	30	20	35·27	3·78	69·06	6·27
	40	10	23·51	2·62	46·04	4·59
	45	5	13·22	1·59	25·90	3·05
	30	30	44·24	4·30	86·95	7·12
60	40	20	39·32	3·85	77·29	6·46
	45	15	33·18	3·29	65·21	5·63
	50	10	24·58	2·58	48·31	4·45

TABLE B-II

Means and standard deviations of T_3 and T_4

(*a*) For T_3, $\mu_1' = (2n-3)\dfrac{n_1 n_2}{n(n-1)}$

$$\sigma^2 = (4n-7)\frac{n_1 n_2}{n(n-1)} + (4n^2-26n+44) \times \frac{n_1(n_1-1)\,n_2(n_2-1)}{n(n-1)(n-2)(n-3)} - \left\{(2n-3)\frac{n_1 n_2}{n(n-1)}\right\}^2$$

(*b*) For T_4, $\mu_1' = 3(n-2)\dfrac{n_1 n_2}{n(n-1)}$

$$\sigma^2 = (9n-22)\frac{n_1 n_2}{n(n-1)} + (9n^2-69n+146) \times \frac{n_1(n_1-1)\,n_2(n_2-1)}{n(n-1)(n-2)(n-3)} - \left\{3(n-2)\frac{n_1 n_2}{n(n-1)}\right\}^2$$

n	n_1	n_2	T_3		T_4	
			Mean μ_1'	Standard deviation σ	Mean μ_1'	Standard deviation σ
	10	10	9·74	1·64	14·21	2·11
20	12	8	9·35	1·59	13·64	2·05
	15	5	7·30	1·30	10·66	1·75
	12	12	11·74	1·79	17·22	2·28
24	16	8	10·43	1·62	15·30	2·09
	18	6	8·80	1·40	12·91	1·86

TABLE B-II (*Contd.*)

n	n_1	n_2	T_3		T_4	
			Mean μ_1'	Standard deviation σ	Mean μ_1'	Standard deviation σ
30	15	15	14·74	1·99	21·72	2·51
	20	10	13·10	1·79	19·31	2·29
	24	6	9·43	1·34	13·90	1·80
	25	5	8·19	1·19	12·07	1·63
40	20	20	19·74	2·28	29·23	2·86
	30	10	14·81	1·75	21·92	2·26
	32	8	12·64	1·52	18·71	2·00
	35	5	8·64	1·09	12·79	1·50
50	25	25	25·46	2·54	36·73	3·17
	30	20	23·76	2·44	35·27	3·06
	40	10	15·84	1·68	23·51	2·18
	45	5	8·91	1·01	13·22	1·40
60	30	30	29·75	2·77	44·24	3·45
	40	20	26·44	2·48	39·32	3·11
	45	15	22·31	2·11	33·18	2·69
	50	10	16·53	1·60	24·58	2·09