ON THE DISTRIBUTION OF THE TWO-SAMPLE CRAMÉR-VON MISES CRITERION[1]

By T. W. Anderson

Columbia University

1. Summary and introduction. The Cramér-von Mises ω^2 criterion for testing that a sample, $x_1, \cdots, x_N$, has been drawn from a specified continuous distribution $F(x)$ is

$$(1) \qquad \omega^2 = \int_{-\infty}^{\infty} [F_N(x) - F(x)]^2 \, dF(x),$$

where $F_N(x)$ is the empirical distribution function of the sample; that is, $F_N(x) = k/N$ if exactly k observations are less than or equal to $x (k = 0, 1, \cdots, N)$. If there is a second sample, $y_1, \cdots, y_M$, a test of the hypothesis that the two samples come from the same (unspecified) continuous distribution can be based on the analogue of $N\omega^2$, namely

$$(2) \qquad T = [NM/(N + M)] \int_{-\infty}^{\infty} [F_N(x) - G_M(x)]^2 \, dH_{N+M}(x),$$

where $G_M(x)$ is the empirical distribution function of the second sample and $H_{N+M}(x)$ is the empirical distribution function of the two samples together [that is, $(N + M)H_{N+M}(x) = NF_N(x) + MG_M(x)$]. The limiting distribution of $N\omega^2$ as $N \to \infty$ has been tabulated [2], and it has been shown ([3], [4a], and [7]) that T has the same limiting distribution as $N \to \infty$, $M \to \infty$, and $N/M \to \lambda$, where λ is any finite positive constant. In this note we consider the distribution of T for small values of N and M and present tables to permit use of the criterion at some conventional significance levels for small values of N and M. The limiting distribution seems a surprisingly good approximation to the exact distribution for moderate sample sizes (corresponding to the same feature for $N\omega^2$ [6]). The accuracy of approximation is better than in the case of the two-sample Kolmogorov-Smirnov statistic studied by Hodges [4].

2. The procedure. The cumulative distribution function $H_{N+M}(x)$ gives weight $1/(N + M)$ at each of the numbers $x_1, \cdots, x_N, y_1, \cdots, y_M$. The (Lesbesgue-Stieltjes) integral (2) is the sum

$$(3) \quad T = [NM/(N + M)^2] \left\{ \sum_{i=1}^{N} [F_N(x_i) - G_M(x_i)]^2 + \sum_{j=1}^{M} F_N(y_j) - G_M(y_j)]^2 \right\}.$$

Received August 28, 1961; revised April 12, 1962.

[1] This research was conducted largely at Stanford University and has been sponsored by the Office of Naval Research. This paper is a technical report of Contracts Nonr-3279(00) and Nonr-266(33). Reproduction in whole or in part is permitted for any purpose of the United States Government.

Let r_i and s_j be the ranks in the pooled sample of the ordered observations of the first and second samples, respectively ($i = 1, \cdots, N$ and $j = 1, \cdots, M$). Then

$$F_N(x) - G_M(x) = (i/N) - [(r_i - i)/M] \tag{4}$$

at the ith x-observation and

$$F_N(x) - G_M(x) = [(s_j - j)/N] - (j/M) \tag{5}$$

at the jth y-observation. (The probability of any two observations being equal is 0 under the null hypothesis.) The criterion is

$$\begin{aligned} T &= \frac{NM}{(N+M)^2}\left\{\sum_{i=1}^{N}\left[\frac{r_i}{M} - i\left(\frac{1}{M}+\frac{1}{N}\right)\right]^2 + \sum_{j=1}^{M}\left[\frac{s_j}{N} - j\left(\frac{1}{M}+\frac{1}{N}\right)\right]^2\right\} \\ &= \frac{1}{(N+M)^2}\left\{\frac{N}{M}\sum_{i=1}^{N}\left(r_i - \frac{N+M}{N}i\right)^2 + \frac{M}{N}\sum_{j=1}^{M}\left(s_j - \frac{N+M}{M}j\right)^2\right\}. \end{aligned} \tag{6}$$

If $M = N$, we can write

$$T = (4N^2)^{-1}\left\{\sum_{i=1}^{N}(r_i - 2i)^2 + \sum_{j=1}^{N}(s_j - 2j)^2\right\}. \tag{7}$$

Using the fact that

$$\sum_{i=1}^{N} r_i^2 + \sum_{j=1}^{M} s_j^2 = \sum_{k=1}^{N+M} k^2 = \frac{(N+M)(N+M+1)(2N+2M+1)}{6}, \tag{8}$$

we can write T as

$$T = \frac{U}{NM(N+M)} - \frac{4MN-1}{6(M+N)}, \tag{9}$$

where

$$U = N\sum_{i=1}^{N}(r_i - i)^2 + M\sum_{j=1}^{M}(s_j - j)^2. \tag{10}$$

To test the null hypothesis that the two samples are drawn from the same distribution, one orders all of the observations, determines the ranks $r_1 < r_2 < \cdots < r_N$ of the N observations from the first sample and the ranks $s_1 < s_2 < \cdots < s_M$ of the M observations from the second sample, and computes U. If U is too large, one rejects the null hypothesis.

When the null hypothesis is true every order of the two sets of observations is equally likely, and, hence, every set of N integers from $1, 2, \cdots, M + N$ is equally likely to be the ranks of the first sample. On this basis the distribution of U under the null hypothesis has been computed[2] for all combinations of sample

[2] I am indebted to Mrs. Ann Kinney Kretschmer for programming these computations. A description of the computational procedure and the complete tables of distributions are given in Mrs. Kretschmer's Master's Essay "Anderson's W-Test for Small Samples," Stanford University, July, 1955. Photostated copies of the tables can be obtained at cost from the Department of Statistics, Stanford University.

sizes N, $M = 1, 2, \cdots, 7$. Since the number of values the statistic takes on increases very rapidly with N and M, it is not feasible to give the full distributions. For some N and M Table 1 gives the larger values that U can take on together with the probabilities of U being that value or greater. In each case at least 10% of the distribution is included.

For larger values of M and N we give values of u such that the probability of observing a value of U at least that large is about 10% in Table 2, 5% in Table 3, 1% in Table 4, and .1% in Table 5. In each case probabilities are given which straddle the stated percentage. (Some additional probabilities are given for the purpose of comparing with the limiting distribution.) If the statistician wishes to achieve exactly a given significance level, he can randomize appropriately.

The expected value of T (under the null hypothesis) is

$$\mathcal{E}T = (1/6) + \{1/[6(M + N)]\}, \tag{11}$$

as compared with the mean of $\frac{1}{6}$ of the limiting distribution. The variance of T is the variance of U divided by $N^2M^2(N + M)^2$; the variance of U is

$$\begin{aligned} \operatorname{Var}(U) = {} & (N - M)^2 \operatorname{Var}(\textstyle\sum r_i^2) + 4N^2 \operatorname{Var}(\sum ir_i) + 4M^2 \operatorname{Var}(\sum js_j) \\ & - 4N(N - M) \operatorname{Cov}(\textstyle\sum r_i^2, \sum ir_i) - 4M(M - N) \operatorname{Cov}(\sum s_j^2, \sum js_j) \\ & + 8NM \operatorname{Cov}(\textstyle\sum ir_i, \sum js_j), \end{aligned} \tag{12}$$

where (8) has been used to reduce the terms. The necessary variances and covariances have been given by Wegner [9] except

$$\operatorname{Cov}(\textstyle\sum ir_i, \sum js_j) = -\dfrac{NM(N + M + 1)(8NM + 7N + 7M + 8)}{360}. \tag{13}$$

Then the variance of T is

$$\operatorname{Var}(T) = \frac{1}{45} \cdot \frac{M + N + 1}{(M + N)^2} \cdot \frac{4MN(M + N) - 3(M^2 + N^2) - 2MN}{4MN}, \tag{14}$$

as compared with the variance of $\frac{1}{45}$ of the limiting distribution. Some values of the means and standard deviations are indicated in Table 6. In Tables 2 to 5 are given some values of t corresponding to u and some values of t adjusted so the resulting quantity has mean $\frac{1}{6}$ and variance $\frac{1}{45}$ (called "normalized t").

For the moderate sample sizes considered here the probabilities are already very close to those of the limiting distribution in the upper tail. The last line of each of Tables 2 to 5 gives the corresponding value of t for which the limiting probability is the desired significance level. It will be seen that for the larger values of N and M the probabilities correspond quite closely to those of the asymptotic distribution.

One way of using the limiting distribution as an approximate distribution for determining whether an observed value is significant is to adjust an observed T as

$$[(T - \mathcal{E}T)/\{45 \operatorname{Var}(T)\}^{\frac{1}{2}}] + \tfrac{1}{6} \tag{15}$$

and compare this value with the desired significance point of the limiting distribution. In Table 7 we give the difference between the actual significance level and the nominal significance level when using such a procedure. Roughly speaking, at the 5% and 1% significance levels and these values of N and M the relative error is about $\frac{1}{10}$. While these numerical results are given only for $M \leqq 8$ and $N \leqq 8$, they suggest that for larger values of N and M the limiting distribution could be used to approximate significance levels between 1% and 10% with a relative error of generally less than $\frac{1}{10}$.

It is inevitable that there is some difference between the distribution for a given N and M and the limiting distribution because the first is discrete and the second is continuous. The values that T can take on are limited to certain rational numbers and jumps in its distribution function are limited in value to certain rational numbers. However, in addition to the differences between distribution functions due to jumps there are more systematic differences. In Table 8 we give bounds on the absolute value of the difference between the limiting distribution and the distribution of $T - 1/[6(M + N)]$ for several ranges in the upper tail.

In general the distribution for $N \neq M$ is smoother than for $N = M$ because in the latter case the number of values U can take on is more limited. At 10% the distribution function is increasing so rapidly that in the discrete cases the jumps are big. Near .1% errors are relatively large because the limiting distribution is unbounded while each of the discrete distributions has one last jump. Continuity corrections do not seem feasible because for a given pair, N and M, the jumps are not equal in size and the intervals between jumps are not equal.

E. J. Burr[3] has more recently extended the computations summarized in this paper to larger values of N and M. On the basis of his further study of the relationship between the limiting distribution and the distributions computed for some values of $N = M$, he has suggested an empirical correction formula for using the limiting distribution to approximate the exact distribution.

3. Some remarks.

3.1. *Other tabulations.* Sundrum [8] has tabulated a closely related statistic suggested by Lehmann for $N = M = 2, 3, 4, 5$ and $N = 2$, $M = 3$, and $N = 3$, $M = 4$, and $N = 4$, $M = 5$. The difference between this statistic and T suggested here is that $\frac{1}{2}[F_N(x) + G_M(x)]$ is used in defining the integral (2) instead of $H_{N+M}(x)$. As Wegner [9] has indicated, when $N = M$ this statistic is T. For the cases tabulated by Sundrum, T takes on more values than this statistic when $N \neq M$. At the present time there is no theoretical basis for choosing between the two statistics for $N \neq M$, but the pooled empirical distribution function $H_{N+M}(x)$ seems more natural than the unweighted average of the $F_N(x)$ and $G_M(x)$.

---

[3] I am indebted to Mr. Burr for checking some of my calculations against his and for seeing his results before publication.

Kurup [5] has tabulated a related statistic suggested earlier by Mood. This statistic is defined by using $F_N(x)$ or $G_M(x)$ instead of $H_{N+M}(x)$. When the null hypothesis is true, all of these statistics have the same limiting distribution. More references are given in the papers cited.

3.2. *Asymptotic power.* It is easy to see that this test procedure is consistent since $F_N(x)$ and $G_M(y)$ are consistent estimates of the distributions from which the samples are drawn. For a more accurate analysis of the asymptotic power of the test consider two sequences of continuous distributions $\{F_N^*(x)\}$ and $\{G_M^*(x)\}$ such that

$$\lim_{N\to\infty} F_N^*(x) = H(x), \qquad \lim_{M\to\infty} G_M^*(x) = H(x), \tag{16}$$

$$\lim_{N\to\infty} N^{\frac{1}{2}}[F_N^*(x) - H(x)] = f[H(x)], \tag{17}$$

$$\lim_{M\to\infty} M^{\frac{1}{2}}[G_M^*(x) - H(x)] = g[H(x)], \tag{18}$$

uniformly, $H(x)$ is continuous, and $f(u)$ and $g(u)$ $(0 \leqq u \leqq 1)$ are square-integrable. Then the limiting distribution of T as $N \to \infty$, $M \to \infty$, and $N/M \to \lambda$ is the distribution of

$$\int_0^1 \{Y(u) + (1+\lambda)^{-\frac{1}{2}} f(u) - [\lambda/(1+\lambda)]^{\frac{1}{2}} g(u)\}^2\, du, \tag{19}$$

where $Y(u)$ is a Gaussian stochastic process with mean 0 and covariance function $\min(u, v) - uv$. The characteristic function of (25) is the product of the characteristic function of $\int_0^1 Y^2(u)du$ (that is, the limiting characteristic function of $N\omega^2$ under the null hypothesis) and

$$\exp\left\{it \int_0^1 k^2(u)\, du - it \int_0^1 \int_0^1 R_{2it}(u, v) k(u) k(v)\, du\, dv\right\}, \tag{20}$$

where

$$\begin{aligned} R_{2it}(u, v) &= R_{2it}(v, u) \\ &= -[(2it)^{\frac{1}{2}}/\sin(2it)^{\frac{1}{2}}] \sin((2it)^{\frac{1}{2}}u) \sin[(2it)^{\frac{1}{2}}(1-v)], \quad u \leqq v, \end{aligned} \tag{21}$$

$$k(u) = (1+\lambda)^{-\frac{1}{2}} f(u) - [\lambda/(1+\lambda)]^{\frac{1}{2}} g(u). \tag{22}$$

This result is discussed in [1].

3.3. *Some modifications.* Statistics slightly different from T can be derived from the integral (2) or sum (3) by discarding the usual convention that a cumulative distribution function is continuous on the right. More precisely, if x_i is the kth x in order of magnitude (that is, if there are $k - 1$ x's less than x_i), we could define $F_N(x_i)$ as any number between $(k-1)/N$ and k/N; similarly $G_M(y_j)$ could be any number between $(k-1)/M$ and k/M if y_j is the kth y in order of magnitude. If these are defined as $(k-a)/N$ and $(k-a)/M$, respectively, T is changed only by subtraction of $a(1-a)/(M+N)$. This fact shows incidentally that the statistic is unchanged by replacing the convention of continuity on the right $(a = 0)$ by continuity on the left $(a = 1)$. If the values of the empirical cumulative distribution function at the jumps are chosen to minimize the expected value of the statistic (under the null hypothesis), the statistic is

$$\frac{NM}{(N+M)^2}\left\{\frac{1}{M^2}\sum_{i=1}^{N}\left(r_i - \frac{M+N+1}{N+1}i\right)^2 + \frac{1}{N^2}\sum_{j=1}^{M}\left(s_j - \frac{M+N+1}{M+1}j\right)^2\right\}; \tag{23}$$

that is, the quantities squared are the differences between the ranks and their expected values.

In the definition of $N\omega^2$ one might replace $dF(x)$ by $dF_N(x)$ to obtain

$$S = N\int_{-\infty}^{\infty}[F_N(x) - F(x)]^2\,dF_N(x) = \sum_{i=1}^{N}[F_N(x_i^*) - F(x_i^*)]^2, \tag{24}$$

where $x_1^* < \cdots < x_N^*$ are the ordered observations. If $F_N(x_i^*)$ is defined as $(i - \frac{1}{2})/N$, the statistic is $N\omega^2 - 1/(12N)$. If $F_N(x_i^*)$ is defined to minimize the expected value of S (under the null hypothesis), the statistic is

$$\sum_{i=1}^{N}\{F(x_i^*) - [i/(N+1)]\}^2; \tag{25}$$

here $\mathcal{E}F(x_i^*) = i/(N+1)$.

These modifications do not affect the asymptotic theory; however, there might be some that have advantages for small samples.

REFERENCES

[1] ANDERSON, T. W. (1954). Distributions of some integrals of certain stochastic processes and limiting distributions of some 'goodness of fit' criteria. Unpublished (abstract, *Ann. Math. Statist.* **25** 174–175).

[2] ANDERSON, T. W. and DARLING, D. A. (1952). Asymptotic theory of certain 'goodness of fit' criteria based on stochastic processes. *Ann. Math. Statist.* **23** 193–212.

[3] FISZ, M. (1960). On a result by M. Rosenblatt concerning the Von Mises-Smirnov test. *Ann. Math. Statist.* **31** 427–429.

[4] HODGES, JR., J. L. (1957). The significance probability of the Smirnov two-sample test. *Ark. Mat.* **3** 469–486.

[4a] KIEFER, J. (1959). *K*-sample analogues of the Kolmogorov-Smirnov and Cramér-v. Mises tests. *Ann. Math. Statist.* **30** 420–447.

[5] KURUP, R. S. (1954). The problem of two samples from non-parametric populations. *Bull. Central Res. Inst., Univ. of Travancore, Trivandrum, Ser. B* **2** 21–59.

[6] MARSHALL, A. W. (1958). The small sample distribution of $n\,\omega_n^2$. *Ann. Math. Statist.* **29** 307–309.

[7] ROSENBLATT, M. (1952). Limit theorems associated with variants of the von Mises statistic. *Ann. Math. Statist.* **23** 617–623.

[8] SUNDRUM, R. M. (1954). On Lehmann's two-sample test. *Ann. Math. Statist.* **25** 139–145.

[9] WEGNER, L. H. (1956). Properties of some two-sample tests based on a particular measure of discrepancy. *Ann. Math. Statist.* **27** 1006–1016.

TABLE 1
Upper Tails of Distributions

N	M	u	$\Pr\{U = u\}$	$\Pr\{U \geqq u\}$	t
2	4	64	2/15	.133 333	.472 222
2	5	100	2/21	.095 238	.500 000
		87	2/21	.190 476	.314 286
2	6	144	2/28	.071 429	.520 833
		128	2/28	.142 859	.354 167
2	7	196	2/36	.055 556	.537 037
		177	2/36	.111 111	.386 243
3	3	81	1/10	.100 000	.527 778
3	4	144	2/35	.057 143	.595 238
		127	2/35	.114 286	.392 857
3	5	225	2/56	.035 714	.645 833
		203	2/56	.071 429	.462 500
		191	2/56	.107 143	.362 500
3	6	324	2/84	.023 810	.685 185
		297	2/84	.047 619	.518 519
		282	2/84	.071 429	.425 926
		279	2/84	.095 238	.407 407
		276	2/84	.119 048	.388 889
3	7	441	2/120	.016 667	.716 667
		409	2/120	.033 333	.564 286
		391	2/120	.050 000	.478 572
		387	2/120	.066 667	.459 524
		383	2/120	.083 333	.440 476
		365	2/120	.100 000	.354 762
4	4	256	1/35	.028 571	.687 500
		232	1/35	.057 143	.500 000
		216	2/35	.114 286	.375 000
4	5	400	2/126	.015 873	.759 259
		369	2/126	.031 746	.587 037
		348	2/126	.047 619	.470 370
		346	2/126	.063 492	.459 259
		337	2/126	.079 365	.409 259
		336	2/126	.095 238	.403 704
		331	2/126	.111 111	.375 926
5	5	625	1/126	.007 937	.850 000
		585	1/126	.015 873	.690 000
		555	2/126	.031 746	.570 000
		535	2/126	.047 619	.490 000
		525	5/126	.087 302	.450 000
		505	3/126	.111 111	.370 000

TABLE 2

Significance Levels Near 10%

N	M	u	Pr $\{U \geqq u\}$	t	Normalized t
4	6	472	$\frac{18}{210} = .085\ 714$	.383 333	
		468	$\frac{22}{210} = .104\ 762$	.366 667	
4	7	634	$\frac{32}{330} = .096\ 970$	.376 623	
		631	$\frac{34}{330} = .103\ 030$	.366 883	
5	6	718	$\frac{46}{462} = .099\ 567$	.372 727	
		710	$\frac{48}{462} = .103\ 896$	.348 485	
5	7	967	$\frac{78}{792} = .098\ 485$	.371 825	
		963	$\frac{80}{792} = .101\ 010$	.362 302	
6	6	1020	$\frac{43}{462} = .093\ 074$	.375 000	.371 314
		1008	$\frac{59}{462} = .127\ 706$	.347 222	.342 078
6	7	1374	$\frac{166}{1716} = .096\ 737$	.375 458	.372 127
		1373	$\frac{172}{1716} = .100\ 233$	.373 626	.370 207
		⋮			
		1362	$\frac{194}{1716} = .113\ 054$	.353 480	.349 084
		1359	$\frac{196}{1716} = .114\ 219$	.347 985	.343 324
7	7	1855	$\frac{160}{1716} = .093\ 240$	.382 653	.379 626
		1841	$\frac{185}{1716} = .107\ 809$	.362 245	.358 330
		1827	$\frac{197}{1716} = .114\ 802$	.341 837	.337 034
∞	∞		.10	.347 30	.347 30

TABLE 3

Significance Levels Near 5%

N	M	u	Pr $\{U \geqq u\}$	t	Normalized t
4	6	498	$\frac{10}{210} = .047\ 619$	.491 667	
		496	$\frac{12}{210} = .057\ 143$	.483 333	
4	7	671	$\frac{16}{330} = .048\ 485$	.496 753	
		654	$\frac{18}{330} = .054\ 545$	.441 558	
5	6	756	$\frac{22}{462} = .047\ 619$	.487 879	
		755	$\frac{24}{462} = .051\ 948$	.484 848	
5	7	1011	$\frac{38}{792} = .047\ 980$	.476 587	
		1009	$\frac{40}{792} = .050\ 505$	.471 826	
6	6	1080	$\frac{18}{462} = .038\ 961$	.513 889	.517 490
		1068	$\frac{25}{462} = .054\ 113$	.486 111	.488 254
		1044	$\frac{31}{462} = .067\ 100$	.430 556	.429 784
6	7	1423	$\frac{84}{1716} = .048\ 951$	.465 202	.466 216
		1419	$\frac{86}{1716} = .050\ 117$	.457 876	.458 535
7	7	1925	$\frac{84}{1716} = .048\ 951$	.484 694	.486 106
		1911	$\frac{96}{1716} = .055\ 944$	.464 286	.464 809
		1897	$\frac{117}{1716} = .068\ 182$	.443 878	.443 514
∞	∞		.05	.461 36	.461 36

TABLE 4

Significance Levels Near 1%

N	M	u	$\Pr\{U \geqq u\}$	t	Normalized t
4	6	576	$\frac{2}{210} = .009\ 524$	.816 667	
		538	$\frac{4}{210} = .019\ 048$	.658 333	
4	7	784	$\frac{2}{330} = .006\ 061$	.863 636	
		739	$\frac{4}{330} = .012\ 121$	.717 532	
5	6	851	$\frac{4}{462} = .008\ 658$	.775 758	
		814	$\frac{6}{462} = .012\ 987$	.663 636	
5	7	1123	$\frac{6}{792} = .007\ 576$	.743 254	
		1119	$\frac{8}{792} = .010\ 101$	.733 730	
6	6	1188	$\frac{4}{462} = .008\ 658$	.763 889	.780 607
		1152	$\frac{6}{462} = .012\ 987$	.680 556	.692 902
6	7	1577	$\frac{14}{1716} = .008\ 159$	.747 253	.761 924
		1564	$\frac{16}{1716} = .009\ 324$	.723 443	.736 962
		1552	$\frac{18}{1716} = .010\ 490$	.701 465	.713 919
7	7	2121	$\frac{14}{1716} = .008\ 159$	.770 408	.784 247
		2107	$\frac{18}{1716} = .010\ 490$	.750 000	.762 952
		2079	$\frac{20}{1716} = .011\ 655$	.709 184	.720 360
8	8	3472	$\frac{63}{6435} = .009\ 790$	.734 375	.744 648
		3456	$\frac{69}{6435} = .010\ 723$	.718 750	.728 443
∞	∞		.01	.743 46	.743 46

TABLE 5

Significance Levels Near .1%

N	M	u	$\Pr\{U \geqq u\}$	t	Normalized t
6	6	1296	$\frac{1}{462} = .002\ 165$	1.013 889	1.043 725
6	7	1764	$\frac{2}{1716} = .001\ 166$	1.089 744	1.120 999
7	7	2401	$\frac{1}{1716} = .000\ 583$	1.178 571	1.210 166
		2317	$\frac{2}{1716} = .001\ 166$	1.056 122	1.082 390
8	8	4096	$\frac{1}{6435} = .000\ 156$	1.343 750	1.376 647
		3984	$\frac{2}{6435} = .000\ 311$	1.234 375	1.263 211
		3888	$\frac{4}{6435} = .000\ 622$	1.140 625	1.165 981
		3808	$\frac{6}{6435} = .000\ 933$	1.062 500	1.084 955
		3792	$\frac{7}{6435} = .001\ 088$	1.046 875	1.068 750
∞	∞		.001	1.167 86	1.167 86

TABLE 6

Mean and Variance of T Related to the Limiting Mean and Variance

N	M	Mean $- \frac{1}{6}$	$\sqrt{\text{Variance} \times 45}$
6	6	$\frac{1}{72} = .013\ 889$	$\sqrt{\frac{65}{72}} = \frac{1}{1.052\ 470}$
6	7	$\frac{1}{78} = .012\ 821$	$\sqrt{\frac{615}{676}} = \frac{1}{1.048\ 421}$
7	7	$\frac{1}{84} = .011\ 905$	$\sqrt{\frac{45}{49}} = \frac{1}{1.043\ 498}$
8	8	$\frac{1}{96} = .010\ 417$	$\sqrt{\frac{119}{128}} = \frac{1}{1.037\ 126}$

TABLE 7

Difference Between Actual Significance Level and Nominal Significance Level Using Limiting Distribution

N	M	Nominal Level	Difference	Difference Relative To Nominal Level
6	6	.10	−.006 926	−.069
		.05	.004 113	.082
		.01	−.001 342	−.134
6	7	.10	.013 054	.131
		.05	−.001 049	−.021
		.01	−.001 841	−.184
7	7	.10	.007 809	.078
		.05	.005 944	.119
		.01	.000 490	.049
		.001	−.000 417	−.417
8	8	.01	−.000 210	−.021
		.001	−.000 689	−.689

TABLE 8

Maximum Absolute Differences Between Distributions and Limiting Distribution

		Range of $T - [1/(6(M+N))]$		
N	M	$.34730 - \infty$	$.46136 - \infty$	$.74346 - \infty$
5	5	.0282	.0149	.0061
5	6	.0090	.0089	.0083
5	7	.0077	.0063	.0036
6	6	.0112	.0079	.0053
6	7	.0058	.0056	.0042
7	7	.0101	.0058	.0040