ON THE KOLMOGOROV-SMIRNOV TEST FOR NORMALITY WITH MEAN AND VARIANCE UNKNOWN

HUBERT W. LILLIEFORS
The George Washington University

The standard tables used for the Kolmogorov-Smirnov test are valid when testing whether a set of observations are from a completely specified continuous distribution. If one or more parameters must be estimated from the sample then the tables are no longer valid.

A table is given in this note for use with the Kolmogorov-Smirnov statistic for testing whether a set of observations is from a normal population when the mean and variance are not specified but must be estimated from the sample. The table is obtained from a Monte Carlo calculation.

A brief Monte Carlo investigation is made of the power of the test.

The Kolmogorov-Smirnov statistic provides a means of testing whether a set of observations are from some completely specified continuous distribution, $F_0(X)$. The usual alternative would be the chi-square test. The Kolmogorov-Smirnov test has at least two major advantages over the chi-square test [ref. 1, 2].

1. It can be used with small sample sizes, where the validity of the chi-square test would be questionable.
2. Often it appears to be a more powerful test than the chi-square test for any sample size.

Unfortunately, when certain parameters of the distribution must be estimated from the sample, then the Kolmogorov-Smirnov test no longer applies—at least not using the commonly tabulated critical points. It is suggested in ref. 2 that if the test is used in this case, the results will be conservative in the sense that the probability of a type I error will be smaller than as given by tables of the Kolmogorov-Smirnov statistic [as found in ref. 2 or 4]. As will be seen below, the results of this procedure will indeed be extremely conservative.

In ref. 1 it is pointed out that if the parameters to be estimated are parameters of scale or location, then one can construct tables for use with the Kolmogorov-Smirnov statistic for that particular distribution.

This note presents a table for use with the Kolmogorov-Smirnov statistic when testing that a set of observations are from a normal population but the mean and variance are not specified.

The procedure is: Given a sample of N observations, one determines $D = \max_X \mid F^*(X) - S_N(X) \mid$, where $S_N(X)$ is the sample cumulative distribution function and $F^*(x)$ is the cumulative normal distribution function with $\mu = \overline{X}$, the sample mean, and $\sigma^2 = s^2$, the sample variance, defined with denominator $n-1$. If the value of D exceeds the critical value in the table, one rejects the hypothesis that the observations are from a normal population.

The values in the table were obtained by a Monte Carlo calculation. For each value of N, 1,000 or more samples were drawn and the distribution of D was thus estimated. The calculations were performed at The George Washington University Computing Center.

When the values are compared with those in the standard table for the Kolmogorov-Smirnov test [ref. 2, 4] it is found that the Monte Carlo critical values are in most cases approximately two-thirds the standard values. Since the ratio of the Monte Carlo values to the standard values remains relatively fixed, especially for the larger values of N, it appeared that the values were then decreasing as $1/\sqrt{N}$. The Monte Carlo values for a sample of size 40 were multiplied by the square root of 40 and the result was used as the numerator for the critical values for sample sizes greater than 30. In ref. 3 values were obtained via a similar calculation for $N=100$ using 400 samples. The values were in accord with the "asymptotic" values given in Table 1.

Comparing Table 1 with the standard table for the Kolmogorov-Smirnov test from ref. 2, it is seen that the critical values in Table 1 for a .01 significance level are for each value of N slightly smaller than critical values for a .20 significance level using the standard tables. Thus the result of using the standard table when values of the mean and standard deviation are estimated from the

TABLE 1. TABLE OF CRITICAL VALUES OF D

The values of D given in the table are critical values associated with selected values of N. Any value of D which is greater than or equal to the tabulated value is significant at the indicated level of significance. These values were obtained as a result of Monte Carlo calculations, using 1,000 or more samples for each value of N.

Sample Size N	Level of Significance for $D=\text{Max}\ \lvert F^*(X)-S_N(X)\rvert$				
	.20	.15	.10	.05	.01
4	.300	.319	.352	.381	.417
5	.285	.299	.315	.337	.405
6	.265	.277	.294	.319	.364
7	.247	.258	.276	.300	.348
8	.233	.244	.261	.285	.331
9	.223	.233	.249	.271	.311
10	.215	.224	.239	.258	.294
11	.206	.217	.230	.249	.284
12	.199	.212	.223	.242	.275
13	.190	.202	.214	.234	.268
14	.183	.194	.207	.227	.261
15	.177	.187	.201	.220	.257
16	.173	.182	.195	.213	.250
17	.169	.177	.189	.206	.245
18	.166	.173	.184	.200	.239
19	.163	.169	.179	.195	.235
20	.160	.166	.174	.190	.231
25	.149	.153	.165	.180	.203
30	.131	.136	.144	.161	.187
Over 30	$\frac{.736}{\sqrt{N}}$	$\frac{.768}{\sqrt{N}}$	$\frac{.805}{\sqrt{N}}$	$\frac{.886}{\sqrt{N}}$	$\frac{1.031}{\sqrt{N}}$

TABLE 2

Probability of rejecting hypothesis of normality using D statistic and chi-square statistic when sample size is 20. The numbers are the result of Monte Carlo calculations with 500 samples for each distribution.

Underlying Distribution	Kolmogorov-Smirnov test		Chi-Square test	
	Using Critical Values From Table 1		Using Monte Carlo Critical Values	
	$\alpha = .05$	$\alpha = .10$	$\alpha = .06$	$\alpha = .12$
Normal	.06	.10	.06	.12
Chi-square, 3 d.f.	.44	.55	.20	.27
Student's t, 3 d.f.	.50	.58	.40	.52
Exponential	.61	.72	.29	.41
Uniform	.12	.22	.10	.18

sample would be to obtain an extremely conservative test in the sense that the actual significance level would be much lower than that given by the table.

It would appear that this specialized Kolmogorov-Smirnov test for normality should have the same advantages over the chi-square test as does the usual Kolmogorov-Smirnov test when testing for a completely specified distribution. Clearly it provides a test which can be used with sample sizes which are too small for use of the chi-square test. It is shown in ref. 3 that asymptotically it is more powerful than the chi-square test.

A brief Monte Carlo investigation was made of the power of this test. Five hundred samples of size 20 were drawn from each of several distributions. The probability of rejection using the Kolmogorov-Smirnov test (Table 1) was determined. The results are given in Table 2. The value of chi-square was also determined for each sample (using four intervals). The intervals were determined so as to have equal probabilities under the fitted normal curve. It was shown in ref. 5 that the asymptotic distribution of chi-square lies between chi-square with one degree of freedom and chi-square with three degrees of freedom. This is due to the use of maximum likelihood estimators based on the individual observations rather than data grouped into cell frequencies (in which case the distribution would be chi-square with one degree of freedom). When

TABLE 3

Probability of rejecting hypothesis of normality using D statistic when sample size is 10. The numbers are the result of Monte Carlo calculations with 500 samples for each distribution.

Underlying Distribution	$\alpha = .05$	$\alpha = .10$
Normal	.05	.10
Chi-Square, 3 d.f.	.23	.35
Student's t, 3 d.f.	.28	.36
Exponential	.34	.46
Uniform	.07	.13

the standard chi-square point for $\alpha = .05$ and one degree of freedom was compared to the Monte Carlo results it was found that the probability of a type I error was .11. Since this probability was so far from the nominal value, rejection points were determined for chi-square from the Monte Carlo calculations. The values of 5.2 and 4.0 were found to give probabilities of type I error of 06 and .12 respectively. The probability of rejection was tabulated using these new critical values.

Probabilities of rejecting the hypothesis of normality were also determined (again using a Monte Carlo calculation) for a sample of size 10 using the Kolmogorov-Smirnov statistic and the critical points of Table 1. These results are given in Table 3.

REFERENCES

[1] David, F. N. and Johnson, N. L., "The Probability Integral Transformation When Parameters Are Estimated From the Sample," *Biometrika*, 35 (1948), 182–90.

[2] Massey, F. J., "The Kolmogorov-Smirnov Test for Goodness of Fit," *Journal of the American Statistical Association*, 46 (1951), 68–78.

[3] Kac, M., Kiefer, J., and Wolfowitz, J., "On Tests of Normality and Other Tests of Goodness of Fit Based on Distance Methods," *The Annals of Mathematical Statistics*, 26 (1955), 189–211.

[4] Birnbaum, Z. W., "Numerical Tabulation of the Distribution of Kolmogorov's Statistic for Finite Sample Size," *Journal of the American Statistical Association*, 47 (1952), 425–41.

[5] Chernoff, H. and Lehmann, E. L., "The Use of Maximum Likelihood Estimates in χ^2 Tests for Goodness of Fit," *The Annals of Mathematical Statistics*, 25 (1954), 579–86.