

On Testing Normality Using Several Samples:

An Analysis of Peanut Aflatoxin Data

by

C. P. Quesenberry, T. B. Whitaker and J. W. Dickens

Institute of Statistics  
Mimeo Series #1013  
Raleigh - 1975

ON TESTING NORMALITY USING SEVERAL SAMPLES:

AN ANALYSIS OF PEANUT AFLATOXIN DATA

C. P. Quesenberry  
Department of Statistics,  
North Carolina State University,  
Raleigh, N. C. 27607

T. B. Whitaker and J. W. Dickens  
USDA-ARS, Biological and Agricultural Engineering Department  
North Carolina State University  
Raleigh, N. C. 27607

SUMMARY

Eight samples of size 16 and 3 samples of size 15 consisting of replicate determinations of aflatoxin in peanut subsamples are considered. This data is analyzed to decide whether the samples could have arisen from two-parameter normal parents, or from two-parameter lognormal parents. The method of analysis consists of first transforming the individual samples in such a way that the transformed values have uniform distributions on the unit interval under the null hypothesis, the transformed values are then pooled and tested for uniformity. The analysis leads to the conclusion that the normal model fits the data quite well, and definitely better than the lognormal model.

## 1. INTRODUCTION

### 1.1 The Biological Problem

Aflatoxin is a toxic material produced in peanuts by the fungus *Aspergillus flavus*, Bampton [1963]. As a precautionary measure all commercial lots of peanuts in the U. S. (approximately 20,000 each crop year) are tested for aflatoxin, Duggan [1970]. A lot of shelled peanuts may vary in size up to 100,000 pounds. The concentration of aflatoxin in a lot of shelled peanuts is estimated by measuring the aflatoxin concentration in a random sample of kernels drawn from the lot. The sample is comminuted in a subsampling mill, Dickens and Satterwhite [1969], and a subsample of approximately 280 g is analyzed by a chemical procedure using thin layer chromatography (TLC), Waliking, et al. [1968]. To facilitate an adequate quality control and consumer protection program, it is desirable to design a sampling plan that will provide a high level of protection for the consumer with reasonable assurance to the processor that lots of good peanuts will not be rejected by the testing program.

Because aflatoxin is often highly concentrated in a small percentage of the kernels, variation among aflatoxin determinations is large, and accurate estimates of the average concentration in a lot is exceedingly difficult, Whitaker, et al. [1972, 1974]. In order to predict the risk levels and costs associated with an aflatoxin sampling plan, computer simulation methods have been developed to evaluate sampling programs, Whitaker, et al. [1970].

One important aspect of the computer model is estimation of the distribution of analytical results about the subsample concentra-

tion. It is assumed that replicated analyses from subsamples with different levels of aflatoxin can all be described by distributions of the same functional form, but, possibly, with different parameters.

The objectives of the present study are to answer the questions:

- (a) Does the normal or the lognormal model give a better fit for the numbers obtained by replicated analyses of peanut subsamples?
- (b) Does either the normal or the lognormal model provide an adequate fit for this data?

## 2.2 The Statistical Problem

Suppose there are available  $k$  independent samples:

$$\begin{array}{cccc} x_{11} & , & x_{12} & , \dots , x_{1n_1} ; \\ x_{21} & , & x_{22} & , \dots , x_{2n_2} ; \\ & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot \\ x_{k1} & , & x_{k2} & , \dots , x_{kn_k} . \end{array} \quad (1)$$

The random variables in (1) will be said to satisfy a normal model if all members of the  $j$ th sample are independently distributed with a  $N(\mu_j, \sigma_j^2)$  distribution for  $j = 1, \dots, k$ . They will be said to satisfy a lognormal model if all members of the  $j$ th sample are independently distributed with a  $LN(\mu_j, \sigma_j^2)$  distribution for  $j = 1, \dots, k$ . (A random variable  $y$  has a  $LN(\mu, \sigma^2)$  distribution if  $\ln y$  has a  $N(\mu, \sigma^2)$  distribution.) In order to answer questions (a) and (b) above for a particular set of data of the structure of (1), we require  $k$ -sample composite goodness-of-fit and model comparison techniques.

There is a small statistical literature that considers these types of problems. Petrov [1956] considers this model for normal

distributions and samples of the same size; and gives some distributional results. Prohorov [1966] mentions but does not develop the type of transformation-pooling approach that will be used here.

The work of a number of writers who have considered the problem of transforming a single sample under a composite null hypothesis assumption is of some interest here. These include papers by David and Johnson [1948], Sarkadi [1960, 1965], Durbin [1961], Störmer [1964], Seshandri, et al. [1969], and O'Reilly and Quesenberry [1973].

## 2. ANALYSIS OF AFLATOXIN DATA

### 2.1 Experimental Procedure

Approximately 6200 g of raw peanut kernels contaminated with aflatoxin were comminuted in a mill similar to that used by the inspection service. The ground meal was then divided into eleven subsamples weighing approximately 560 g each. Each subsample of comminuted peanuts was blended with 2800 ml methanol-water-hexane solution for 2 minutes. The blended material was divided equally among 16 centrifuge bottles. The content of each centrifuge bottle was analyzed by a modified version of the Walkking method, Walkking, et al., [1968]. One observation was lost from each of 3 subsamples leaving 8 subsamples with 16 determinations and 3 subsamples with 15 determinations. The data is shown in Table 1.

-----  
Table 1 near here  
-----

TABLE 1

REPLICATED AFLATOXIN DETERMINATIONS

## SAMPLE NUMBER

1	2	3	4	5	6	7	8	9	10	11
121.23	95.33	20.04	22.35	30.02	10.84	28.60	34.30	16.71	7.63	52.56
71.69	55.94	20.30	32.54	26.26	19.31	32.59	27.90	18.11	6.37	112.86
117.91	72.01	24.69	23.23	26.48	13.12	37.24	34.87	11.19	6.97	69.89
91.09	58.96	22.26	68.31	36.38	13.20	24.61	32.56	11.98	5.28	100.15
104.86	114.62	24.92	27.51	48.01	12.39	24.71	31.67	24.86	6.74	87.14
151.00	41.64	21.45	27.44	49.50	17.20	35.83	29.06	20.18	6.74	33.42
125.40	98.76	19.44	28.53	15.68	12.61	48.93	32.00	17.49	12.06	82.88
83.94	53.62	24.04	29.49	30.79	18.07	37.99	33.40	9.25	7.53	64.63
137.53	90.23	24.40	51.98	22.03	18.19	29.19	30.55	15.35	6.81	73.98
83.49	91.92	19.85	28.87	26.54	16.65	28.89	31.87	21.12	12.58	112.11
116.78	66.88	11.88	21.80	22.74	15.85	32.46	29.17	17.23	12.64	97.59
90.72	91.81	24.34	29.12	35.36	13.53	40.31	26.05	18.67	10.71	65.36
100.54	100.37	15.32	36.65	51.61	11.06	35.88	36.50	23.17	5.41	32.19
74.59	77.26	14.85	40.76	27.53	15.25	31.04	27.51	17.05	7.46	94.55
137.19	91.56	23.10	23.81	36.71	16.25	32.05	30.67	15.43	4.39	60.24
146.25	66.25	22.21	36.19		12.39		31.76	16.82	11.85	

## 2.2 Statistical Analysis

To simplify notation we shall first consider a single sample  $x_1, \dots, x_n$ . Let

$$\bar{x}_r = (x_1 + \dots + x_r)/r,$$

$$s_r^2 = (1/r) \sum_{i=1}^r (x_i - \bar{x}_r)^2,$$

for  $r = 3, 4, \dots, n$ ; and  $G_\nu$  denote the distribution function of a Student-t distribution with  $\nu$  degrees of freedom. Then put

$$u_{r-2} = G_{r-2} \left\{ (r-2)^{1/2} (x_r - \bar{x}_r) / [(r-1)s_r^2 - (x_r - \bar{x}_r)^2]^{1/2} \right\}, \quad (2)$$

for  $r = 3, \dots, n$ .

It follows from O'Reilly and Quesenberry [1973], Corollary 2.1 and example 4.1, that if  $x_1, \dots, x_n$  is a random sample from a member of the  $N(\mu, \sigma^2)$  family then  $u_1, \dots, u_{n-2}$  are independently and identically distributed as uniform random variables on the unit interval (0,1)--i.i.d.  $U(0,1)$  rv's.

Let each sample in (1) be transformed using the transformations in (2). Then from the  $j$ th sample ( $j = 1, \dots, k$ ) we obtain  $n_j - 2$  i.i.d.  $U(0,1)$  random variables, and a total of  $N \equiv n_1 + \dots + n_k - 2k$  pooled i.i.d.  $U(0,1)$  random variables. Let  $u_{(1)} \leq u_{(2)} \leq \dots \leq u_{(N)}$  denote the ordered values of this pooled set of random variables. Under the normal model hypothesis then  $E(u_{(j)}) = j/(N+1)$  for  $j = 1, \dots, N$ ; and if the pairs  $(u_{(j)}, j/(N+1))$  are plotted on Cartesian axes the points should approximate the line  $g(u) = u$  for  $0 < u < 1$ . This gives a graphical procedure for judging

whether the parent distribution is normal. (See Epstein [1960], section 1, for a similar graphical procedure.)

As mentioned earlier in subsection 2.1, Table 1 gives the data obtained from repeated determinations of aflatoxin from 11 subsamples. The physical interpretation of the parent distribution of interest is that it is the distribution of analytical "errors" of the determination process.

The transformations of (2) have been performed on each of the samples in Table 1. The u-values obtained for each sample have been pooled and ranked and are given in Table 2.

-----  
Table 2 near here  
-----

The 151 values in Table 2 have been used to plot the 151 points  $(u_{(j)}, j/(N+1))$  in Figure 1.

-----  
Figure 1 near here  
-----

In addition to this graphical procedure, we compute two goodness-of-fit statistics as measures of the uniformity of the u-data of Table 2. We compute the Pearson  $\chi^2$  statistic using 10 cells of equal probability, and the Watson  $U^2$  statistic, Watson [1961].

To compute the chi-squared statistic the interval (0,1) is partitioned into ten intervals of length 0.1 each and the vector of cell frequencies from the data of Table 2 is (12, 14, 10, 20, 19, 15, 16, 13, 15, 17). This gives a chi-squared value of



TABLE 2

POOLED AND RANKED NORMAL U - DATA

.00082	.13930	.29274	.36630	.46620	.56100	.69103	.80297	.92663
.01938	.15006	.29946	.36659	.46746	.56220	.69229	.80575	.92823
.06128	.16654	.30140	.37203	.47610	.56773	.69461	.81878	.92964
.06530	.17336	.30337	.38696	.47701	.57158	.69685	.82131	.94233
.06809	.17461	.31949	.39961	.48015	.58405	.70536	.82447	.95420
.07161	.19171	.33819	.40249	.48427	.60007	.70923	.82545	.96132
.07785	.19340	.34183	.41576	.49125	.60922	.71191	.83398	.96301
.08120	.19548	.34335	.41709	.49125	.61159	.71418	.83989	.96326
.08777	.19958	.34465	.41993	.50913	.62691	.71461	.84354	.96553
.08790	.22247	.34954	.42337	.51608	.63754	.72050	.85736	.97956
.09628	.22411	.34994	.42405	.52162	.64758	.72308	.85575	.98416
.09995	.23680	.35816	.42552	.52176	.65618	.73236	.86864	.98556
.10666	.23982	.35827	.42885	.53436	.65654	.73270	.87260	.98551
.12083	.24195	.35924	.43051	.53895	.66946	.73510	.8726	.99032
.12112	.25510	.35943	.43872	.54633	.68423	.75552	.89042	.99934
.13489	.26693	.36443	.43943	.55576	.68819	.77980	.90173	
.13601	.28429	.36591	.46419	.55815	.69003	.78882	.90852	

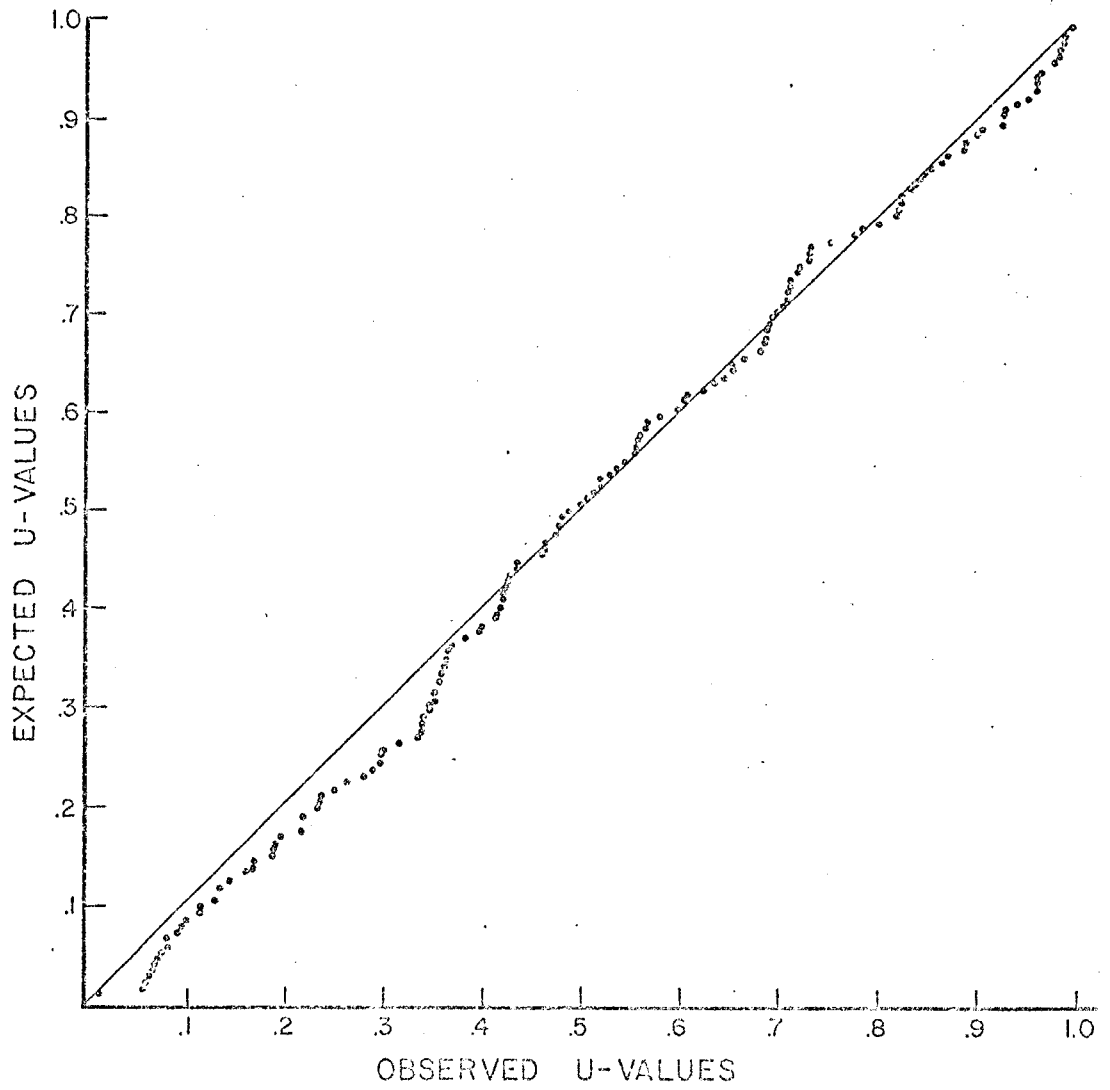


FIGURE 1

Ranked u-values Against Expectations, Normal Model

$$\begin{aligned} \chi^2 &= (k/N) \sum_{j=1}^k N_j^2 - N , \\ &= (10/151)(2365) - 151 , \\ &= 5.62 . \end{aligned}$$

The Watson  $U^2$  statistic has been shown by Miller and Quesenberry [1975] to have attractive power properties as a test of uniformity and is given by

$$U^2 = (1/12N) + \sum_{j=1}^N [u_{(j)} - (2j-1)/2N]^2 - N(\bar{u}-1/2)^2 ,$$

where  $\bar{u} = (1/N)(u_{(1)} + \dots + u_{(N)})$ . The value of  $U^2$  for the data of Table 2 is 0.0785.

To study the fit of the lognormal model to the data of Table 1, we proceeded as follows. First, the natural logarithm of each number in Table 1 was taken, and these logarithms were then analyzed for normality just as described above. Figure 2 is a graph of the pooled and ranked u-values obtained from these logarithms using equation (2), plotted against their null hypothesis expected values. The Pearson  $\chi^2$  and the Watson  $U^2$  statistics were found to have values  $\chi^2 = 12.64$  and  $U^2 = 0.1089$  for this case.

- - - - -  
Figure 2 near here  
- - - - -

The values of  $\chi^2$  and  $U^2$  for both the normal and lognormal models are given in Table 3. The value in parentheses is the significance level of the observed value.

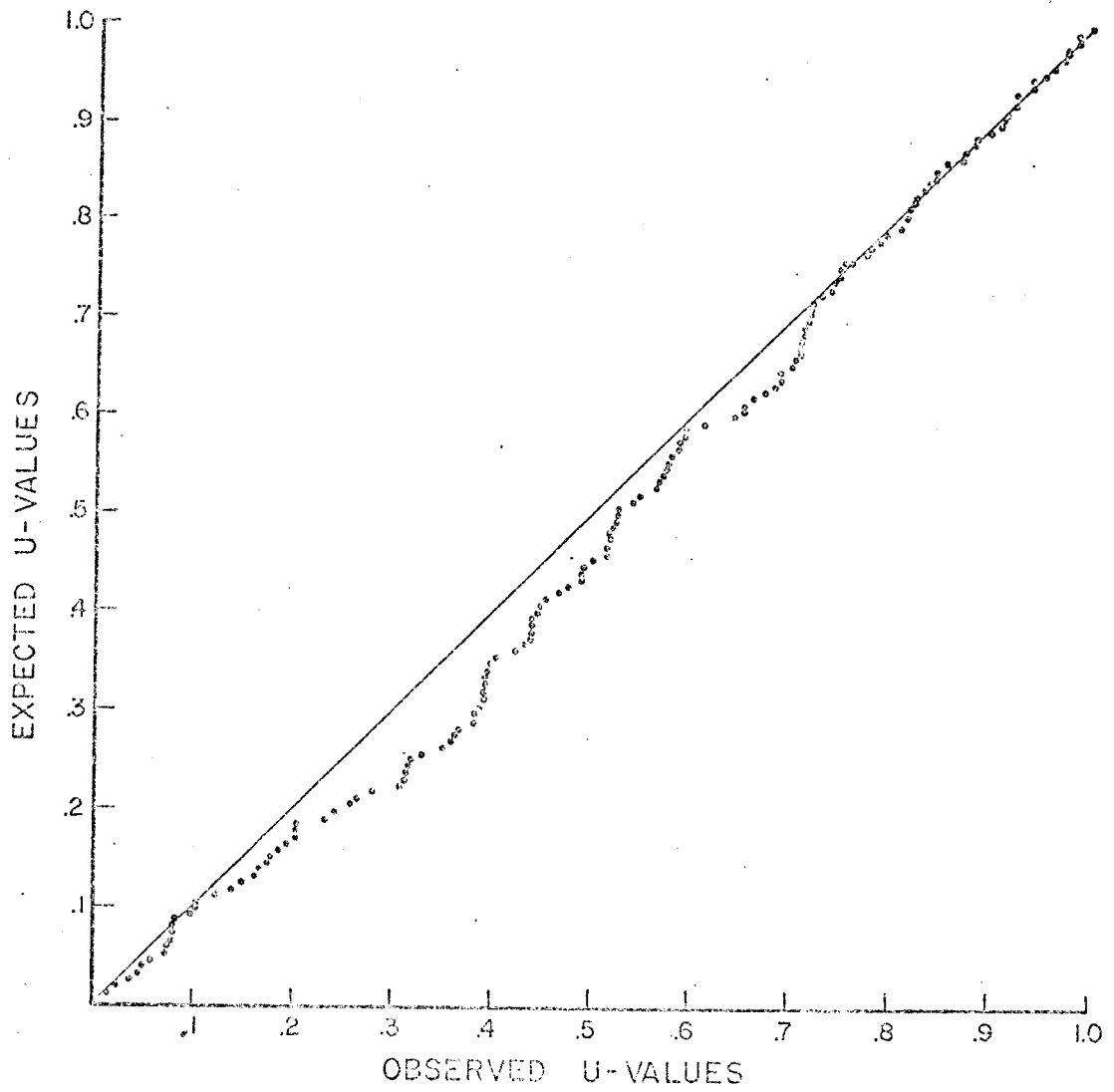


FIGURE 2

Ranked u-values Against Expectations, Lognormal Model

For the  $\chi^2$  values, this is the probability that a chi-squared random variable with 9 d.f. is less than the observed value. The probabilities for  $U^2$  have been evaluated using equation (22) in Watson [1961].

TABLE 3  
OBSERVED VALUES OF TEST STATISTICS FOR  
NORMAL AND LOGNORMAL MODELS

Statistic	Normal	Lognormal
$\chi^2$	5.62 (.78)	12.64 (.18)
$U^2$	.0785 (.421)	.1089 (.237)

### 2.3 Analysis Summary

Examination of Figures 1 and 2 shows that the normal model fits quite well, and that the lognormal model is not as good. Exactly the same conclusion is reached by examination of Table 3. Both of the test statistics are smaller for the normal model than for the lognormal model and so the normal must be preferred. While the values of  $\chi^2$  and  $U^2$  are not significant at the usual levels (.1, .05, etc.), the values for the lognormal model are sufficiently large to cause suspicion. We feel that this analysis supports the following responses to questions (a) and (b) of section 1.1:

- (a) the normal model fits the data better than the lognormal model, and
- (b) the fit of the normal model is quite satisfactory.

3. SOME REMARKS CONCERNING THE  
STATISTICAL PROCEDURE

The approach to model testing and goodness-of-fit used here is to transform each individual sample, and then to pool the transformed sets. This approach may also be useful for testing models other than the normal and lognormal.

We call attention to the fact that the transformations of equation (2) are not symmetric in the observations. It is necessary that the values  $x_1, \dots, x_n$  as they are used in (2) be i.i.d. random variables, and must, therefore, be in random order. They must not, for example, be arranged according to magnitude.

The computational work described in the previous sections, including the transformations in (2), the pooling and ranking of the u-values, and the computation of test statistics have been carried out using a Fortran program written by the authors for this purpose. This work was performed at the Triangle Universities Computation Center (TUCC).

The authors are grateful to Professor F. G. Giesbrecht for helpful discussions.

REFERENCES

- Bampton, S. C. [1963]. Growth of *Aspergillus flavus* and Production of Aflatoxin in Groundnuts - Part I. Tropical Science 5, 74-81.
- David, F. N. and N. L. Johnson [1948]. The probability integral transformation when parameters are estimated from the sample. Biometrika 35, 182-192.
- Dickens, J. W. and J. B. Satterwhite [1969]. Subsampling Mill for Peanut Kernels. Food Technology 23, 90-92.

- Duggan, R. E. [1970]. Controlling Aflatoxin. FDA Papers, April, 13-18.
- Durbin, J. [1961]. Some methods of constructing exact tests. Biometrika 40, 41-55.
- Epstein, B. [1960]. Tests for the validity of the assumption that the underlying distribution of life is exponential: Part I. Technometrics 2, 83-101.
- Gibbons, J. D. and J. W. Pratt [1975]. P-values: interpretation and methodology. American Statistician 29, No. 1, 20-25.
- Miller, F. L., Jr. and C. P. Quesenberry [1975]. Statistics for testing uniformity on the unit interval. Technical Report UNCCNC-CSD-12, Union Carbide Corporation, Nuclear Division.
- O'Reilly, F. and C. P. Quesenberry [1973]. The conditional probability integral transformation and applications to obtain composite chi-square goodness of fit tests. Annals of Statistics 1, 1-10.
- Petrov, A. A. [1956]. Verification of statistical hypotheses on the type of a distribution based on small samples. Theory of Probability and Its Applications 1, 223-245.
- Prohorov, Yu [1966]. Some characterization problems in statistics. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability I, 341-349.
- Sarkadi, K. [1960]. On testing for normality. Publications of the Mathematical Institute of the Hungarian Academy of Sciences 3, 269-275.
- Sarkadi, K. [1965]. On testing for normality. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 373-387.
- Seshadri, V., M. Csorgo and M. A. Stephens [1969]. Tests for the exponential distribution using Kolmogorov-type statistics. Journal of the Royal Statistical Society, Series B 31, 499-509.
- Stephens, M. A. [1974]. EDF statistics for goodness of fit and some comparisons. Journal of the American Statistical Association 69, 730-737.
- Störmer, Horand [1964]. Ein test zum erkennen Normalverteilungen. Zeitschrift Wahrscheinlichkeitstheorie und Verwandte Gebiete 2, 420-428.
- Watson, G. S. [1961]. Goodness-of-fit tests on a circle I. Biometrika 48, 109-114.

- Waltking, A. E., G. Bleffert and M. Kiernan [1968]. An improved rapid physiochemical assay method for aflatoxin in peanuts and peanut products. Journal of the American Oil Chemists' Society 45, 880-883.
- Whitaker, T. B., J. W. Dickens and R. J. Monroe [1974]. Variability of aflatoxin test results. Journal of the American Oil Chemists' Society 49, 590-593.
- Whitaker, T. B., J. W. Dickens, R. J. Monroe and E. H. Wiser [1972]. Comparison of the observed distribution of aflatoxin in shelled peanuts to the negative binomial distribution. Journal of the American Oil Chemists' Society 49, 590-593.
- Whitaker, T. B., J. W. Dickens and E. H. Wiser [1970]. Design and analysis of sampling plans to estimate aflatoxin concentrations in shelled peanuts. Journal of the American Oil Chemists' Society 47, 501-504.