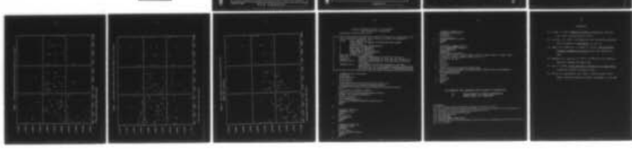
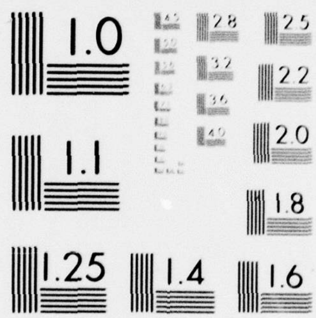
AD-A059 935 ROCHESTER UNIV NY DEPT OF STATISTICS F/G 12/1
ASYMPTOTIC BEHAVIOR OF THE Z-TEST FOR NORMALITY.(U)
1978 C LIN, G S MUDHOLKAR AFOSR-77-3360

UNCLASSIFIED AFOSR-TR-78-1356 NL

1 OF 1
AD
A059935



END
DATE
FILMED
12-78
DDC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR-78-1356	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ASYMPTOTIC BEHAVIOR OF THE Z-TEST FOR NORMALITY		5. TYPE OF REPORT & PERIOD COVERED Interim rept.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Ching-Chung Lin and Govind S. Mudholkar		8. CONTRACT OR GRANT NUMBER(s) AFOSR-77-3360
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Rochester, Department of Statistics, Rochester, New York 14626		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM, Bolling AFB, Washington, DC 20332		12. REPORT DATE 1978
		13. NUMBER OF PAGES 16
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Motivated by the independence of the mean and the variance of a sample from a normal population, Mudholkar and Lin (4) proposed a statistic Z for testing the composite hypothesis of normality. This test is based on the Fisher transform of the product moment correlation coefficient r between X_i and Y_i, where X_i, i=1, 2,...,n is the random sample and Y_i is the Wilson-Hilferty transform of the sample variance with the i^{th} observation delected

DD FORM 1 JAN 73 1473 UNCLASSIFIED

20. Abstract continued

from the sample. In this note we show that r and Z both are asymptotically normally distributed. It is then demonstrated that the Z-test of normality is consistent against any asymmetric alternative. The large sample null distribution of Z is used to reconstruct the expression for the finite sample estimate of the variance of Z which agrees with its asymptotic value. Fortran and APL routines for computing the statistic and its P-value are given.

ASYMPTOTIC BEHAVIOR OF THE Z-TEST FOR NORMALITY*

by

Ching-Chuong Lin** and Govind S. Mudholkar

University of Maryland Baltimore County and University of Rochester

ABSTRACT

Motivated by the independence of the mean and the variance of a sample from a normal population, Mudholkar and Lin [4] proposed a statistic Z for testing the composite hypothesis of normality. This test is based on the Fisher transform of the product moment correlation coefficient r between X_i and Y_i, where X_i, $i=1, 2, \ldots, n$ is the random sample and Y_i is the Wilson-Hilferty transform of the sample variance with the i^{th} observation deleted from the sample. In this note we show that r and Z both are asymptotically normally distributed. It is then demonstrated that the Z-test of normality is consistent against any asymmetric alternative. The large sample null distribution of Z is used to reconstruct the expression for the finite sample estimate of the variance of Z which agrees with its asymptotic value. Fortran and APL routines for computing the statistic and its P-value are given.

*Research sponsored in part by the Air Force Office of Scientific Research, Air Force Systems Command, USAF under Grant No. AFOSR-77-3360.

**Research aupported in part by University of Maryland Baltimore County Faculty Summer Research Fellowship.

ASYMPTOTIC BEHAVIOR OF THE Z-TEST FOR NORMALITY

1. INTRODUCTION

Let $X_1, X_2, \ldots, X_n$ be n independent identically distributed random variables with distribution function $F(\cdot)$, and consider the composite hypothesis

$$H_0: \ F(x) = \Phi\left(\frac{X-\mu}{\sigma}\right), \ -\infty < \mu < \infty, \ \sigma > 0, \qquad (1.1)$$

where $\Phi(\cdot)$ is the standard normal distribution. Mudholkar and Lin [4], motivated by the independence of the mean and the variance of a sample from a normal population, proposed using the Fisher transform Z of the product moment correlation coefficient

$$r = \frac{\sum_{i=1}^{n} (X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n} (X_i-\bar{X})^2} \ \sqrt{\sum_{i=1}^{n} (Y_i-\bar{Y})^2}} \qquad (1.2)$$

between X_i and Y_i, where $Y_i = \{[\sum_{j\neq i} X_j^2 - (\sum_{j\neq i} X_j)^2/(n-1)]/n\}^{\frac{1}{3}}$ is the Wilson-Hilferty transform [6] of the sample variance with the i^{th} observation deleted from the sample, as a statistic for testing normality. They show that even in small samples the null distribution of Z is near normal with mean zero. A Monte Carlo experiment was conducted to obtain the variance and the excess of Kurtosis of Z for various values of the sample size n. From these using weighted regression, it was proposed that we may take

$$\sigma_n = \hat{\sigma}_n(Z) = (.591730 + .143559\,n - .002235\,n^2 + .000016\,n^3)^{-1}, \qquad (1.3)$$

$$\gamma_{2,n} = \hat{\gamma}_{2,n}(Z) = -11.697157/n + 55.059097/n^2, \qquad (1.4)$$

as the finite sample estimates of the standard deviation and the excess of Kurtosis of Z, respectively. Even though (1.3) gives a good estimate for $\sigma_n(Z)$ over the range of the simulation, it suggests that asymptotically it is $O(1/n^3)$. In this note, we show that r and Z both are asymptotically normally distributed. It is then demonstrated that the Z-test of normality is consistent against any asymmetric alternative. The large sample null distribution of Z is used to reconstruct the expression for σ_n which agrees with its asymptotic value. Fortran and APL routines for computing the statistic and its P-value are given in Appendix.

2. THE LARGE SAMPLE THEORY FOR THE Z-TEST

Consider the problem of testing the composite hypothesis that a sample $X_1, X_2, \ldots, X_n$ of size n is from a normal population. It is well known, (e.g., Cramér (1946), and Kagan, Linnik and Rao (1973)), that the hypothesis is true if and only if the sample mean $\overline{X}$ and the sample variance S^2 are independently distributed. Thus a test for the independence of $\overline{X}$ and S^2 is also a goodness-of-fit test for the normality. The apparent limitation that we have only one $(\overline{X}, S^2)$ for testing the independence of the pair can be circumvented in many ways. The most convenient of these is to consider the n means and the corresponding variances computed from the n samples obtained by deleting one observation at a time. Even though the n pairs so obtained are not independent, they can be used to estimate the extent of the dependence between the mean and the variance of samples from the population. Because of its simplicity, we wish to use the product moment correlation coefficient as a measure of the dependence.

Now the test based upon a product moment correlation coefficient is appropriate for testing independence in a bivariate normal distribution. However, one of the marginal distributions in the present case, the distribution of variance, is nonnormal even if the parent population is normal. This can be remedied to a considerable extent by applying the famous Wilson and Hilferty (1931) cube root transformation to the n variances in order approximately to normalize them. Because of its invariance with respect to changes in scale and origin, the correlation coefficient between the n sample means and the cube roots of the corresponding sample variances mentioned above, equals the correlation coefficient $r = \mathrm{corr}(X_i, Y_i)$, where

$$Y_i = \{\frac{1}{n}[\sum_{j\neq i} X_j^2 - (\sum_{j\neq i} X_j)^2 / (n-1)]\}^{1/3} \quad i = 1,2,\ldots,n . \qquad (2.1)$$

Abnormally large values of r^2 would cast doubt on the normality of the X's.

<u>Asymptotic Distribution of r</u>. Without any loss of generality, suppose that $X_1, X_2, \ldots, X_n$ are independent identically distributed random variables with $EX_1 = 0$, $EX_1^2 = 1$ and denote $\mu_k = EX_1^k$, $k = 3,4,5,6$. In order to understand the large sample distribution of the coefficient

$$r = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n} (X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}} , \qquad (2.2)$$

We note that as $n \to \infty$ $\frac{1}{n}\sum_{i=1}^{n} (X_1 - \overline{X})^2 \to 1$ in probability and examine the large sample behavior of the other two terms.

Lemma 2.1. As $n \to \infty$

$$\text{(a)} \quad Y_i - \bar{Y} = S^{2/3}\left[\frac{1}{3(n-1)}\left(1 - \frac{(X_i - \bar{X})^2}{S^2}\right) + O_p\left(\frac{1}{n^2}\right)\right] \tag{2.3}$$

for $i = 1,2,\ldots,n$ and

$$\text{(b)} \quad n\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = S^{4/3}\left\{\frac{n^2}{9(n-1)^2}\left[\frac{m_4}{S^4} - 1\right] + O_p\left(\frac{1}{n}\right)\right\}, \tag{2.4}$$

where $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ and $m_4 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^4$.

(a) Proof. The variable Y_i defined in (2.1) can be expressed as

$$Y_i = \left[\frac{1}{n}\left(\sum_{j=1}^{n}(X_j - \bar{X})^2 - \frac{n}{n-1}(X_i - \bar{X})^2\right)\right]^{1/3}$$

$$= S^{2/3}\left[1 - \frac{1}{(n-1)}\frac{(X_i - \bar{X})^2}{S^2}\right]^{1/3}$$

$$= S^{2/3}\left[1 - \frac{1}{3(n-1)}\frac{(X_i - \bar{X})^2}{S^2} + O_p\left(\frac{1}{n^2}\right)\right] \tag{2.5}$$

Hence $\bar{Y} = \sum_{i=1}^{n} Y_i/n = S^{2/3}[1 - (3(n-1))^{-1} + O_p(1/n^2)]$ and

$$Y_i - \bar{Y} = S^{2/3}[(3(n-1))^{-1}(1 - (X_i - \bar{X})^2 / S^2) + O_p(1/n^2)] .$$

(b) From (a) we have

$$n\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = n\,S^{4/3}\left\{\frac{1}{9(n-1)^2}\sum_{i=1}^{n}\left(1 - \frac{(X_i - \bar{X})^2}{S^2}\right)^2 + O_p\left(\frac{1}{n^2}\right)\right\}$$

$$= n\,S^{4/3}\left\{\frac{1}{9(n-1)^2}\left[n - 2n + \frac{N\cdot m_4}{S^4}\right] + O_p\left(\frac{1}{n^2}\right)\right\}$$

$$= S^{4/3}\left\{\frac{n^2}{9(n-1)^2}\left[\frac{m_4}{S^4} - 1\right] + O_p\left(\frac{1}{n}\right)\right\} .$$

Lemma 2.2. As $n \to \infty$

$$\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y}) = - \frac{n}{3(n-1)} S^{-4/3} \cdot m_3 + O_p(\frac{}{n}) , \qquad (2.6)$$

where $m_3 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^3$.

Proof. From Lemma 2.1 (a) we have

$$\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y}) = S^{2/3} \sum_{i=1}^{n} (X_i - \overline{X})(\frac{1}{3(n-1)} (1 - \frac{(X_i-\overline{X})^2}{S^2}) + O_p (\frac{1}{n^2}))$$

$$= S^{2/3}(\frac{1}{3(n-1)} \sum_{i=1}^{n} (X_i - \overline{X})(1 - \frac{(X_i-\overline{X})^2}{S^2}) + O_p (\frac{1}{n})$$

$$= \frac{1}{3(n-1)} S^{2/3}(- \frac{\sum_{i=1}^{n} (X_i-\overline{X})^3}{S^2}) + O_p (\frac{1}{n})$$

$$= - \frac{n}{3(n-1)} S^{-4/3} m_3 + O_p (\frac{1}{n}) .$$

Theorem 2.3. If $\mu_k = EX^k < \infty$, for $1 \leq k \leq 6$, then as $n \to \infty$ the correlation coefficient r in (2.2) is asymptotically normally distributed with mean

$$E(r) = - \frac{\mu_3}{\sqrt{\mu_4 - 1}} = - \frac{\beta_1}{\sqrt{\beta_2 - 1}} \qquad (2.7)$$

and variance

$$Var(r) = \frac{\mu_6 - 6\mu 4 - \mu_3^2 + 9}{(\mu_4 - 1)n} = (\beta_3 - \frac{\beta_1^2}{\beta_2 - 1} + \frac{3}{\beta_2 - 1} - 6) / n , \qquad (2.8)$$

where $\beta_3 = \mu_6/(\mu_4 - 1)$, β_1 and β_2 are respectively the coefficients of skewness and kurtosis.

Proof. Let $T_n = (-3 S^{4/3}((n-1)/n)) \sum_{i=1}^{n} (X_i-\overline{X})(Y_i-\overline{Y})$. Then from Lemma 2.2, we see that $T_n = m_3 + O_p(1/n)$. It is well known (see, for example, Cramér (1946) p. 365) that the third central moment m_3 is asymptotically normal with mean μ_3 and variance $(\mu_6 - 6\mu_4 - \mu_3^2 + 9)/n$. Hence T_n is asymptotically normal with the same mean and variance as m_3. Now

$$r = T_n[-3\, S^{4/3}(\frac{n-1}{n})\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2}\sqrt{n\sum_{i=1}^{n}(Y_i-\overline{Y})^2}]^{-1} . \qquad (2.9)$$

Hency by Slutsky's theorem and Lemma 2.1 (b), r is asymptotically normal with mean $-\mu_3 / \sqrt{\mu_4-1} = -\beta_1 / \sqrt{\beta_2-1}$ and variance $(\mu_6 - 6\mu_4 - \mu_3^2 + 9)/(n(\mu_4-1) = (\beta_3 - \beta_1^2 / (\beta_2-1) + 3/(\beta_2-1) - 6)/n$.

Corollary 2.4. If $X_1, X_2, \ldots, X_n$ are independent and identical normal random variables then r is asymptotically normally distributed with mean 0 and variance 3/n.

Theorem 2.5. (Consistency). The test based on the correlation coefficient r is consistent against all asymmetric alternatives with first four moments finite.

Proof. In view of Lemma 2.1 and Lemma 2.2, $\sqrt{n\sum_{i=1}^{n}(Y_i-\overline{Y})^2}$ and $\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})$ converge in probability to $\sqrt{\mu_4-1} / 3$ and $-\mu_3/3$, respectively, as $n \to \infty$. Consequently the correlation coefficient r converges in probability to $-\mu_3 / \sqrt{\mu_4-1}$. Because the critical constant of the test converges to 0 as $n \to \infty$, the test based upon r is consistent for alternatives with $\mu_3 \neq 0$.

Remark 2.6. From the asymptotic distribution of r given in Theorem 2.3, it can be seen that the test is asymptotically unbiased against symmetric alternatives provided $\beta_3 - \beta_1^2 / (\beta_2-1) + 3/(\beta_2-1) - 6 > 3$. It also follows that the test is not consistent against symmetric alternatives.

Asymptotic Distribution of Z. It is well known that in samples of moderate size from a bivariate normal population the product moment correlation coefficient is very nonnormal, but the distribution of its Fisher transform Z is practically normal for remarkably small values of the sample size. In view of the robust nature of this phenomenon, and marginally near normal distributions of X_i and Y_i, the observation may be expected to remain valid in case of the statistic r defined in (1.2) especially when the null hypothesis H_o is true. Hence, we propose using

$$Z = \frac{1}{2} \log \frac{1+r}{1-r} \tag{2.10}$$

as the test statistic and rejecting the normality of the X's if $|Z| >$ constant.

Theorem 3. Under the conditions of Theorem 1, as $n \to \infty$ the statistic Z is asymptotically normal with mean

$$E(Z) = \frac{1}{2} \log \left(\frac{\sqrt{\beta_2-1} - \beta_1}{\sqrt{\beta_2-1} + \beta_1} \right) \tag{2.11}$$

and variance

$$\mathrm{Var}(Z) = \left(\frac{\beta_2-1}{\beta_2-\beta_1^2-1} \right)^2 \left(\beta_3 - \frac{\beta_1^2}{\beta_2-1} + \frac{3}{\beta_2-1} - 6 \right) / n \tag{2.12}$$

Proof. The result follows from Theorem 2.3 and the well known convergence theorem due to Mann-Wald (see, Rao (1973) p. 385), which states that $g(T)$ is asymptotically $N(g(\theta); (g'(\theta)\,\sigma(T))^2)$ if T is asymptotically $N(\theta, \sigma^2(T))$. In the present case $T = r$, $g(T) = Z$, and $g'(\theta) = 1/(1-\theta^2)$, where $\theta = E(r) = -\beta_1/\sqrt{\beta_2-1}$.

Corollary 2. If $X_1, X_2, \ldots, X_n$ are independent and identical normal rnadom variables, then Z is asymptotically normally distributed with mean 0 and variance $3/n$.

3. AN ADJUSTMENT OF VARIANCE OF THE NULL DISTRIBUTION IN SMALL SAMPLES AND A MONTE CARLO POWER STUDY

Mudholkar and Lin [4], using an extensive Monte Carlo study for the null distribution of Z in small samples, have demonstrated that the statistic Z can be regarded as practically normally distributed with mean zero and standard deviation σ_n expressed as in (1.3). This expression of σ_n is inappropriate when the sample size n is large because it is $O(1/n^3)$ asymptotically. The large sample theory of the Z-test developed in Section 2 is now used to adjust the expression for σ_n so that it agrees with its asymptotic value. Toward this end a polynomial regression of $\sigma_n^2(Z)$, obtained in [3] by simulation for various values of n, is performed by taking the asymptotic variance $3/n$ as the leading term. The analysis yields

$$\sigma_n^2 = \hat{\sigma}_n^2(Z) = \frac{3}{n} - \frac{7.324}{n^2} + \frac{53.005}{n^3} . \tag{3.1}$$

A Monte Carlo experiment is conducted with a view to comparing the power of the Z-test with the powers of the other well known tests for normality. In this experiment, 1000 samples each of size $n = 20$ are obtained by simulation from a broad class of alternative distributions which includes both symmetric and asymmetric, light- and heavy-tailed distributions. The Z-test using the new formula for σ_n in (3.1) at 5% level of significance is performed on each of the samples inorder to estimate its power at these alternatives. The power values of the Z-test, along with the corresponding values for competing tests, are presented in Table 1. Some of the values of the power functions are taken from Filliben [2], and the remaining are estimated by us on the basis 1000 samples.

From the table it may be concluded that the Z-test is superior to or comparable with other tests against asymmetric distributions and is fair against heavy-tailed symmetric alternatives and poor in detecting light-tailed symmetric distributions.

Table 1

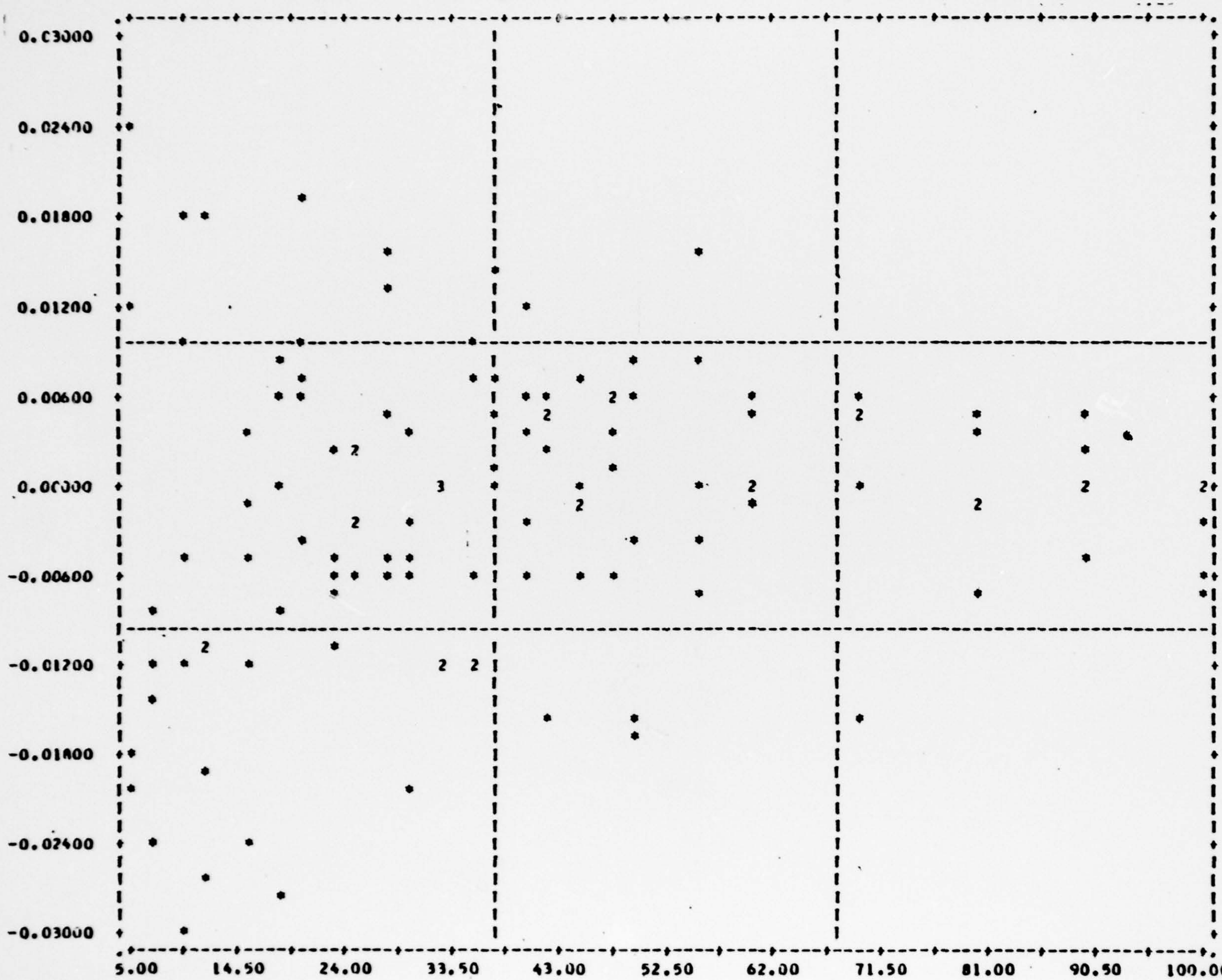
Power Functions of Some Tests for Normality

$n = 20,\ \alpha = 5\%$

Population	Skewness β_1	Kurtosis β_2	KS	CM	K_3	W	W'	R	b_1	Z
Uniform	0	1.8	.09	.14	.42	.15	.04	.04	.01	.04
Tukey (.25)	0	2.539	.05	.05	.06	.01	.01	.01	.02	.03
Logistic	0	4.2	.10	.13	.05	.07	.12	.12	.13	.12
Laplace	0	6.0	.23	.29	.10	.26	.33	.33	.24	.23
Cauchy	0	---	.84	.88	.74	.89	.91	.92	.79	.70
Weibull (10)	- .638	3.57	.12	.14	.10	.14	.15	.16	.18	.15
Exponential	2.0	9.0	.59	.73	.85	.86	.82	.82	.71	.83
Gamma (2)	1.414	6.0	.32	.41	.44	.50	.48	.50	.45	.54
Gamma (3)	1.155	5.0	.24	.32	.28	.33	.32	.33	.37	.43
Beta (2,1)	- .566	2.40	.15	.21	.43	.35	.18	.20	.10	.22

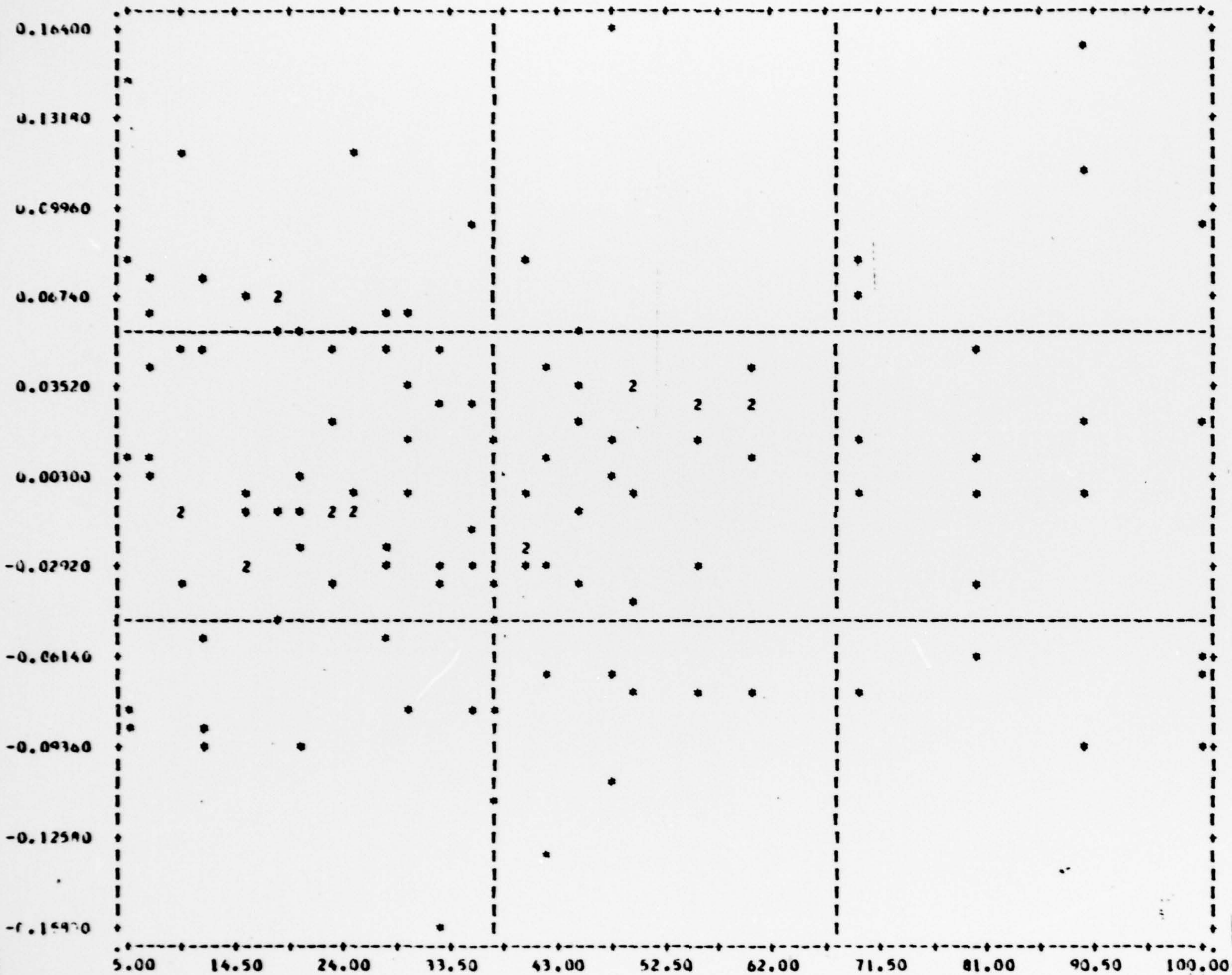
KS: Kolmogorov-Smirnov
CM: Cramér-von Mises
K_3: Vasicek
W : Shapiro-Wilk
W': Shapiro-Francia
R : Filliben
b_1: Sample Skewness
Z : New test

Figure 1 Scattergram of Mean (down) versus Sample Size N (across)



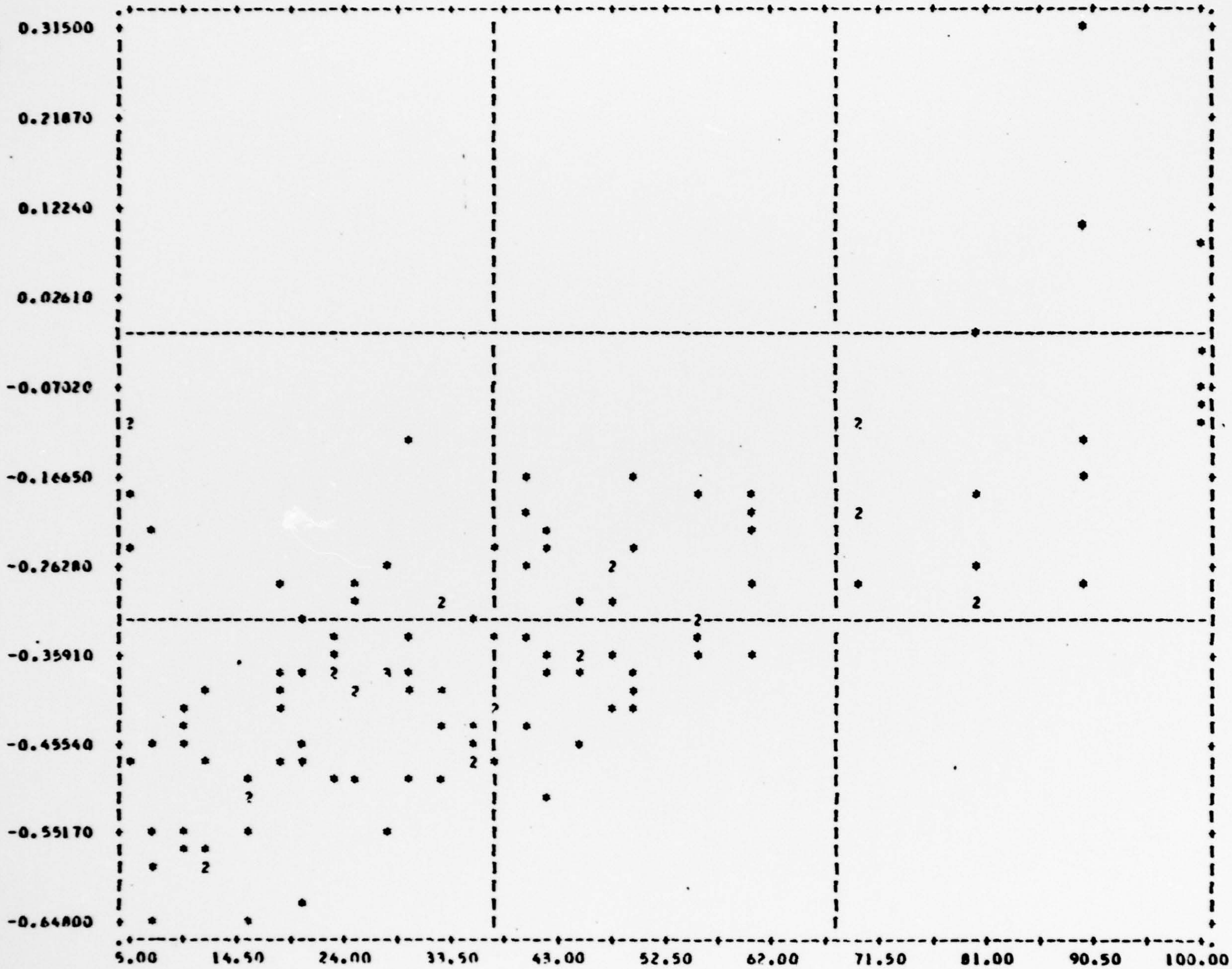
* Each point is based upon 1000 values.

Figure 2 Scattergram of Skewness (down) versus Sample Size N (across)



* Each point is based upon 1000 values.

Figure 3 Scattergram of Excess of Kurtosis (down) versus Sample Size N (across)



* Each point is based upon 1000 values.

APPENDIX:FORTRAN ROUTINE FOR COMPUTING
THE P-VALUE OF Z-STATISTIC

```
      SUBROUTINE ZTEST(X,N,U,V,Z,P)
C
C*****************************************************************************
C     THIS SUBROUTINE COMPUTES THE P-VALUE OF Z-TEST FOR NORMALITY
C     USAGE     CALL ZTEST(X,N,Z,P)
C     X         INPUT VECTOR OF LENGTH N CONTAINING THE SAMPLE
C               OBSERVATIONS
C     N         NUMBER OF SAMPLE OBSERVATIONS
C     U         WORK AREA OF LENGTH N
C     V         WORK AREA OF LENGTH N
C     Z         SAMPLE VALUE OF Z-STATISTIC
C     P         P-VALUE OF Z-STATISTIC
C     REQUIRED ROUTINES ε FUNCTIONS
C     SUM(X,N)          FUNCTION SUBPROGRAM TO GIVE THE SUM OF X
C     SUMSQ(X,N)        FUNCTION SUBPROGRAM TO GIVE THE SUM SQUARE OF X
C     CROSPR(X,Y,N)     FUNCTION SUBPROGRAM TO GIVE THE SUM OF PRODUCT
C                       OF X AND Y
C     NORMAL(X,Y,P)     SUBROUTINE TO GIVE THE ORDINATE Y AND THE
C                       PROBABILITY P OF THE STANDARD NORMAL DISTRIBUTION
C                       AT X AND MAY BE SUBSTITUTED BY COMPARABLE ROUTINE
C*****************************************************************************
C
      DIMENSION X(N),U(N),V(N)
      S=SUM(X,N)
      SSQ=SUMSQ(X,N)
      DO 10 I=1,N
      U(I)=S-X(I)
10    V(I)=(SSQ-X(I)*X(I)-U(I)*U(I)/(N-1))**(1./3.)
      CORR=(CROSPR(U,V,N)-SUM(U,N)*SUM(V,N)/N)/
     1((SUMSQ(U,N)-SUM(U,N)**2/N)*(SUMSQ(V,N)-SUM(V,N)**2/N))
     2**.5
      Z=0.5*ALOG((1+CORR)/(1-CORR))
      SIGMA=(3.0/N-7.324/N**2+53.005/N**3)**0.5
      GAMMA2=-11.697/N+55.059/N**2
      ZN=ABS(Z/SIGMA)
      CALL NORMAL(ZN,Y,P)
      P=2.-2.*(P-GAMMA2*(ZN*ZN*ZN-3*ZN)*Y/24.)
      RETURN
      END
C
C
      FUNCTION SUM(X,N)
      DIMENSION X(N)
      SUM=0.
      DO 10 I=1,N
10    SUM=SUM+X(I)
      RETURN
      END
C
      FUNCTION SUMSQ(X,N)
      DIMENSION X(N)
      SUMSQ=0.
      DO 10 I=1,N
10    SUMSQ=SUMSQ+X(I)*X(I)
      RETURN
      END
C
```

```
      FUNCTION CROSPR(X,Y,N)
      DIMENSION X(N),Y(N)
      CROSPR=0.
      DO 10 I=1,N
10    CROSPR=CROSPR+X(I)*Y(I)
      RETURN
      END
C
      SUBROUTINE NORMAL(X,Y,P)
      G=1.1283792*EXP(-(X*X/2.))
      Y=G/2.82842712
      XA=ABS(X)
      IF(XA.LT.2.5) GO TO 2
      U=1./(XA+1./(XA+2./(XA+3./(XA+4./(XA+5./(XA+6./(XA+7./(XA+8./(XA+9
     1./(XA+10./(XA+11./(XA+12./XA))))))))))))
      IF(X.GE.0) GO TO 1
      P=U*Y
      GO TO 4
 1    P=1.-U*Y
      GO TO 4
 2    ET=1.4142136/(1.4142136+0.3275911*XA)
      U=G*((((0.94064607*ET-1.28782245)*ET+1.2596513)*ET-0.252128668)*
     1ET+0.225836846)*ET
      IF(X.GE.0) GO TO 3
      P=U/2.
      GO TO 4
 3    P=1.-U/2.
 4    RETURN
      END
```

APL FUNCTION FOR COMPUTING THE P-VALUE OF Z-STATISTICS

X: INPUT VECTOR OF SAMPLE OBSERVATIONS
Z: SAMPLE VALUE OF Z-STATISTIC

```
∇ P←ZTEST X
 [1] V←(((+/X*2)-(X*2))-(U*2)÷N-1)*(1÷3)⋄ U←(+/X)-X⋄ N←⍴X
 [2] R←((+/U×V)-(+/U)×(+/V)÷N)÷(((+/U*2)-((+/U)*2)÷N)×(+/V*2)-((+/V)*2)÷N)*0.5
 [3] SIGMA←((3.0÷N)+(¯7.324÷N×N)+53.005÷N×N×N)*0.5
 [4] GAMMA2←(¯11.697÷N)+55.059÷N*2
 [5] Z←0.5×⍟(1+R)÷1-R
 [6] ZN←|Z÷SIGMA
 [7] ⍎')LOAD PUBLIC,STAT:NORM'
 [8] NPROB←(NORM ZN)-(1÷24)×GAMMA2×((ZN*3)-3×ZN)×(2.718281828*(-(ZN*2)÷2))÷(○2)*.5
 [9] P←2-2×NPROB
∇
```

REFERENCES

[1] Cramer, H. (1946), Mathematical Methods of Statistics, Princeton, New Jersey: Princeton University Press.

[2] Filliben, J.J. (1975), "The Probability Plot Correlation Coefficient Test for Normality", Technometrics, **17**, 111-117.

[3] Kagan, A.M., Linnik, Yu. V. and Rao C.R. (1973), Characterization Problems in Mathematical Statistics, New York: John Wiley & Sons.

[4] Mudholkar, G.S. and Lin, C.C. (1977), "A Simple Test for Normality", Manuscript submitted for publication.

[5] Rao, C.R. (1973), Linear Statistical Inference and its Applications, 2nd Edition, New York: John Wiley & Sons.

[6] Wilson, E.B. and Hilferty, M.M. (1931), "The Distribution of Chi-Square", National Academy of Sciences, Proceedings, **17**, 684-688.