Journal of the Korean
Data & Information Science Society
1999, Vol. 10, No. 1, pp. 1 ~ 10

Quantiles for Shapiro-Francia W' Statistic

Mezbahur Rahman [1] · **Mir Masoom Ali** [2]

Abstract

Table of the empirical quantiles for the well known Shapiro-Francia W' goodness of fit statistic is produced which is more accurate than the existing ones. Prediction equation for the quantiles of W' statistic for sample sizes 30 or more are developed. The process of computing the expected values for the standard normal variate is discussed. This work is intended to make the Shapiro-Francia W' statistic more accessible to the practitioner.

Key Words and Phrases : Empirical distribution, Expected values for normal order statistics, Gauss-Legendre quadrature, Goodness of fit tests.

1. Introduction

A test for goodness of fit usually involves examining a random sample from some unknown distribution in order to test the null hypothesis that the unknown distribution function is in fact a known, specified function.

Testing for distributional assumptions in general and for normality in particular has been a major area of continuing statistical research. A possible cause of such sustained interest is that many statistical procedures have been derived based on particular distributional assumptions, especially that of normality. In many cases the techniques are both scale and origin invariant and hence the statistics are appropriate for a test of the composite hypothesis of normality.

The Kolmogorov (1933) Goodness-of-fit test (see Conover (1980) for details) is perhaps the most useful, partly because it furnishes us with an alternative to the chi-square test for goodness of fit, designed for ordinal data, and partly because the Kolmogorov test statistic enables us to form a "confidence band" for the unknown distribution function. The Kolmogorov test is intended for use only when

[1]Institute for Mathematical Sciences and Applications, California State University Monterey Bay. Seaside, CA 93955, USA
[2]Department of Mathematical Sciences, Ball State University, Muncie, IN 47306, USA

the hypothesized distribution is completely specified, that is, when there are no unknown parameters that must be estimated from the sample. The Kolmogorov test has been modified to allow it to be used in several situations where parameters are estimated from the data. The first such modification was designed to test the composite hypothesis of normality. This test was first presented by Lilliefors (1967).

The chi-square goodness-of-fit test is flexible enough to allow some parameters to be estimated from the data. One degree of freedom is simply subtracted for each parameter estimated in the "minimum chi-square". However, the chi-square test requires that the data be grouped, and such a grouping of data is usually arbitrary. The distribution of the test statistic is known only approximately, and sometimes the power of the chi-square test is not very good.

A well-known goodness-of-fit test for normality is the Shapiro-Wilk (1965, 1968) test. Some empirical studies indicate that this test has good power in many situations when compared with other tests of the composite hypothesis of normality, including the Lilliefors test and the chi-square test (Shapiro et al., 1968; La Brecque, 1977). Other goodness-of-fit tests for the same composite hypothesis of normality have been offered by Hartley and Pfaffenberger (1972), Bowman and Shenton (1975) and Pearson et al. (1977). Recently, an alternative method was proposed by Fan (1994). In that paper she used a kernel estimate of the distribution function instead of the usual empirical distribution function.

2. Shapiro-Wilk W Statistic

The theory behind the Shapiro-Wilk test is too lengthy to present here, but the interested reader is referred to the original papers by Shapiro and Wilk (1965, 1968). The test statistic is obtained by dividing the square of an appropriate linear combination of the sample order statistics by the usual symmetric estimate of the variance. This ratio is both scale and origin invariant and hence the statistic is appropriate for a test of the composite hypothesis of normality. One useful feature of the Shapiro-Wilk test is that several independent goodness-of-fit tests may be combined into one overall test of normality. This is convenient when several small samples from possibly different populations are insufficient by themselves to reject the hypothesis of normality, but their combined evidence is enough to disprove normality. To avoid the main disadvantage of nonavailability of the necessary tables, Shapiro and Francia (1972) suggested an approximate W′ test for normality.

3. An Approximate Shapiro-Francia (1972) W' Test For Normality

Let $(x_1, \cdots, x_n)$ be a random sample to be tested for departure from normality, ordered $x_{(1)} < x_{(2)} < \cdots < x_{(n)}$, and let $\mathbf{m}'$ denote the vector of expected values of the standard normal order statistics.

Define

$$W' = \frac{(\sum_1^n m_i x_{(i)})^2}{\sum_1^n m_i^2 \times \sum_1^n (x_i - \bar{x})^2}. \tag{1}$$

Note that W' equals the square of the standard product-moment correlation coefficient between the $x_{(i)}$ and m_i, and therefore measures the straightness of the normal probability plot of the $x_{(i)}$; small values of W' indicate non-normality. The null distribution of W' is itself far from normal, but Royston (1982) showed by Monte Carlo simulation that for Shapiro-Wilk W the transformed variable

$$y = (1 - W)^\lambda \tag{2}$$

(where λ is a function of sample size, n) was approximately normal. He provided polynomials to evaluate λ and the moments μ_y and σ_y of y for $7 \leq n \leq 2000$, enabling to use

$$z = (y - \mu_y)/\sigma_y \tag{3}$$

as a standard normal deviate for test purposes. λ was positive for all n.

The same approach was taken by Royston (1983) with W', in the range $5 \leq n \leq 1000$. Eight thousand values of W' were simulated from pseudo-random standard normal deviates for each selected sample size in the above range. Then values of λ were estimated by regression of quantiles of W' on normal quantiles, and then smoothed using the following polynomial in $X = log_e(n) - 5$:

$$\hat{\lambda} = -0.048157 + 0.019720X - 0.011907X^3 \tag{4}$$

Since $\hat{\lambda}$ became negative at $n = 22$, the transformation

$$y = [(1 - W')^\lambda - 1]/\lambda \tag{5}$$

was preferred to $(1 - W')^\lambda$ for practical purposes. The mean μ_y and s.d. σ_y of y using smoothed λ's were then obtained, and themselves smoothed with the following polynomials:

$$\hat{\mu}_y = -e^{1.693067 + 0.144167X - 0.018493X^2 + 0.031074X^3 + 0.005572X^4} \tag{6}$$

and

$$\hat{\sigma}_y = e^{-0.510725 - 0.116036X - 0.006702X^2 + 0.054466X^3 + 0.008740X^4} \tag{7}$$

The test statistic z is, as before,

$$z = (y - \hat{\mu}_y)/\hat{\sigma}_y \tag{8}$$

referred to the upper tail of $N(0, 1)$ (since y moves in the opposite direction to W').

4. Motivation

Among practitioners, Shapiro-Francia W' test statistic is more familiar than that of the original Shapiro-Wilk W statistic. The convergence of the asymptotic distribution of the W statistic and hence the W' statistic is surprisingly slow (see Verrill and Johnson (1988)). The dependence on the empirical distribution of the W' statistic through Monte Carlo simulation is high. The accuracy of the W' statistic depends on the accuracy of the moments of the standard normal distribution. This paper addresses how to compute the expected values of the standard normal variate. And then uses these expected values to generate a table of empirical W' percentiles through Monte Carlo simulation. The present table gives more accurate percentiles for a wide range of sample sizes. Finally, prediction equations for W' percentiles are given for the sample sizes which are absent in this table.

5. Expected Values of Normal Order Statistics

The computation of W' and its empirical distribution solely depend on the expected values of the standard normal order statistics, m_i's. Values of m_i's are given by Harter (1961), using David-Johnson (1954) series for the expected values of normal order statistics. Recently, Parrish (1992) computed expected values of normal order statistics using Gauss-Legendre quadrature techniques. Expected values of normal order statistics are given in Parrish (1992) to 25 decimal places for samples of sizes 2(1)50. The expected value of the ith smallest order statistic in a sample of size n from a normal parent distribution is given by

$$E[X_{i|n}] = K_{in} \int_{-\infty}^{\infty} x f(x)[F(x)]^{i-1}[1 - F(x)]^{n-i} dx \tag{9}$$

where

$$K_{in} = \frac{n!}{[(n-i)!(i-1)!]},$$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2},$$

$$F(x) = \int_{-\infty}^{x} f(t)dt.$$

The Gauss-Legendre quadrature method is employed to compute (9) numerically as

$$E(X_{i|n}) = K_{in} \int_{-A}^{A} x f(x)[F(x)]^{i-1}[1 - F(x)]^{n-i} dx + e_{in} \tag{10}$$

where K_{in} is defined above, A is a suitably chosen constant, and e_{in} is the error due to bounded range. Since the limits on the integral define a finite integration range, the Gauss-Legendre approach (see Press et al. (1986), Section 4.5) can be applied using only a linear transformation of abscissas. Thus, the numerical approximation of the expected value is given, generally, by the summation expression

$$K_{in}\sum_{j=1}^{N} w_j x_j f(x_j)[F(x_j)]^{i-1}[1-F(x_j)]^{n-i} \tag{11}$$

where the w_j represent appropriate weights and the x_j represent appropriate integration points. The reader is referred to Press et al. (1986), Section 4.5, for details about Gauss-Legendre polynomials and the corresponding weights. This type of numerical integration produces values that converge as the number of points N increases. The accuracy of such computed expected values are very high and are addressed in Parrish (1992) using the standard relations between the expected values of the standard normal variate.

Royston (1992) suggested an approximation for the coefficients required to compute the Shapiro-Wilk W statistic. A normalizing transformation for the W statistic is also given which is considered a wide range of sample sizes.

6. Empirical W' Quantiles

The available tables such as in Conover (1980) produced by Pearson and Hartley (1976) used expected values computed by Harter (1961). More recently, Verrill and Johnson (1988) commented about a surprisingly slow rate of the asymptotic distribution of the W statistic and gave an empirical table through Monte Carlo simulation for a modification of the W statistic. Here we used the expected values computed by Parrish (1992) which are more accurate. 10000 samples are generated using Monte-Carlo simulation under the null hypothesis, that is, from the standard normal population. W' was computed using the expression given in equation (1) for each sample using Parrish (1992) expected values for the samples of sizes 3(1)50. And for the samples of sizes 60(10)100, 150(50)500, and 600(100)1000, W' values are computed using the FORTRAN77 code by G. Rybicki given in Press et al. (1986) on pages 125-126. In expected value computations, the abscissa is considered to be in the range -12.2 to 12.2 as in Parrish (1992). The empirical W' quantiles are given in Table 1.

An exploratory regression analysis is done for each quantile values with respect to different sample sizes 30(1)50, 60(10)100, 150(50)500, and 600(100)1000 using the SPSS (statistical package for social sciences) software. The best fit prediction equation found to be

$$\hat{q} = \hat{\beta}_0 + \hat{\beta}_1 \ln n + \hat{\beta}_2 (\ln n)^2 \tag{12}$$

where $\hat{q}$ represent the predicted quantile value for a given sample size n ($\geq$) and 'ln' represents the natural logarithm. The R^2 (the coefficient of determination) values are very close to 1 and the mean squared errors (MSE) are very small. The regression coefficients, R^2, and MSE are displayed in Table 2.

In table 2, P represents the corresponding percentiles. It is to be noted that the p-values for all F and t tests for the fitted models were 0.0000.

7. Conclusion

The Table 1 gives more accurate quantile values as they are computed using more accurate expected values. And the empirical quantiles are computed using 10000 samples in each case. Finally, prediction relations are established for computing W' quantiles for arbitrary samples. These relations are simpler to use than that of Royston (1983) and have very high R^2 values and very small MSE values. To compute the expected values of the standard normal variate, the practitioner may use the table given by Parrish (1992) for sample sizes 2(1)50 or may use the FORTRAN77 code by G. Rybicki given in Press et al. (1986) pages 125-126. It is to be noted that similar code in C (programming language) also given by G. Rybicki and available in a seperate book by the same authors (Press et al.).

Table 1: W' Quantiles

n	0.01	0.02	0.05	0.10	0.50	0.90	0.95	0.98	0.99
3	0.754	0.759	0.774	0.793	0.935	0.997	0.999	1.000	1.000
4	0.682	0.706	0.753	0.800	0.918	0.984	0.992	0.997	0.998
5	0.679	0.714	0.769	0.813	0.921	0.975	0.985	0.992	0.995
6	0.699	0.735	0.789	0.829	0.926	0.973	0.981	0.988	0.992
7	0.718	0.752	0.800	0.839	0.928	0.972	0.980	0.986	0.989
8	0.735	0.771	0.815	0.851	0.933	0.972	0.979	0.985	0.989
9	0.754	0.785	0.831	0.863	0.938	0.973	0.980	0.985	0.988
10	0.769	0.801	0.843	0.873	0.942	0.974	0.980	0.984	0.988
11	0.791	0.816	0.851	0.880	0.945	0.975	0.980	0.985	0.987
12	0.795	0.828	0.861	0.888	0.947	0.976	0.980	0.985	0.988
13	0.812	0.834	0.867	0.893	0.950	0.976	0.981	0.985	0.988
14	0.817	0.841	0.876	0.900	0.952	0.977	0.982	0.986	0.988
15	0.825	0.849	0.880	0.903	0.954	0.978	0.982	0.986	0.988
16	0.839	0.859	0.888	0.909	0.957	0.979	0.983	0.987	0.989
17	0.836	0.861	0.889	0.909	0.957	0.979	0.983	0.987	0.989
18	0.838	0.862	0.893	0.914	0.959	0.980	0.984	0.937	0.989
19	0.856	0.874	0.899	0.919	0.961	0.981	0.984	0.987	0.989
20	0.861	0.880	0.902	0.921	0.962	0.981	0.985	0.988	0.989

Table 1 continued: W' Quantiles

n	0.01	0.02	0.05	0.10	0.50	0.90	0.95	0.98	0.99
21	0.863	0.882	0.906	0.924	0.963	0.982	0.985	0.988	0.990
22	0.872	0.887	0.908	0.926	0.965	0.982	0.985	0.989	0.990
23	0.876	0.891	0.913	0.929	0.966	0.983	0.986	0.989	0.990
24	0.878	0.894	0.916	0.932	0.967	0.983	0.986	0.989	0.991
25	0.880	0.896	0.916	0.932	0.968	0.984	0.986	0.989	0.990
26	0.884	0.899	0.921	0.935	0.969	0.984	0.987	0.990	0.991
27	0.889	0.904	0.923	0.937	0.969	0.984	0.987	0.990	0.991
28	0.895	0.908	0.927	0.939	0.971	0.985	0.988	0.990	0.991
29	0.885	0.909	0.927	0.941	0.971	0.985	0.988	0.990	0.991
30	0.895	0.911	0.929	0.941	0.971	0.986	0.988	0.990	0.992
31	0.898	0.913	0.931	0.944	0.972	0.986	0.988	0.990	0.992
32	0.900	0.916	0.933	0.945	0.973	0.986	0.988	0.991	0.992
33	0.906	0.918	0.934	0.947	0.973	0.986	0.989	0.991	0.992
34	0.907	0.919	0.935	0.947	0.974	0.987	0.989	0.991	0.992
35	0.910	0.921	0.937	0.949	0.975	0.987	0.989	0.991	0.992
36	0.911	0.924	0.938	0.950	0.975	0.987	0.989	0.991	0.993
37	0.913	0.925	0.941	0.952	0.976	0.987	0.989	0.991	0.992
38	0.919	0.927	0.942	0.952	0.976	0.988	0.990	0.991	0.993
39	0.916	0.929	0.943	0.954	0.977	0.988	0.990	0.992	0.993
40	0.920	0.933	0.945	0.955	0.977	0.988	0.990	0.992	0.993
41	0.925	0.933	0.946	0.956	0.978	0.988	0.990	0.992	0.993
42	0.924	0.934	0.946	0.956	0.978	0.988	0.990	0.992	0.993
43	0.926	0.936	0.948	0.957	0.978	0.989	0.990	0.992	0.993
44	0.929	0.937	0.949	0.958	0.979	0.989	0.991	0.992	0.993
45	0.927	0.937	0.949	0.958	0.979	0.989	0.991	0.992	0.993
46	0.928	0.939	0.950	0.959	0.979	0.989	0.991	0.993	0.994
47	0.930	0.940	0.951	0.960	0.980	0.989	0.991	0.993	0.993
48	0.933	0.941	0.952	0.960	0.980	0.989	0.991	0.993	0.994
49	0.934	0.942	0.954	0.962	0.980	0.990	0.991	0.993	0.994
50	0.933	0.943	0.953	0.962	0.981	0.990	0.991	0.993	0.994
60	0.944	0.951	0.961	0.968	0.983	0.991	0.992	0.994	0.995
70	0.951	0.958	0.965	0.972	0.985	0.992	0.993	0.994	0.995
80	0.957	0.962	0.969	0.974	0.987	0.993	0.994	0.995	0.995
90	0.961	0.966	0.973	0.977	0.988	0.993	0.994	0.995	0.996
100	0.965	0.969	0.975	0.979	0.989	0.994	0.995	0.996	0.996

Table 1 continued: W' Quantiles

n	0.01	0.02	0.05	0.10	0.50	0.90	0.95	0.98	0.99
150	0.975	0.978	0.982	0.985	0.992	0.996	0.996	0.997	0.997
200	0.981	0.984	0.986	0.989	0.994	0.997	0.997	0.998	0.998
250	0.985	0.987	0.989	0.991	0.995	0.997	0.998	0.998	0.998
300	0.987	0.989	0.991	0.992	0.996	0.998	0.998	0.998	0.998
350	0.989	0.990	0.992	0.993	0.996	0.998	0.998	0.998	0.999
400	0.990	0.991	0.993	0.994	0.997	0.998	0.998	0.999	0.999
450	0.992	0.992	0.994	0.995	0.997	0.998	0.998	0.999	0.999
500	0.992	0.993	0.994	0.995	0.997	0.998	0.999	0.999	0.999
600	0.993	0.994	0.995	0.996	0.998	0.999	0.999	0.999	0.999
700	0.994	0.995	0.996	0.996	0.998	0.999	0.999	0.999	0.999
800	0.995	0.996	0.996	0.997	0.998	0.999	0.999	0.999	0.999
900	0.996	0.996	0.997	0.997	0.998	0.999	0.999	0.999	0.999
1000	0.996	0.996	0.997	0.997	0.999	0.999	0.999	0.999	0.999

Table 2: $\hat{q} = \hat{\beta}_0 + \hat{\beta}_1 \ln n + \hat{\beta}_2 (\ln n)^2$

P	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	R^2	MSE
0.01	0.555057	0.139587	-0.011077	0.99176	0.00321
0.02	0.620266	0.118585	-0.009375	0.99343	0.00247
0.05	0.699734	0.093229	-0.007336	0.99466	0.00178
0.10	0.758424	0.074571	-0.005840	0.99514	0.00138
0.50	0.885117	0.034852	-0.002691	0.99673	0.00055
0.90	0.943888	0.016622	-0.001258	0.99819	0.00021
0.95	0.954141	0.013450	-0.001008	0.99857	0.00015
0.98	0.963031	0.010761	-0.000802	0.99875	0.00012
0.99	0.968407	0.009087	-0.000669	0.99823	0.00012

Bibliography

1. Bowman, K. O. and Shenton, L. R. (1975). Omnibus test contours for departures from normality based on $\sqrt{b_1}$, b_2, *Biometrika*, 62(2), 243-250.

2. Conover, W. J. (1980). *Practical Nonparametric Statistics*, 2nd ed. John Wiley & Sons, New York.

3. David, F. N. and Johnson, N. L. (1954). Statistical treatment of censored data, *Biometrika*, 41(1 and 2), 228-240.

4. Fan, Y. (1994). Testing the goodness-of-fit of a parametric density function by kernel method, *Econometric Theory*, 10, 316-356.

5. Harter, H. L. (1961). Expected values of normal order statistics, *Biometrika*, 48(1 and 2), 151-159.

6. Hartley, H. O. and Pfaffenberger, R. C. (1972). Quadratic forms in order statistics used as goodness-of-fit criteria, *Biometrika*, 59(3), 605-611.

7. Kolmogorov, A. (1933). Sulla determinazione empirica di una legge di distribuzione, *G. Ist. Ital. Attuari*, 4, 83.

8. La Brecque, J. (1977). Goodness-of-fit tests based on non-linearity in probability plots, *Technometrics*, 19(3), 293-306.

9. Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown, *Journal of the American Statistical Association*, 62(318), 399-402.

10. Parrish, R. S. (1992). Computing expected values of normal order statistics, *Communications in Statistics — Simulation and Computation*, 21(1), 57-70.

11. Pearson, E. S., D'Agostino, R. B., and Bowman, K. O. (1977). Tests for departure from normality: Comparison of powers, *Biometrika*, 64(2), 231-246.

12. Pearson, E. S. and Hartley, H. O. (1976). *Biometrika Tables for Statisticians*, vol. 1, 3rd ed. Cambridge University Press, Cambridge.

13. Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1986). *Numerical Recipes: The art of Scientific Computing*, Cambridge University Press, Cambridge.

14. Royston, J. P. (1982). An extension of Shapiro-Wilk W test for normality to large samples, *Applied Statistics*, 31(2), 115-124.

15. Royston, J. P. (1983). A simple method for evaluating the Shapiro-Francia W' test of non-normality, *The Statistician*, 32(3), 297-300.

16. Royston, J. P. (1992). Approximating the Shapiro-Wilk W-test for non-normality, *Statistics and Computing*, 2, 117-119.

17. Shapiro, S. S. and Francia, R. S. (1972). An approximate analysis of variance test for normality, *Journal of the American Statistical Association*, 67(337), 215-216.

18. Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples), *Biometrika*, 52(3 and 4), 591-611.

19. Shapiro, S. S. and Wilk, M. B. (1968). Approximations for the null distribution of the W statistic, *Technometrics*, 10(4), 861-866.

20. Shapiro, S. S., Wilk, M. B., and Chen, Mrs. H. J. (1968). A comparison study of various tests for normality, *Journal of the American Statistical Association*, 63(324), 1343-1372.

21. Verrill, S. and Johnson, R. A. (1988). Tables and Large-Sample Distribution Theory for Censored-Data Correlation Statistics for Testing Normality, *Journal of the American Statistical Association*, 83(404), 1192-1197.