Journal of the Korean
Data & Information Science Society
2000, Vol. 11, No. 2, pp. 139 ~ 155

Shapriro-Francia W' Statistic Using Exclusive Monte Carlo Simulation

Mezbahur Rahman · Larry M. Pearson [1]

Abstract

An exclusive simulation study is conducted in computing means for order statistics in standard normal variate. Monte Carlo moments are used in Shapiro-Francia W' statistic computation. Finally, quantiles for Shapiro-Francia W' are generated. The study shows that in computing means for order statistics in standard normal variate, complicated distributions and intensive numerical integrations can be avoided by using Monte Carlo simulation. Lack of accuracy is minimal and computation simplicity is noteworthy.

Key Words and Phrases: Gauss-Legendre quadrature method; Goodness-of-fit tests; Shapiro-Wilk W Statistic.

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]Department of Mathematics and Statistics, Minnesota State University, Mankato, MN 56002

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 $(x_1, x_2, \cdots, x_n)$ be a random sample to be tested for departure from normality and $x_{(1)} < x_{(2)} < \cdots < x_{(n)}$ are the corresponding order statistics. Let

$m_1, m_2, \cdots, m_n$ are the expected values of the standard normal order statistics. Define

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

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 the Shapiro-Wilk W the transformed variable

$$y = (1 - W)^\lambda$$

(where λ is a function of the 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

$$z = (y - \mu_y)/\sigma_y$$

as a standard normal deviate for test purposes. For all values of $n \lambda$ was positive.

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 $X = log_e(n) - 5$:

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

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

$$y = [(1 - W')^\lambda - 1]/\lambda$$

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:

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

and

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

The test statistic z is, as before,

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

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

Among practitioners, the 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. Recently, Rahman and Ali (1999) generated a table of empirical W' percentiles through Monte Carlo simulation using the most accurate expected values for order statistics in standard normal variate available.

4. 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 the m_i's are given by Harter (1961), using the 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

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

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 (1) numerically as

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

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

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

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.

Here we compute the expected values for order statistics in standard normal variate using exclusive Monte Carlo simulation. That is, we simulate 1,000,000 samples of different sizes from the standard normal distribution and save the ordered measurements. Finally, the means of each ordered positions are considered as the Monte-Carlo means of order statistics. It is to be noted that due to simulation, symmetry of the moments are lost. Symmetry of the moments are reinforced by using $m_{(n+1)/2} = 0$ when n is odd and by using $m_i = (m_i - m_{n-i+1})/2$ due to the fact that $m_i = -m_{n-i+1}$. And hence only m_i's for $i < [\frac{n}{2}]$, where $[x]$ is the largest integer less than or equal to x are presented. The Monte-Carlo means and their standard errors are displayed for samples of sizes 3 through 50 in Table 1. It is to be noted that the means are very close to those computed by Parish (1992) and are with very small standard errors.

Table 1: Means and their Standard Errors for Order Statistics

i	n	$\hat{E}(X_{i:n})$	$S.E.(\hat{E}(X_{i:n}))$	i	n	$\hat{E}(X_{i:n})$	$S.E.(\hat{E}(X_{i:n}))$
1	3	-0.846219	0.000955659	2	14	-1.208030	0.001240609
1	4	-1.029580	0.001119622	3	14	-0.901287	0.000931656
2	4	-0.296747	0.000387372	4	14	-0.661790	0.000692184
1	5	-1.162670	0.001240149	5	14	-0.455393	0.000486937
2	5	-0.494853	0.000570559	6	14	-0.267050	0.000300198
1	6	-1.267450	0.001336442	7	14	-0.088111	0.000120856
2	6	-0.641676	0.000706243	1	15	-1.735610	0.001776297
3	6	-0.211568	0.000268500	2	15	-1.247830	0.001278840
1	7	-1.352370	0.001415217	3	15	-0.947686	0.000976395
2	7	-0.757180	0.000813384	4	15	-0.714776	0.000743167
3	7	-0.352199	0.000410883	5	15	-0.515601	0.000544791
1	8	-1.423570	0.001481472	6	15	-0.335130	0.000365675
2	8	-0.852396	0.000902562	7	15	-0.165126	0.000196856
3	8	-0.472863	0.000524429	1	16	-1.765560	0.001804944
4	8	-0.152651	0.000205641	2	16	-1.284430	0.001313946
1	9	-1.485540	0.001539412	3	16	-0.989955	0.001017112
2	9	-0.932217	0.000977710	4	16	-0.762898	0.000789617
3	9	-0.571762	0.000617546	5	16	-0.569587	0.000596843
4	9	-0.274259	0.000322576	6	16	-0.395907	0.000424226
1	10	-1.538740	0.001589512	7	16	-0.233568	0.000263219
2	10	-1.001030	0.001042898	8	16	-0.077209	0.000106316
3	10	-0.655714	0.000697053	1	17	-1.794150	0.001832404
4	10	-0.375865	0.000419350	2	17	-1.319120	0.001347502
5	10	-0.122833	0.000166743	3	17	-1.029640	0.001055464
1	11	-1.586620	0.001634682	4	17	-0.807569	0.000832853
2	11	-1.061490	0.001100400	5	17	-0.619631	0.000645185
3	11	-0.728165	0.000765917	6	17	-0.451453	0.000477853
4	11	-0.461665	0.000500743	7	17	-0.295213	0.000322828
5	11	-0.224795	0.000265939	8	17	-0.145951	0.000174490
1	12	-1.629100	0.001675022	1	18	-1.819810	0.001856898
2	12	-1.115710	0.001152073	2	18	-1.350360	0.001377742
3	12	-0.792778	0.000827633	3	18	-1.065790	0.001090472
4	12	-0.536769	0.000572478	4	18	-0.848199	0.000872224
5	12	-0.312273	0.000349766	5	18	-0.664714	0.000688815
6	12	-0.102659	0.000140161	6	18	-0.501212	0.000525981
1	13	-1.667900	0.001711955	7	18	-0.350525	0.000376344
2	13	-1.163650	0.001198023	8	18	-0.207579	0.000234485
3	13	-0.849442	0.000881927	9	18	-0.068778	0.000094897
4	13	-0.602459	0.000635298	1	19	-1.844550	0.001880663
5	13	-0.387909	0.000422255	2	19	-1.379830	0.001406193
6	13	-0.190326	0.000226173	3	19	-1.099260	0.001122877
1	14	-1.703950	0.001746229	4	19	-0.885540	0.000908346

Table 1 Contd.: Means and their Standard Errors for Order Statistics

i	n	$\hat{E}(X_{i:n})$	$S.E.(\hat{E}(X_{i:n}))$	i	n	$\hat{E}(X_{i:n})$	$S.E.(\hat{E}(X_{i:n}))$
5	19	-0.706294	0.000729113	7	23	-0.569043	0.000588403
6	19	-0.547096	0.000570429	8	23	-0.446200	0.000466152
7	19	-0.401338	0.000425502	9	23	-0.329720	0.000350394
8	19	-0.263437	0.000288679	10	23	-0.217649	0.000239122
9	19	-0.130698	0.000156568	11	23	-0.108233	0.000130095
1	20	-1.867660	0.001902764	1	24	-1.947330	0.001979522
2	20	-1.407890	0.001433366	2	24	-1.503120	0.001525862
3	20	-1.131180	0.001153858	3	24	-1.238910	0.001258746
4	20	-0.921434	0.000943214	4	24	-1.040820	0.001059460
5	20	-0.745790	0.000767442	5	24	-0.876812	0.000895056
6	20	-0.590538	0.000612559	6	24	-0.733427	0.000751673
7	20	-0.448435	0.000471173	7	24	-0.603738	0.000622243
8	20	-0.314797	0.000338491	8	24	-0.483756	0.000502713
9	20	-0.186918	0.000211501	9	24	-0.370287	0.000389860
10	20	-0.062002	0.000085742	10	24	-0.261612	0.000281932
1	21	-1.888890	0.001923348	11	24	-0.155809	0.000176793
2	21	-1.433370	0.001458081	12	24	-0.051789	0.000071891
3	21	-1.160500	0.001182437	1	25	-1.965730	0.001997524
4	21	-0.953733	0.000974658	2	25	-1.524430	0.001546648
5	21	-0.781411	0.000802108	3	25	-1.262840	0.001282102
6	21	-0.629702	0.000650602	4	25	-1.067020	0.001085082
7	21	-0.491411	0.000512874	5	25	-0.905108	0.000922731
8	21	-0.362116	0.000384387	6	25	-0.764235	0.000781752
9	21	-0.238575	0.000261800	7	25	-0.636942	0.000654628
10	21	-0.118411	0.000142085	8	25	-0.519426	0.000537489
1	22	-1.909870	0.001943456	9	25	-0.408653	0.000427281
2	22	-1.457940	0.001481911	10	25	-0.302662	0.000321953
3	22	-1.188230	0.001209349	11	25	-0.200113	0.000220070
4	22	-0.984570	0.001004591	12	25	-0.099491	0.000119727
5	22	-0.815287	0.000835006	1	26	-1.982020	0.002013235
6	22	-0.666642	0.000686536	2	26	-1.544110	0.001565734
7	22	-0.531557	0.000551925	3	26	-1.284960	0.001303697
8	22	-0.405498	0.000426559	4	26	-1.091330	0.001108794
9	22	-0.285768	0.000307697	5	26	-0.931642	0.000948626
10	22	-0.170027	0.000192753	6	26	-0.792890	0.000809735
11	22	-0.056439	0.000078217	7	26	-0.668027	0.000685016
1	23	-1.929030	0.001961919	8	26	-0.552754	0.000570034
2	23	-1.481290	0.001504650	9	26	-0.444450	0.000462208
3	23	-1.214630	0.001235100	10	26	-0.341199	0.000359554
4	23	-1.013710	0.001033000	11	26	-0.241389	0.000260387
5	23	-0.847061	0.000865994	12	26	-0.143892	0.000163428
6	23	-0.701269	0.000720277	13	26	-0.047825	0.000066417

Table 1 Contd.: Means and their Standard Errors for Order Statistics

i	n	$\hat{E}(X_{i:n})$	$S.E.(\hat{E}(X_{i:n}))$	i	n	$\hat{E}(X_{i:n})$	$S.E.(\hat{E}(X_{i:n}))$
1	27	-1.998570	0.002029009	2	30	-1.615890	0.001635800
2	27	-1.563270	0.001584451	3	30	-1.364920	0.001381887
3	27	-1.306520	0.001324762	4	30	-1.178680	0.001194347
4	27	-1.114710	0.001131701	5	30	-1.026190	0.001041190
5	27	-0.956969	0.000973384	6	30	-0.894555	0.000909273
6	27	-0.820139	0.000836380	7	30	-0.776791	0.000791490
7	27	-0.697103	0.000713397	8	30	-0.668907	0.000683688
8	27	-0.583914	0.000600466	9	30	-0.568327	0.000583355
9	27	-0.477945	0.000494905	10	30	-0.473263	0.000488640
10	27	-0.377017	0.000394521	11	30	-0.382413	0.000398234
11	27	-0.279268	0.000297911	12	30	-0.294521	0.000310849
12	27	-0.185226	0.000204003	13	30	-0.208929	0.000225296
13	27	-0.092296	0.000111218	14	30	-0.124779	0.000142013
1	28	-2.013560	0.002043617	15	30	-0.041497	0.000057772
2	28	-1.581320	0.001602063	1	31	-2.056480	0.002085151
3	28	-1.326790	0.001344596	2	31	-1.631750	0.001651281
4	28	-1.137190	0.001153722	3	31	-1.382740	0.001399331
5	28	-0.981442	0.000997399	4	31	-1.198040	0.001213389
6	28	-0.846389	0.000862098	5	31	-1.047110	0.001061745
7	28	-0.725238	0.000740980	6	31	-0.916948	0.000931265
8	28	-0.614054	0.000629962	7	31	-0.800701	0.000814929
9	28	-0.509885	0.000526142	8	31	-0.694368	0.000708680
10	28	-0.411034	0.000427794	9	31	-0.595425	0.000609915
11	28	-0.315951	0.000333256	10	31	-0.501996	0.000516784
12	28	-0.223867	0.000241741	11	31	-0.412789	0.000427974
13	28	-0.133635	0.000151998	12	31	-0.326798	0.000342442
14	28	-0.044472	0.000061853	13	31	-0.243154	0.000259270
1	29	-2.029080	0.002058737	14	31	-0.161187	0.000177744
2	29	-1.599100	0.001619370	15	31	-0.080301	0.000096872
3	29	-1.346210	0.001363594	1	32	-2.069610	0.002097994
4	29	-1.158360	0.001174402	2	32	-1.647280	0.001666418
5	29	-1.004150	0.001019585	3	32	-1.400000	0.001416242
6	29	-0.870856	0.000886063	4	32	-1.216770	0.001231692
7	29	-0.751447	0.000766597	5	32	-1.067330	0.001081578
8	29	-0.642040	0.000657363	6	32	-0.938517	0.000952450
9	29	-0.539711	0.000555304	7	32	-0.823580	0.000837399
10	29	-0.442806	0.000458790	8	32	-0.718753	0.000732621
11	29	-0.350031	0.000366518	9	32	-0.621284	0.000635253
12	29	-0.260123	0.000277145	10	32	-0.529430	0.000543657
13	29	-0.172380	0.000189955	11	32	-0.441799	0.000456384
14	29	-0.085860	0.000103534	12	32	-0.357430	0.000372440
1	30	-2.043310	0.002072362	13	32	-0.275669	0.000291126

Table 1 Contd.: Means and their Standard Errors for Order Statistics

i	n	$\hat{E}(X_{i:n})$	$S.E.(\hat{E}(X_{i:n}))$	i	n	$\hat{E}(X_{i:n})$	$S.E.(\hat{E}(X_{i:n}))$
14	32	-0.195775	0.000211702	7	35	-0.886713	0.000899380
15	32	-0.117069	0.000133335	8	35	-0.785657	0.000798291
16	32	-0.038979	0.000054333	9	35	-0.692022	0.000704733
1	33	-2.081800	0.002109863	10	35	-0.604185	0.000617046
2	33	-1.662050	0.001680966	11	35	-0.520848	0.000533922
3	33	-1.416690	0.001432674	12	35	-0.440964	0.000454319
4	33	-1.234980	0.001249580	13	35	-0.363772	0.000377483
5	33	-1.086750	0.001100659	14	35	-0.288663	0.000302785
6	33	-0.959347	0.000972907	15	35	-0.215161	0.000229721
7	33	-0.845652	0.000859037	16	35	-0.142765	0.000157704
8	33	-0.742171	0.000755559	17	35	-0.071199	0.000086073
9	33	-0.645961	0.000659500	1	36	-2.118090	0.002144964
10	33	-0.555372	0.000569094	2	36	-1.703690	0.001721617
11	33	-0.469314	0.000483330	3	36	-1.462820	0.001477896
12	33	-0.386601	0.000400996	4	36	-1.284890	0.001298589
13	33	-0.306405	0.000321213	5	36	-1.140480	0.001153467
14	33	-0.228239	0.000243525	6	36	-1.016530	0.001029107
15	33	-0.151407	0.000167117	7	36	-0.906421	0.000918784
16	33	-0.075521	0.000091252	8	36	-0.806397	0.000818677
1	34	-2.094750	0.002122405	9	36	-0.713950	0.000726296
2	34	-1.676550	0.001695133	10	36	-0.627242	0.000639723
3	34	-1.432420	0.001448110	11	36	-0.544982	0.000557659
4	34	-1.252220	0.001266482	12	36	-0.466212	0.000479132
5	34	-1.105370	0.001118989	13	36	-0.390251	0.000403481
6	34	-0.979119	0.000992327	14	36	-0.316627	0.000330239
7	34	-0.866691	0.000879251	15	36	-0.244622	0.000258611
8	34	-0.764430	0.000777452	16	36	-0.173810	0.000188168
9	34	-0.669649	0.000682760	17	36	-0.103992	0.000118594
10	34	-0.580390	0.000593677	18	36	-0.034608	0.000048279
11	34	-0.495690	0.000509238	1	37	-2.129090	0.002155673
12	34	-0.414435	0.000428306	2	37	-1.716620	0.001734318
13	34	-0.335747	0.000350030	3	37	-1.476800	0.001491629
14	34	-0.259178	0.000273914	4	37	-1.300150	0.001313575
15	34	-0.184094	0.000199246	5	37	-1.156650	0.001169317
16	34	-0.110123	0.000125561	6	37	-1.033820	0.001046091
17	34	-0.036584	0.000051065	7	37	-0.924793	0.000936829
1	35	-2.106300	0.002133568	8	37	-0.825861	0.000837824
2	35	-1.692010	0.001708453	9	37	-0.734526	0.000746500
3	35	-1.447590	0.001462911	10	37	-0.648855	0.000660930
4	35	-1.268430	0.001282414	11	37	-0.567771	0.000580048
5	35	-1.122900	0.001136162	12	37	-0.490356	0.000502861
6	35	-0.997856	0.001010723	13	37	-0.415779	0.000428589

Table 1 Contd.: Means and their Standard Errors for Order Statistics

i	n	$\hat{E}(X_{i:n})$	$S.E.(\hat{E}(X_{i:n}))$	i	n	$\hat{E}(X_{i:n})$	$S.E.(\hat{E}(X_{i:n}))$
14	37	-0.343321	0.000356463	19	39	-0.063955	0.000141702
15	37	-0.272604	0.000286116	1	40	-2.160900	0.002186641
16	37	-0.203314	0.000217231	2	40	-1.753230	0.001770139
17	37	-0.135035	0.000149264	3	40	-1.517230	0.001531358
18	37	-0.067370	0.000081547	4	40	-1.343810	0.001356493
1	38	-2.140460	0.002166701	5	40	-1.203320	0.001215221
2	38	-1.729380	0.001746779	6	40	-1.083180	0.001094684
3	38	-1.490930	0.001505472	7	40	-0.977178	0.000988442
4	38	-1.315230	0.001328422	8	40	-0.881123	0.000892284
5	38	-1.172920	0.001185359	9	40	-0.792611	0.000803729
6	38	-1.051060	0.001063059	10	40	-0.709851	0.000720993
7	38	-0.943098	0.000954886	11	40	-0.631708	0.000642963
8	38	-0.845211	0.000856872	12	40	-0.557363	0.000568766
9	38	-0.754781	0.000766474	13	40	-0.485953	0.000497570
10	38	-0.670160	0.000681925	14	40	-0.416958	0.000428832
11	38	-0.590028	0.000601931	15	40	-0.349822	0.000361970
12	38	-0.513630	0.000525721	16	40	-0.284287	0.000296760
13	38	-0.440106	0.000452445	17	40	-0.219934	0.000232742
14	38	-0.368839	0.000381493	18	40	-0.156439	0.000169579
15	38	-0.299418	0.000312434	19	40	-0.093555	0.000106854
16	38	-0.231512	0.000244890	20	40	-0.031119	0.000043499
17	38	-0.164705	0.000178429	1	41	-2.170620	0.002196157
18	38	-0.098548	0.000112477	2	41	-1.764450	0.001781164
19	38	-0.032800	0.000045782	3	41	-1.529750	0.001543661
1	39	-2.151090	0.002177053	4	41	-1.357600	0.001370123
2	39	-1.741570	0.001758772	5	41	-1.218010	0.001229759
3	39	-1.504120	0.001518460	6	41	-1.098920	0.001110226
4	39	-1.329690	0.001342671	7	41	-0.993673	0.001004665
5	39	-1.188280	0.001200490	8	41	-0.898462	0.000909320
6	39	-1.067470	0.001079246	9	41	-0.810722	0.000821531
7	39	-0.960547	0.000972039	10	41	-0.728891	0.000739748
8	39	-0.863519	0.000874908	11	41	-0.651636	0.000662563
9	39	-0.774043	0.000785435	12	41	-0.578133	0.000589186
10	39	-0.690407	0.000701866	13	41	-0.507622	0.000518862
11	39	-0.611349	0.000622935	14	41	-0.439526	0.000450997
12	39	-0.535953	0.000547706	15	41	-0.373528	0.000385272
13	39	-0.463439	0.000475438	16	41	-0.309037	0.000321070
14	39	-0.393365	0.000405629	17	41	-0.245890	0.000258240
15	39	-0.325181	0.000337758	18	41	-0.183529	0.000196205
16	39	-0.258524	0.000271456	19	41	-0.121926	0.000134924
17	39	-0.192918	0.000206186	20	41	-0.060905	0.000073762
18	39	-0.128172	0.000141702	1	42	-2.180550	0.002205793

Table 1 Contd.: Means and their Standard Errors for Order Statistics

i	n	$\hat{E}(X_{i:n})$	$S.E.(\hat{E}(X_{i:n}))$	i	n	$\hat{E}(X_{i:n})$	$S.E.(\hat{E}(X_{i:n}))$
2	42	-1.775840	0.001792401	2	44	-1.797280	0.001813439
3	42	-1.541850	0.001555506	3	44	-1.565380	0.001578665
4	42	-1.370540	0.001382796	4	44	-1.395750	0.001407697
5	42	-1.232080	0.001243565	5	44	-1.258830	0.001269990
6	42	-1.113760	0.001124857	6	44	-1.142140	0.001152837
7	42	-1.009350	0.001020140	7	44	-1.039230	0.001049619
8	42	-0.914823	0.000925501	8	44	-0.946289	0.000956504
9	42	-0.827974	0.000838558	9	44	-0.860912	0.000871069
10	42	-0.746880	0.000757480	10	44	-0.781425	0.000791556
11	42	-0.670471	0.000681148	11	44	-0.706524	0.000716695
12	42	-0.597890	0.000608690	12	44	-0.635513	0.000645759
13	42	-0.528258	0.000539195	13	44	-0.567527	0.000577898
14	42	-0.461092	0.000472233	14	44	-0.502165	0.000512684
15	42	-0.396053	0.000407408	15	44	-0.438756	0.000449472
16	42	-0.332625	0.000344273	16	44	-0.377158	0.000388104
17	42	-0.270490	0.000282437	17	44	-0.316970	0.000328169
18	42	-0.209344	0.000221612	18	44	-0.257804	0.000269282
19	42	-0.148980	0.000161544	19	44	-0.199658	0.000211456
20	42	-0.089130	0.000101821	20	44	-0.142118	0.000154187
21	42	-0.029672	0.000041471	21	44	-0.085050	0.000097222
1	43	-2.189970	0.002214981	22	44	-0.028307	0.000039616
2	43	-1.786820	0.001803101	1	45	-2.207730	0.001232258
3	43	-1.554060	0.001567594	2	45	-1.807480	0.001823394
4	43	-1.383710	0.001395832	3	45	-1.576650	0.001589816
5	43	-1.246030	0.001257375	4	45	-1.408070	0.001419790
6	43	-1.128470	0.001139336	5	45	-1.271920	0.001282910
7	43	-1.024640	0.001035208	6	45	-1.155950	0.001166388
8	43	-0.931104	0.000941475	7	45	-1.053710	0.001063896
9	43	-0.844971	0.000855289	8	45	-0.961472	0.000971508
10	43	-0.764668	0.000775015	9	45	-0.876850	0.000886769
11	43	-0.689031	0.000699411	10	45	-0.798035	0.000807940
12	43	-0.617268	0.000627753	11	45	-0.723930	0.000733853
13	43	-0.548454	0.000559108	12	45	-0.653621	0.000663621
14	43	-0.482218	0.000493071	13	45	-0.586358	0.000596465
15	43	-0.418078	0.000429125	14	45	-0.521760	0.000532015
16	43	-0.355510	0.000366819	15	45	-0.459175	0.000469594
17	43	-0.294304	0.000305885	16	45	-0.398376	0.000408992
18	43	-0.234182	0.000246082	17	45	-0.338979	0.000349833
19	43	-0.174877	0.000187047	18	45	-0.280783	0.000291881
20	43	-0.116274	0.000128673	19	45	-0.223600	0.000234975
21	43	-0.058020	0.000070316	20	45	-0.167078	0.000178744
1	44	-2.198850	0.002223625	21	45	-0.111050	0.000122937

Table 1 Contd.: Means and their Standard Errors for Order Statistics

i	n	$\hat{E}(X_{i:n})$	$S.E.(\hat{E}(X_{i:n}))$	i	n	$\hat{E}(X_{i:n})$	$S.E.(\hat{E}(X_{i:n}))$
22	45	-0.055446	0.000067215	19	47	-0.268537	0.000279274
1	46	-2.217150	0.002241516	20	47	-0.213908	0.000224897
2	46	-1.817590	0.001833335	21	47	-0.159894	0.000171144
3	46	-1.587660	0.001600673	22	47	-0.106296	0.000117710
4	46	-1.419960	0.001431563	23	47	-0.053087	0.000064375
5	46	-1.284560	0.001295367	1	48	-2.232850	0.002256772
6	46	-1.169180	0.001179501	2	48	-1.836230	0.001851677
7	46	-1.067560	0.001077570	3	48	-1.608250	0.001620926
8	46	-0.976056	0.000985847	4	48	-1.442240	0.001453507
9	46	-0.892022	0.000901727	5	48	-1.308210	0.001318697
10	46	-0.813839	0.000823526	6	48	-1.194490	0.001204439
11	46	-0.740400	0.000750124	7	48	-1.094220	0.001103883
12	46	-0.670774	0.000680532	8	48	-1.004020	0.001013479
13	46	-0.604368	0.000614209	9	48	-0.921295	0.000930670
14	46	-0.540469	0.000550465	10	48	-0.844461	0.000853754
15	46	-0.478718	0.000488861	11	48	-0.772352	0.000781661
16	46	-0.418754	0.000429086	12	48	-0.704004	0.000713347
17	46	-0.360241	0.000370795	13	48	-0.638821	0.000648228
18	46	-0.302969	0.000313775	14	48	-0.576309	0.000585828
19	46	-0.246637	0.000257722	15	48	-0.516005	0.000525651
20	46	-0.191083	0.000202426	16	48	-0.457452	0.000467251
21	46	-0.136076	0.000147679	17	48	-0.400390	0.000410392
22	46	-0.081528	0.000093229	18	48	-0.344660	0.000354878
23	46	-0.027172	0.000038023	19	48	-0.289989	0.000300437
1	47	-2.224880	0.002248922	20	48	-0.236150	0.000246848
2	47	-1.827150	0.001842749	21	48	-0.182996	0.000193933
3	47	-1.598450	0.001611264	22	48	-0.130368	0.000141502
4	47	-1.431300	0.001442694	23	48	-0.078037	0.000089240
5	47	-1.296570	0.001307173	24	48	-0.025991	0.000036370
6	47	-1.182120	0.001192256	1	49	-2.241560	0.002265195
7	47	-1.081070	0.001090927	2	49	-1.845780	0.001861143
8	47	-0.990206	0.000999845	3	49	-1.618870	0.001631350
9	47	-0.906880	0.000916415	4	49	-1.453250	0.001464365
10	47	-0.829457	0.000838959	5	49	-1.319970	0.001330291
11	47	-0.756627	0.000766152	6	49	-1.206650	0.001216482
12	47	-0.687752	0.000697296	7	49	-1.107080	0.001116603
13	47	-0.621950	0.000631593	8	49	-1.017530	0.001026806
14	47	-0.558775	0.000568515	9	49	-0.935413	0.000944586
15	47	-0.497700	0.000507581	10	49	-0.859076	0.000868170
16	47	-0.438350	0.000448405	11	49	-0.787462	0.000796547
17	47	-0.380585	0.000390874	12	49	-0.719715	0.000728838
18	47	-0.324070	0.000334556	13	49	-0.655152	0.000664327

Table 1 Contd.: Means and their Standard Errors for Order Statistics

i	n	$\hat{E}(X_{i:n})$	$S.E.(\hat{E}(X_{i:n}))$	i	n	$\hat{E}(X_{i:n})$	$S.E.(\hat{E}(X_{i:n}))$
14	49	-0.593330	0.000602578	8	50	-1.030490	0.001039628
15	49	-0.533595	0.000542971	9	50	-0.948900	0.000957945
16	49	-0.475781	0.000485318	10	50	-0.873273	0.000882205
17	49	-0.419511	0.000429220	11	50	-0.802282	0.000811218
18	49	-0.364491	0.000374408	12	50	-0.735157	0.000744088
19	49	-0.310561	0.000320689	13	50	-0.671223	0.000680209
20	49	-0.257499	0.000267879	14	50	-0.609897	0.000618973
21	49	-0.205172	0.000215787	15	50	-0.550731	0.000559910
22	49	-0.153416	0.000164257	16	50	-0.493533	0.000502855
23	49	-0.102074	0.000113101	17	50	-0.437880	0.000447361
24	49	-0.050934	0.000061786	18	50	-0.383535	0.000393198
1	50	-2.249040	0.002272591	19	50	-0.330334	0.000340216
2	50	-1.854980	0.001870090	20	50	-0.277994	0.000288086
3	50	-1.628610	0.001640857	21	50	-0.226496	0.000236822
4	50	-1.463710	0.001474668	22	50	-0.175559	0.000186106
5	50	-1.331070	0.001341240	23	50	-0.125052	0.000135808
6	50	-1.218480	0.001228128	24	50	-0.074926	0.000085749
7	50	-1.119570	0.001128896	25	50	-0.024974	0.000034967

For samples larger than 50, one can follow the process presented here or use the approximate formula

$$\tilde{m}_i = \Phi^{-1}\left(\frac{i-\frac{3}{8}}{n+\frac{1}{4}}\right), \quad i = 1, 2, \cdots, n$$

given by Blom (1958).

5. 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. Rahman and Ali (1999) computed the expected values using Parrish (1992). Here, we generate a similar table using Monte-Carlo moments. We generate 10,000 samples under the null hypothesis, that is, from the standard normal population. W' was computed using the expression

$$W' = \frac{\left(\sum_{i=1}^{n}(m_i - \bar{m})(x_{(i)} - \bar{x})\right)^2}{\sum_{i=1}^{n}(m_i - \bar{m})^2 \times \sum_{1}^{n}(x_i - \bar{x})^2}$$

where $\bar{m} = \sum_{i=1}^{n} m_i/n$, is needed to adjust for variations due to simulation. And the symmetricity of the moments are deviated slightly for the same reason. The empirical W' quantiles are given in Table 2.

Table 2: 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	0.772	0.793	0.933	0.997	0.999	1.000	1.000
4	0.677	0.702	0.749	0.796	0.917	0.983	0.991	0.996	0.998
5	0.684	0.723	0.774	0.815	0.922	0.975	0.984	0.991	0.994
6	0.699	0.743	0.791	0.832	0.925	0.973	0.980	0.987	0.990
7	0.723	0.756	0.803	0.841	0.930	0.973	0.980	0.987	0.990
8	0.748	0.777	0.823	0.855	0.934	0.973	0.979	0.985	0.989
9	0.754	0.788	0.830	0.861	0.937	0.973	0.979	0.985	0.988
10	0.775	0.805	0.842	0.872	0.941	0.974	0.979	0.984	0.988
11	0.784	0.809	0.852	0.880	0.944	0.975	0.980	0.985	0.988
12	0.795	0.825	0.861	0.886	0.947	0.976	0.981	0.985	0.988
13	0.807	0.831	0.865	0.892	0.950	0.977	0.982	0.986	0.988
14	0.814	0.836	0.875	0.899	0.952	0.977	0.982	0.986	0.988
15	0.826	0.854	0.883	0.904	0.954	0.978	0.982	0.986	0.988
16	0.832	0.852	0.882	0.906	0.956	0.979	0.983	0.986	0.988
17	0.842	0.864	0.890	0.910	0.957	0.979	0.983	0.987	0.989
18	0.849	0.871	0.896	0.915	0.959	0.980	0.984	0.987	0.989
19	0.859	0.878	0.900	0.918	0.961	0.981	0.984	0.987	0.989
20	0.855	0.874	0.902	0.921	0.962	0.981	0.985	0.988	0.990
21	0.866	0.886	0.908	0.924	0.963	0.982	0.985	0.988	0.990
22	0.874	0.889	0.911	0.927	0.965	0.983	0.986	0.989	0.990
23	0.872	0.891	0.913	0.929	0.966	0.983	0.986	0.989	0.991
24	0.879	0.896	0.916	0.931	0.967	0.983	0.986	0.989	0.990
25	0.878	0.895	0.917	0.933	0.968	0.984	0.987	0.989	0.990
26	0.881	0.898	0.920	0.935	0.969	0.984	0.987	0.989	0.991
27	0.889	0.905	0.923	0.937	0.969	0.985	0.987	0.990	0.991
28	0.892	0.907	0.926	0.939	0.970	0.985	0.987	0.990	0.991
29	0.896	0.910	0.929	0.942	0.971	0.985	0.987	0.990	0.991
30	0.896	0.911	0.929	0.943	0.972	0.986	0.988	0.990	0.991
31	0.903	0.916	0.932	0.944	0.972	0.986	0.988	0.991	0.992
32	0.906	0.917	0.933	0.946	0.973	0.986	0.988	0.991	0.992
33	0.906	0.918	0.935	0.947	0.973	0.986	0.989	0.991	0.992
34	0.908	0.921	0.937	0.948	0.974	0.987	0.989	0.991	0.992
35	0.910	0.921	0.938	0.949	0.975	0.987	0.989	0.991	0.992
36	0.910	0.924	0.939	0.950	0.976	0.987	0.990	0.991	0.992
37	0.917	0.927	0.942	0.951	0.976	0.987	0.989	0.991	0.992
38	0.915	0.927	0.942	0.953	0.976	0.988	0.990	0.992	0.992
39	0.916	0.929	0.943	0.953	0.977	0.988	0.990	0.992	0.993
40	0.922	0.931	0.944	0.954	0.977	0.988	0.990	0.992	0.993

Table 2 Contd.: W' Quantiles

n	0.01	0.02	0.05	0.10	0.50	0.90	0.95	0.98	0.99
41	0.922	0.931	0.945	0.956	0.978	0.988	0.990	0.992	0.993
42	0.919	0.932	0.946	0.956	0.978	0.988	0.990	0.992	0.993
43	0.926	0.935	0.948	0.957	0.979	0.989	0.991	0.992	0.993
44	0.926	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	0.994
46	0.927	0.938	0.950	0.959	0.979	0.989	0.991	0.993	0.993
47	0.930	0.940	0.951	0.959	0.980	0.989	0.991	0.993	0.993
48	0.932	0.940	0.952	0.960	0.980	0.989	0.991	0.993	0.994
49	0.932	0.940	0.952	0.961	0.981	0.990	0.991	0.993	0.994
50	0.932	0.941	0.953	0.962	0.981	0.990	0.992	0.993	0.994
60	0.945	0.952	0.961	0.968	0.984	0.991	0.993	0.994	0.994
70	0.951	0.957	0.965	0.971	0.985	0.992	0.993	0.995	0.995
80	0.956	0.961	0.969	0.974	0.987	0.993	0.994	0.995	0.995
90	0.960	0.966	0.972	0.977	0.988	0.993	0.994	0.995	0.996
100	0.964	0.969	0.975	0.979	0.989	0.994	0.995	0.996	0.996
150	0.975	0.978	0.982	0.985	0.992	0.996	0.996	0.997	0.997
200	0.981	0.983	0.986	0.988	0.994	0.997	0.997	0.998	0.998
250	0.984	0.986	0.989	0.991	0.995	0.997	0.997	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	0.991	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	0.995	0.996	0.997	0.998	0.999	0.999	0.999	0.999
900	0.995	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

6. Conclusion

The intention of this work is two-fold. First, a practitioner can apply the Shapiro-Francia W' using the tables produced here without any other table(s). Secondly, this Monte-Carlo method can be applied using minimal computation and without access of any table.

7. Acknowledgements

We thank the referee for constructive comments and suggestions which improved the earlier version of the paper.

Bibliography

1. Blom, G. (1958). *Statistical Estimates and Transformed Beta-variables.* Wiley, New York.

2. 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.

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

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

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

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

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

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

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

10. 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.

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

12. 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.

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

14. 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.

15. Rahman, M. and Ali, M. M. (1999). Quantiles for Shapiro-Francia W' Statistic, *Journal of the Korean Data & Information Science Society*, 10(1), 1-10.

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

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

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

19. 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.

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

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

22. 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.

23. 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.