Appl. Statist. (1982),
31, *No.* 2, *pp.* 115–124

An Extension of Shapiro and Wilk's *W* Test for Normality to Large Samples

By J. P. Royston

MRC Clinical Research Centre, Watford Road, Harrow HA1 3UJ, Middx., UK

[Received February 1981. Revised December 1981]

Summary

Shapiro and Wilk's (1965) W statistic arguably provides the best omnibus test of normality, but is currently limited to sample sizes between 3 and 50. W is extended up to $n = 2000$ and an approximate normalizing transformation suitable for computer implementation is given. A novel application of W in transforming data to normality is suggested, using the three-parameter lognormal as an example.

Keywords: TESTS OF NORMALITY; OMNIBUS TEST; W STATISTIC; NORMALIZING TRANSFORMATION; POLYNOMIAL SMOOTHING

1. Introduction

Research into tests of non-normality was given new impetus with the introduction of the so-called analysis of variance test by Shapiro and Wilk (1965). The test statistic W was constructed by considering the regression of ordered sample values on corresponding expected normal order statistics, which for a sample from a normally distributed population is linear; W was obtained as an F-ratio from generalized least-squares analysis to judge the adequacy of the linear fit. Percentage points of the null distribution of W were tabulated for $p = 0{\cdot}01, 0{\cdot}02, 0{\cdot}05, 0{\cdot}1, 0{\cdot}5, 0{\cdot}9, 0{\cdot}95, 0{\cdot}98$ and $0{\cdot}99$, for sample sizes $n = 3\,(1)\,50$. A normalizing transformation for W using a Johnson S_B approximation in the region $n = 7\,(1)\,50$ was later proposed (Shapiro and Wilk, 1968), though tables were still required for $n = 4\,(1)\,6$.

Extensive empirical comparisons of W with other tests of non-normality using computer-generated pseudo-random numbers indicated that W had good power properties against a wide range of alternative distributions, and was therefore truly an omnibus test (Shapiro *et al.*, 1968). Subsequently, other statistics of the W type, namely Y (D'Agostino, 1971), W' (Shapiro and Francia, 1972) and r (Filliben, 1975), were developed and shown to have power properties broadly comparable with those of W. Another approach was to combine $\sqrt{b_1}$ and b_2, the standard sample measures of skewness and kurtosis, into an omnibus test (D'Agostino and Pearson, 1973; Bowman and Shenton, 1975). One such statistic is K^2 (Pearson *et al.*, 1977).

The power characteristics of the omnibus tests may be summarized as follows. For symmetrical platykurtic distributions ($\sqrt{\beta_1} = 0$, $\beta_2 < 3$), and for most skew distributions, W is optimal (Pearson *et al.*, 1977), and is sometimes markedly more powerful than its nearest rival. K^2, W' and r are all more powerful than W for symmetric, leptokurtic populations ($\beta_2 > 3$), though W seems little inferior for very long-tailed alternatives, e.g. $\beta_2 = 11$. W' and r appear to have very similar power (Filliben, 1975).

With the exception of W for $n = 7\,(1)\,50$, the existing omnibus test statistics suffer from the serious disadvantage that their significance levels have to be obtained from tables, and therefore the tests are inconvenient for computer implementation. Furthermore W, which it may be argued forms the best general test, cannot be calculated for samples larger than 50. The purpose of the present paper is to extend W to sample sizes $n \leqslant 2000$, and to develop a transformation of the null distribution of W to approximate normality throughout the range $7 \leqslant n \leqslant 2000$. A method for calculating the significance level of W for $n < 7$ is also derived. All

 0035–9254/82/31115 $2.00

the techniques are easily programmable. Finally, a novel application of using W to fit the three-parameter lognormal distribution is described.

2. Calculation of W

Let $m' = (m_1, \ldots, m_n)$ denote the vector of expected values of standard normal order statistics, and let $V = (v_{ij})$ be the corresponding $n \times n$ covariance matrix; that is

$$E(x_i) = m_i \; (i = 1, \ldots, n) \quad \text{and} \quad \operatorname{cov}(x_i, x_j) = v_{ij} \; (i, j = 1, \ldots, n),$$

where $x_1 < x_2 < \ldots < x_n$ is an ordered random sample from a standard normal distribution $N(0, 1)$. Suppose $y' = (y_1, \ldots, y_n)$ is a random sample on which the W test of normality is to be carried out, ordered $y_{(1)} < y_{(2)} < \ldots < y_{(n)}$. Then

$$W = \left[\sum_1^n a_i \, y_{(i)}\right]^2 \Big/ \sum_1^n (y_i - \bar{y})^2$$

where

$$a' = (a_1, \ldots, a_n) = m' V^{-1} [(m' V^{-1})(V^{-1} m)]^{-\frac{1}{2}}.$$

The coefficients $\{a_i\}$ are the normalized "best linear unbiased" coefficients tabulated for $n \leqslant 20$ by Sarhan and Greenberg (1956). The covariance matrix V which features in a may be obtained using the algorithm of Davis and Stephens (1978). However, V is not required explicitly, and Shapiro and Wilk (1965) offer a satisfactory approximation for a which

Table 1

Some theoretical moments (μ_i) and simulation moments ($\hat{\mu}_i$) of W. Values in the present study are based on 6000 simulations per sample size

		$\hat{\mu}_1$		$\sqrt{\hat{\mu}_2}$		$\sqrt{b_1}$		b_2	
n	μ_1	*Shapiro and Wilk*	*This study*	*Shapiro and Wilk*	*This study*	*Shapiro and Wilk*	*This study*	*Shapiro and Wilk*	*This study*
7	0·9123	0·9120	0·9113	0·0547	0·0542	−1·32	−1·23	6·41	5·44
8	0·9174	0·9175	0·9174	0·0497	0·0499	−1·38	−1·27	7·11	5·44
9	0·9221	0·9215	0·9227	0·0479	0·0466	−1·60	−1·48	8·45	6·49
10	0·9264	0·9260	0·9277	0·0444	0·0432	−1·67	−1·47	9·28	6·45
15	0·9424	0·9422	0·9423	0·0320	0·0323	−1·89	−1·52	16·74	6·47
20	0·9523	0·9527	0·9523	0·0255	0·0253	−2·28	−1·55	32·59	7·76
30	—	0·9626	0·9631	0·0185	0·0183	−2·73	−1·46	71·77	6·73
40	—	0·9682	0·9684	0·0151	0·0147	−3·17	−1·34	136·48	6·17
50	—	0·9714	0·9718	0·0124	0·0126	−3·32	−1·20	212·43	6·05
60	—	—	0·9738	—	0·0110	—	−1·04	—	5·08
70	—	—	0·9755	—	0·0099	—	−1·02	—	5·64
80	—	—	0·9768	—	0·0090	—	−0·92	—	5·02
90	—	—	0·9778	—	0·0083	—	−0·84	—	4·17
100	—	—	0·9786	—	0·0078	—	−0·73	—	3·86
125	—	—	0·9799	—	0·0069	—	−0·58	—	3·48
150	—	—	0·9810	—	0·0063	—	−0·56	—	3·69
200	—	—	0·9821	—	0·0054	—	−0·44	—	3·60
300	—	—	0·9835	—	0·0044	—	−0·25	—	3·18
400	—	—	0·9844	—	0·0037	—	−0·16	—	3·17
500	—	—	0·9850	—	0·0033	—	−0·10	—	3·00
750	—	—	0·9860	—	0·0026	—	−0·02	—	3·06
1000	—	—	0·9866	—	0·0022	—	0·05	—	3·16
1500	—	—	0·9876	—	0·0018	—	0·18	—	3·09
2000	—	—	0·9884	—	0·0015	—	0·24	—	3·16

improves with increasing sample size n, which we adopt. By definition, a has the property $a'a = 1$. Let $a^* = m'V^{-1}$; approximations $\hat{a}^*$ for a^* are

$$\hat{a}_i^* = \begin{cases} 2m_i, & i = 2, 3, \ldots, n-1, \\ \left(\dfrac{\hat{a}_1^2}{1-2\hat{a}_1^2} \displaystyle\sum_2^{n-1} \hat{a}_i^{*2}\right)^{\frac{1}{2}}, & i = 1, i = n, \end{cases} \tag{2.1}$$

where

$$\hat{a}_1^2 = \hat{a}_n^2 = \begin{cases} g(n-1), & n \leqslant 20, \\ g(n), & n > 20 \end{cases}$$

and

$$g(n) = \frac{\Gamma(\frac{1}{2}[n+1])}{\sqrt{2}\,\Gamma(\frac{1}{2}n+1)}.$$

Using Stirling's formula, $g(n)$ may be approximated and simplified:

$$g(n) = \left[\frac{6n+7}{6n+13}\right]\left(\frac{\exp(1)}{n+2}\left[\frac{n+1}{n+2}\right]^{n-2}\right)^{\frac{1}{2}}. \tag{2.2}$$

For $n = 6$, the exact value of $g(n)$ is 0·39166 compared with an approximate value given by (2.2) of 0·39164. We have used the approximations (2.1) and (2.2) throughout the range $7 \leqslant n \leqslant 2000$, but exact values of the $\{a_i\}$ for $n < 7$.

Expected normal order statistics

The values of the m_i for sample sizes up to 400 were given by Harter (1961). To facilitate computer calculations of W, however, we constructed a new approximation for m_i in the range $2 \leqslant n \leqslant 2000$, accurate to 0·0001 (Royston, 1982). The algorithm, based on Blom's (1958, pp. 69–71) formula, has the advantage of rapidity when compared with "exact" calculations involving numerical integration or summation of series.

3. Extension of W

W was calculated using the formulae (2.1) and (2.2). Values of W were generated from sets of normal deviates, calculated as the normal integral transform of uniform pseudo-random numbers. Wichmann and Hill's (1982) random number generator was used. Six thousand values of W for each sample size $n = 7(1)30(5)100$, 125, 150, 200(100)600, 750, 1000, 1250, 1500 and 2000 were simulated. Different sets of pseudo-random numbers were used for each simulation, to avoid dependence between results. Selected empirical first and second moments of W are given in Table 1, and may be compared with Shapiro and Wilk's (1965, Table 4) values, which are exact for E(W) at $n \leqslant 20$ but otherwise based on extensive simulation. Use of the approximation (2.1) for the defining coefficients $\{a_i\}$ has had little effect on the moments of W, even at $n = 7$. A curious feature of Table 1 is the strong disagreement between our estimates of $\sqrt{\beta_1}$ and β_2 and those given by Shapiro and Wilk. According to our data, skewness and kurtosis both reduce in magnitude as n increases from 10, reaching approximately normal values at about $n = 750$ and then diverging. Shapiro and Wilk give $\sqrt{\beta_1} = -3{\cdot}32$, $\beta_2 = 212{\cdot}43$ at $n = 50$, and still rising sharply. We feel that their values must be wrong, and perhaps that rounding error spoiled their calculations. A glance at the empirical cumulative distribution of W (Fig. 1, and Shapiro and Wilk, Fig. 4) confirms that W is becoming *less* skew as n goes from 5 to 50.

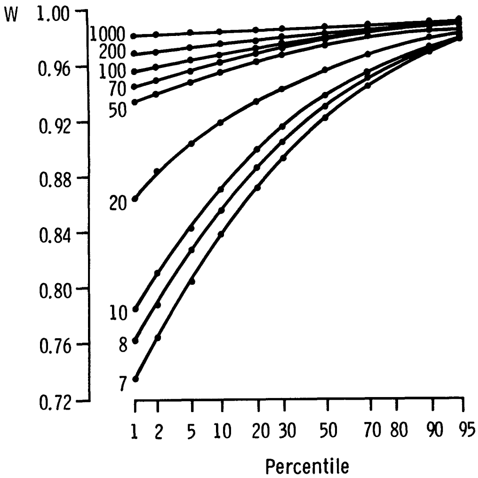

FIG. 1. Normal probability plot of the empirical cumulative distribution function of W, with sample sizes 7, 8, 10, 20, 50, 70, 100, 200, 1000.

4. Approximations for Significance Level of W

Shapiro and Wilk (1965) used Johnson's (1949) S_B system to approximate the null distribution of W, viz.

$$z = \gamma + \delta \log\{(W-\varepsilon)/(1-W)\},$$

where z is a standard normal deviate and γ, δ, ε are constants for each n. This transformation provided an apparently adequate fit for the percentage points of W in the range $n = 7(1)50$. The distribution of W is not asymptotically normal (De Wet and Venter, 1972), but our simulation results indicate that the distribution appears to pass quite close to normality around $n = 750$ (Table 1). To reflect this behaviour, we selected the following approximate normalizing transformation for W:

$$y = (1-W)^{\lambda} \quad \text{and} \quad z = (y-\mu_y)/\sigma_y \tag{4.1}$$

where z is a standard normal deviate, and μ_y and σ_y are the mean and s.d. of y. Equation (4.1) allows one to calculate the significance level of the original W. *Large* values of z (as opposed to small values of W) indicate departure from normality. The quantities λ, μ_y and σ_y are all functions of n.

Estimates of λ were obtained as follows. For each n, the correlation coefficient between selected empirical quantiles of $(1-W)^{\lambda}$ (the 0·1, 0·5, 1, 2, 3, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95 and 99 percentage points) and their normal equivalent deviates was maximized with respect to λ using simple linear search. Differential weighting according to the variance of a normal quantile was applied. Since the definition of the defining coefficients $\{a_i\}$ changes at $n = 20$, a "kink" in the function $\lambda = \lambda(n)$ was anticipated; therefore the number of simulated values of W used to estimate λ was doubled to 12 000 for $n = 7(1)14$, trebled for $n = 15(1)17$ and $21(1)23$, and quadrupled for $n = 18(1)20$. A slight minimum for λ appears at $n = 17$ and a sudden downward step at $n = 21$ (Fig. 2). The estimates of λ were smoothed using separate polynomials in $\log(n)-d$, where $d = 3$ for $n =$ 7–20 and $d = 5$ for $n =$ 21–2000 (Table 2).

The parameters of the distribution of transformed W, μ_y and σ_y were estimated as the mean and s.d. of the y's, recalculated using smoothed values of λ (Fig. 3). Their logarithms were smoothed with polynomials in $\log(n)-d$, again in two parts, $n \leqslant 20$ and $n > 20$ (Table 2).

TABLE 2

Polynomial coefficients for calculation of λ, $\log_e(\mu_y)$ and $\log_e(\sigma_y)$ for the normalizing transformation $y = (1-W)^{\lambda}$. Values of $z = (y-\mu_y)/\sigma_y$ have standard normal distribution if the original sample on which W was calculated is normal; large values of z indicate non-normality. The polynomials are calculated as $\Sigma c_i(\log n - d)^i$, where $d = 3$ for $n \leqslant 20$, $d = 5$ for $21 \leqslant n \leqslant 2000$

		Coefficients						
Parameter	*Range of n*	0	1	2	3	4	5	6
λ	7–20	0·118898	0·133414	0·327907				
	21–2000	0·480385	0·318828	0	−0·0241665	0·00879701	0·002989646	
$\log_e(\mu_y)$	7–20	−0·37542	−0·492145	−1·124332	−0·199422			
	21–2000	−1·91487	−1·37888	−0·04183209	0·1066339	−0·03513666	−0·01504614	
$\log(\sigma_y)$	7–20	−3·15805	0·729399	3·01855	1·558776			
	21–2000	−3·73538	−1·015807	−0·331885	0·1773538	−0·01638782	−0·03215018	0·003852646

TABLE 3

Coefficients for the calculation of u_3 from u_n, $n = 4, 5, 6$. u_3 may then be converted to a significance level for W (see text)

Sample size, n	*Range of u_n*	a_0	a_1	a_2	a_3	a_4
4	−3·8 to 1·4†	−1·26233	1·87969	0·0649583	−0·0475604	−0·0139682
	1·4 to 8·6‡	−0·287696	1·78953	−0·180114	0	0
5	−3·0 to 1·4†	−2·28135	2·26186	0	0	−0·00865763
	1·4 to 5·8‡	−1·63638	5·60924	−3·63738	1·08439	0
6	−1·0 to 1·4†	−3·30623	2·76287	−0·83484	1·20857	−0·507590
	1·4 to 5·4‡	−5·991908	21·04575	−24·58061	13·78661	−2·835295

† $u_3 = a_0 + a_1 u_n + a_2 u_n^2 + a_3 u_n^3 + a_4 u_n^4$.
‡ $u_3 = \exp(a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4)$; $x = \log_e(u_n)$.

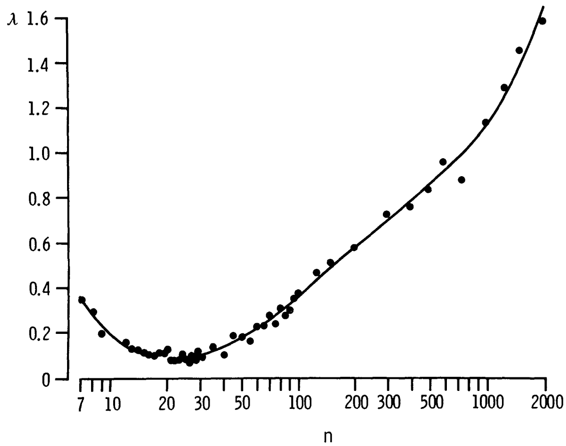

FIG. 2. Power, λ, required in the transformation of W to normality, as a function of sample size. $y = (1-W)^{\lambda}$. Abscissa is sample size (log scale). Solid line represents smoothing polynomials, with a break at $n = 20$.

Operational example

Suppose a sample of size $n = 137$ yields $W = 0{\cdot}9620$. Referring to Table 2, $x = \log(137) - 5 = -0{\cdot}080$. Then

$$\begin{aligned}
\lambda &= 0{\cdot}480385 + (0{\cdot}318828)(-0{\cdot}080) + \ldots + (0{\cdot}002989646)(-0{\cdot}080)^5 \\
&= 0{\cdot}4549, \\
\log \mu_y &= -1{\cdot}91487 + (-1{\cdot}37888)(-0{\cdot}080) + \ldots + (-0{\cdot}01504614)(-0{\cdot}080)^5 \\
&= -1{\cdot}8049, \\
\mu_y &= 0{\cdot}1645, \\
\log \sigma_y &= -3{\cdot}73538 + (-1{\cdot}015807)(-0{\cdot}080) + \ldots + (0{\cdot}003852646)(0{\cdot}080)^6 \\
&= -3{\cdot}6563, \\
\sigma_y &= 0{\cdot}02583.
\end{aligned}$$

So

$$\begin{aligned}
z &= (y - \mu_y)/\sigma_y \\
&= [(1 - 0{\cdot}9620)^{0{\cdot}4549} - 0{\cdot}1645]/0{\cdot}02583 \\
&= 2{\cdot}38 \quad (p < 0{\cdot}01).
\end{aligned}$$

The sample would be judged significantly non-normal at the 1 per cent level.

Effectiveness of the normalizing transformation

Adequate transformations of W were achieved for $n \geqslant 7$, as judged by the skewness and kurtosis of y (Fig. 4). The distributions of y are slightly left-skewed and leptokurtic; over the 50 sample sizes, the mean (s.d.) of $\sqrt{b_1}$ was $-0{\cdot}012$ ($0{\cdot}032$), and of b_2 was $3{\cdot}081$ ($0{\cdot}066$). Approximate 95 per cent ranges for $\sqrt{b_1}$ and b_2 in normal samples of size 6000 are $0 \pm 0{\cdot}063$ and $3 \pm 0{\cdot}126$ respectively.

The smallest sample sizes

Unfortunately, the power transformation of W was considered inadequate for $3 < n < 7$. Wilk and Shapiro (1968) tabulated selected values of $u = \log\{(W - \varepsilon_n)/(1-W)\}$ against W and the standard normal distribution, where $\varepsilon_n = a_1^2\, n/(n-1)$ is the minimum attainable value of W

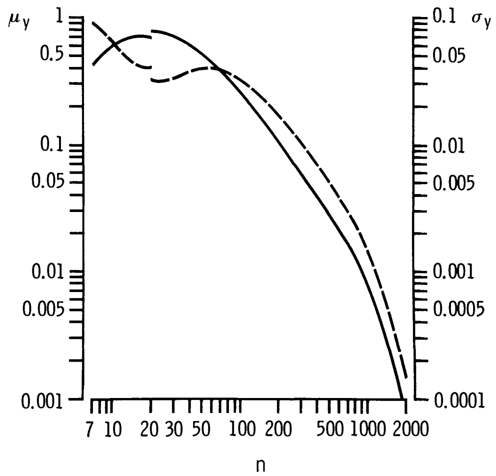

FIG. 3. Mean (solid line) and s.d. (broken line) of $y = (1-W)^{\lambda}$. Values are for smoothed values of λ. All scales are logarithmic.

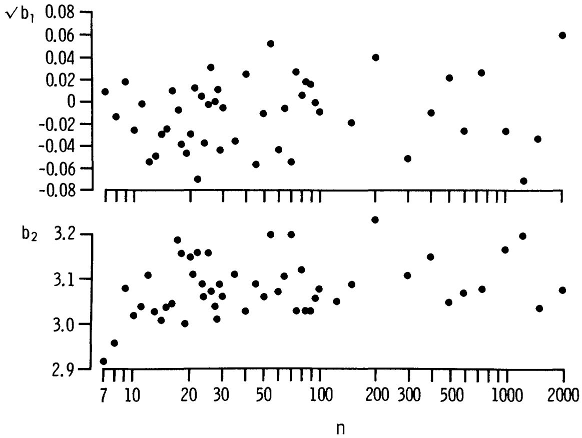

FIG 4. Skewness and kurtosis of y. Each point represents 6000 simulated values of y.

for sample size n (Shapiro and Wilk, 1965). For $n = 3$ the exact probability density function of W is known (Shapiro and Wilk, 1965, p. 595) to be

$$f(W) = (3/\pi)(1-W)^{-\frac{1}{2}}W^{-\frac{1}{2}}, \quad \tfrac{3}{4} \leqslant W \leqslant 1 \tag{4.2}$$

The corresponding cumulative distribution function is

$$p = F(W) = (6/\pi)(\sin^{-1}\sqrt{W} - \sin^{-1}\sqrt{\tfrac{3}{4}}). \tag{4.3}$$

In terms of u, with $\varepsilon = \frac{3}{4}$ for $n = 3$, one obtains

$$p = \frac{6}{\pi}\left\{\sin^{-1}\sqrt{\left(\frac{\exp(u)+\frac{3}{4}}{1+\exp(u)}\right)} - \sin^{-1}\sqrt{\tfrac{3}{4}}\right\}. \tag{4.4}$$

Wilk and Shapiro (1968, Table 1) give values of $u_n (n = 3, 4, 5, 6)$ and corresponding standard normal quantiles Z_n obtained by simulation. Given Z_n, one can calculate an equivalent value of u_3 via the standard normal integral, W_3 ((4.3) inverted) and

$u_3 = \log[(W_3 - \frac{3}{4})/(1-W_3)]$. This provides sets of paired values (u_n, u_3) for a range of p values at each sample size $n = 4, 5, 6$.

One then finds suitable functions to predict u_3 from u_n. Given an observed W whose p value is required, the sequence of operations is

$$W \to u_n \to u_3 \to p.$$

The last stage is accomplished by (4.4).

Polynomials were used to fit u_3 to u_n (Table 3). The range of application is restricted to approximately $0{\cdot}0002 < p < 0{\cdot}9998$.

As an example, suppose $n = 5$, $W = 0{\cdot}690$. For $n = 5$, $\varepsilon = 0{\cdot}5523$, so

$$u_5 = \log\frac{0{\cdot}690 - 0{\cdot}5523}{1 - 0{\cdot}690} = -0{\cdot}8115,$$

$$\begin{aligned} u_3 &= -2{\cdot}28135 + 2{\cdot}26186(-0{\cdot}8115) + (-0{\cdot}00865763)(-0{\cdot}8115)^4 \\ &= -4{\cdot}121, \end{aligned}$$

$$\begin{aligned} p &= \frac{6}{\pi}\left\{\sin^{-1}\sqrt{\left(\frac{\exp(-4{\cdot}121) + \frac{3}{4}}{1 + \exp(-4{\cdot}121)}\right)} - \sin^{-1}\sqrt{\tfrac{3}{4}}\right\} \\ &= 0{\cdot}00883. \end{aligned}$$

This result is close to the value $p = 0{\cdot}0088$ obtained by linear interpolation in Wilk and Shapiro's Table 1.

5. Application

Besides its role as a test of normality, W may be put to work in the field of estimation and data transformation. Quite frequently, one wishes to transform data to (approximate) normality, and in so doing values for several parameters may have to be found. For example, the Johnson S_B transformation, which aims to generate a standard normal deviate z using

$$z = \gamma + \delta y,$$

where

$$y = \log[(x - \xi)/(\xi + \lambda - x)]$$

necessitates estimation of the four parameters γ, δ, ξ and λ. Several methods (e.g. maximum likelihood, method of moments) have been devised to do this. An alternative way to maximize the normality of the variable y is to find values of ξ and λ which maximize W for the y's. In general, this procedure has the advantages of reducing by two the number of parameters to be

Table 4

Values of c in log(*weight* − *c*) *as estimated by the method of maximum W* (*MMW*). *Data from Rona and Altman* (*1977*)

Age (years)	Number of boys	Raw data $\sqrt{b_1}$	Raw data b_2	Transformed data $\sqrt{b_1}$	Transformed data b_2	Transformed data W	Estimates of c: MMW	Estimates of c: Rona and Altman
5	570	0·91	4·83	−0·01	3·29	0·9866	9·1	8
6	606	0·59	3·86	0·01	3·19	0·9851	5·2	8
7	659	1·06	5·28	−0·06	3·68	0·9900	12·2	12
8	612	0·90	4·59	0·02	3·48	0·9844	10·9	12
9	616	1·60	7·55	−0·05	3·99	0·9850	15·8	16
10	611	1·44	5·58	−0·04	4·19	0·9834	17·8	16
11	309	1·39	6·09	0·02	3·83	0·9777	15·2	16

estimated (since W is scale and origin independent), and of providing an immediate measure of the success of the normalizing transformation (via the p value for W). As an illustration, we now apply this "method of maximum W" (MMW) to data for which the three-parameter lognormal distribution was an adequate fit.

In a study of children's height, weight and triceps skinfold, Rona and Altman (1977) calculated constants c such that

$$\log(\text{weight} - c)$$

was reasonably normally distributed. A different value of c was obtained for each one-year age group in the range 5–12, for boys and girls separately. Using the raw data (by courtesy of the authors), we give the estimated values of c for boys aged 5–12 according to the MMW in Table 4. The values of c were found by simple one-dimensional search. For all age groups, the residual skewness is very small, but some leptokurtosis ($b_2 > 3$) remains (not platykurtosis, as Rona and Altman incorrectly state). The values of c agree well with the rounded values chosen by Rona and Altman, whose estimates were a subjective compromise based on $\sqrt{b_1}$ and b_2 for the transformed data.

The W method is easy to apply and, in this example, produces sensible results.

6. Concluding Remarks

The emphasis in this paper has been on developing the W test in a form that can easily be programmed on a computer, for the entire chosen range of sample sizes (3–2000). As far as I know, there is no other omnibus test of normality which can be calculated and assigned an adequately exact level of statistical significance without recourse to some table. D'Agostino (1971), explaining why he considered it unwise to try to extend W to large samples, said "Each sample size requires a new set of weights for calculating W. The proliferation of tables is obvious and undesirable . . . and there would still be the uninviting problem of finding the appropriate null distribution of W." We have circumvented these problems by extending Shapiro and Wilk's (1965) approximations for the weights and by providing a good normalizing transformation for W up to sample size 2000.

We aim to publish an algorithm embodying these methods for calculating W and its significance level in this *Journal* in due course.

References

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

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

D'Agostino, R. B. (1971). An omnibus test of normality for moderate and large sample sizes. *Biometrika*, **58**, 341–348.

D'Agostino, R. B. and Pearson, E. S. (1973). Tests for departure from normality. Empirical results for the distributions of b_2 and $\sqrt{b_1}$. *Biometrika*, **60**, 613–622.

Davis, C. S. and Stephens, M. A. (1978). Algorithm AS 128. Approximating the covariance matrix of normal order statistics. *Appl. Statist.*, **27**, 206–212.

De Wet, T. and Venter, J. H. (1972). Asymptotic distributions of certain test criteria of normality. *S. Africa Statist. J.*, **6**, 135–150.

Filliben, J. J. (1975). The probability plot correlation coefficient test for normality. *Technometrics*, **17**, 111–117.

Harter, H. L. (1961). Expected values of normal order statistics. *Biometrika*, **48**, 151–165.

Johnson, N. L. (1949). Systems of frequency curves generated by methods of translation. *Biometrika*, **36**, 149–176.

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

Rona, R. J. and Altman, D. G. (1977). National Study of Health and Growth: standards of attained height, weight and triceps skinfold in English children 5 to 11 years old. *Ann. Hum. Biol.*, **4**, 501–523.

Royston, J. P. (1982). Algorithm AS 177. Expected normal order statistics (exact and approximate). *Appl. Statist.*, **31**, 161–165.

Sarhan, A. E. and Greenburg, B. G. (1956). Estimation of location and scale parameters by order statistics from singly and doubly censored samples. Part I. *Ann. Math. Statist.*, **27**, 427–451.

Shapiro, S. S. and Francia, R. S. (1972). An approximate analysis of variance test for normality. *J. Amer. Statist. Ass.*, **67**, 215–216.

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

—— (1968). Approximations for the null distribution of the *W* statistic. *Technometrics*, **10**, 861–866.

Shapiro, S. S., Wilk, M. B. and Chen, H. J. (1968). A comparative study of various tests of normality. *J. Amer. Statist. Ass.*, **63**, 1343–1372.

Wichmann, B. A. and Hill, I. D. (1982). Algorithm AS 183. An efficient and portable pseudo-random number generator. *Appl. Statist.*, **31**, 188–190.

Wilk, M. B. and Shapiro, S. S. (1968). The joint assessment of normality of several independent samples. *Technometrics*, **10**, 825–839.