Tests for Departure from Normality. Empirical Results for the Distributions of b_2 and $\sqrt{b_1}$

Biometrika (1973), **60**, 3, *p.* 613
With 2 text-figures
Printed in Great Britain

Tests for departure from normality. Empirical results for the distributions of b_2 and $\sqrt{b_1}$

By RALPH D'AGOSTINO
Boston University
AND
E. S. PEARSON
University College London

SUMMARY

This paper is a preliminary to a detailed survey of the relative powers of a number of omnibus and directional tests of nonnormality. The probability integrals of $\sqrt{b_1}$ and b_2, the standardized third and fourth moment statistics, are found for random samples from a normal distribution. Main attention is given to b_2. Extensive computer simulation and curve fitting have been used to provide charts of probability levels out to the 0·1 % point, for $20 \leqslant n \leqslant 200$. For $\sqrt{b_1}$, the parameters of Johnson's symmetrical S_U approximation are tabled for values of n between 8 and 1000. An illustration is given of two 'omnibus' tests applying the charts and table, involving the joint use of $\sqrt{b_1}$ and b_2.

Some key words: Testing for departure from normality; Omnibus tests; Directional tests; Skewness and kurtosis; Third and fourth moment statistics; Simulation.

1. INTRODUCTION

The earliest work on testing for departure from normality was based on the distributions of the standardized 3rd and 4th moments,

$$\sqrt{b_1} = m_3/m_2^{\frac{3}{2}}, \quad b_2 = m_4/m_2^2$$

in random samples of n observations from a univariate normal population, where for a sample $x_1, \ldots, x_n$ of mean m_1, we define $m_r = \Sigma(x_i - m_1)^r/n$. The null hypothesis distributions used were approximate and apart from Geary's (1947*a*, *b*) work for very small samples, approximation has been more and more uncertain the smaller is n and the further it is taken out into the tails of the distributions. For example, Pearson (1965) gave only 5 and 1 % points for $\sqrt{b_1}$ with $n \geqslant 25$ and b_2 with $n \geqslant 50$.

The paper of Shapiro & Wilk (1965) has stimulated fresh interest in the subject. Their test criterion, W, is based on an entirely new principle involving the ratio of two estimates of variance and makes no direct use of the sample third and fourth moments. Computer simulations by Shapiro, Wilk & Chen (1968) and Chen (1971), using samples from a wide variety of nonnormal populations suggested that their criterion compared favourably as regards power with a number of alternative tests, including two-tail tests using $\sqrt{b_1}$ and b_2 separately. The test criterion, W, in fact, seems to provide what might be termed a very good 'omnibus' test. The original test covered the range $3 \leqslant n \leqslant 50$. Shapiro & Francia (1972) have shown how the test may be extended on certain assumptions to larger samples, using an analogous criterion W'. Here, if $x_{(i)}$ are the ordered observations,

$$W' = \left\{\sum_{i=1}^{n} b_i x_{(i)}\right\}^2 \Big/ \left\{\sum_{i=1}^{n} (x_i - \bar{x})^2\right\},$$

where the b_i depend only on the expected values of the normal order statistics (Harter, 1961; Pearson & Hartley, 1972, Table 9).

It has seemed to us that a rather fuller, comparative investigation would be valuable, involving also the power of 'directional' tests; these it may be appropriate to use if, as will sometimes be the case, prior knowledge suggests, for example, that if a population is not normal it is likely to be skew with the longer tail of its distribution in a specified direction. To carry out this survey it has been necessary to derive as reliable an approximation as possible to the probability integrals of $\sqrt{b_1}$ and b_2, for samples as small as $n = 20$, and going out not merely to the 1 % but to the 0·1 % points. Some work on these lines involving computer simulations had already been published by one of us (D'Agostino & Tietjen, 1971).

The present paper is primarily concerned with outlining the procedure leading to the two charts, Figs. 1 and 2, from which an empirical probability integral for b_2 can be derived for $20 \leqslant n \leqslant 200$. A brief statement is also included on the approximation to the distribution of $\sqrt{b_1}$ already described elsewhere. Finally we illustrate the combined use of these two statistics in two 'omnibus' tests comparable with those based on W and W'.

Work, now almost complete, will provide a critical study of the relative power of a variety of 'omnibus' and 'directional' tests of departure from normality for samples of 20, 50 and 100 using fresh simulation samples, drawn mainly from the same populations as were considered by Shapiro *et al.* (1968) and Chen (1971).

2. EMPIRICAL CUMULATIVE DISTRIBUTION FUNCTION FOR b_2 IN NORMAL SAMPLES

Because we could not derive the true mathematical distribution of b_2, it was clear that the foundation of our investigation must be some very extensive computer simulations; these were particularly necessary in the lower, abrupt tails of the distributions and, to a lesser extent, in the fine, upper tails. For a large stretch of the central portion less extensive simulation showed that use might be made of curves of the Johnson S_U or Pearson systems, as most appropriate, having the correct first four moments. The simulations were for samples of size $n = 20(5)50$, 100, 200. Full details will be given in a Report issued by the Mathematics Department of Boston University, but we give in Table 1 a summary of the simulations involved. Several points should be noted:

(*a*) At $n = 20$ the point $\beta_1(b_2) = 3{\cdot}019$, $\beta_2(b_2) = 8{\cdot}540$, falls into the Johnson S_B area; we did not attempt to provide a 4-moment graduating curve because here the simulation results were sufficiently numerous and smooth for a graduating curve to be drawn with the help of a spline.

(*b*) For $n = 25(5)50$, S_U curves were drawn having the correct first four moments, the parameters being determined as suggested by Johnson (1965). While failing at the lower tail and to a lesser extent at the upper tail of the cumulative distribution curve, $P = F(b_2|n)$, they fitted the simulation points very well in the main central position.

(*c*) To draw a smooth curve, combining the simulation and S_U data, needed the exercise of personal judgement which clearly, in places, must have involved some error.

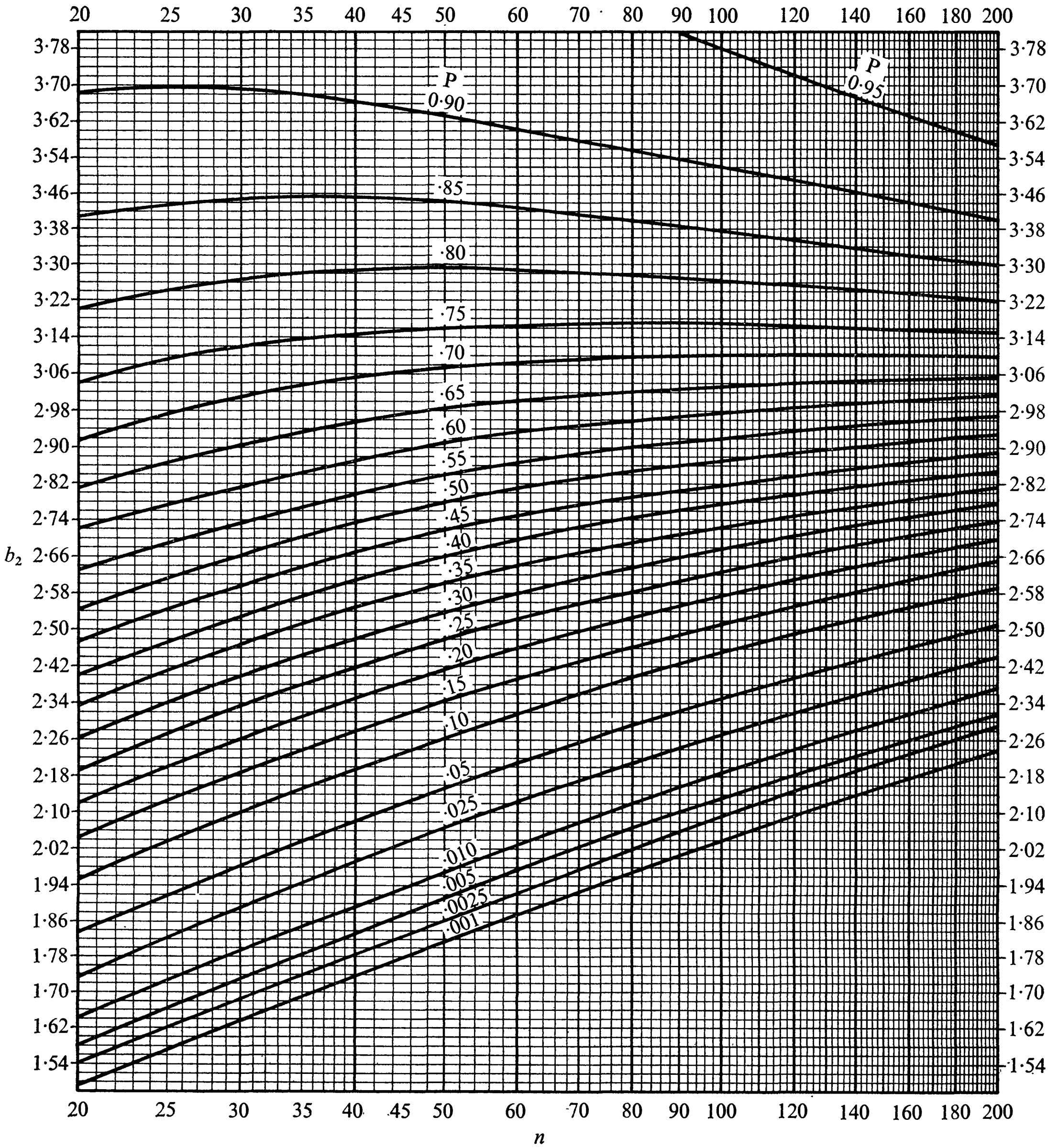

Fig. 1. Empirical probability integral, P, of the distribution of b_2 in samples of n observations from a normal population.

(*d*) For $n = 100$ and 200, Type IV Pearson curves with the correct first four moments provided better fits than S_U curves to the available simulation points. The drawing of these Pearson curves involved the use of the standardized percentage points, now given to 4 decimal places in Table 32 of Pearson & Hartley (1972); to these D. E. Amos & S. L. Daniel, in an unpublished Sandia Laboratory report, have added the extreme points at $P = 0{\cdot}001$ and $0{\cdot}999$.

(*e*) To ease the splining, some intermediate points on the $P = F(b_2|n)$ curves were obtained using 5-point Lagrangian interpolation, for which appropriate coefficients are provided by Pearson & Hartley (1972, Table 69).

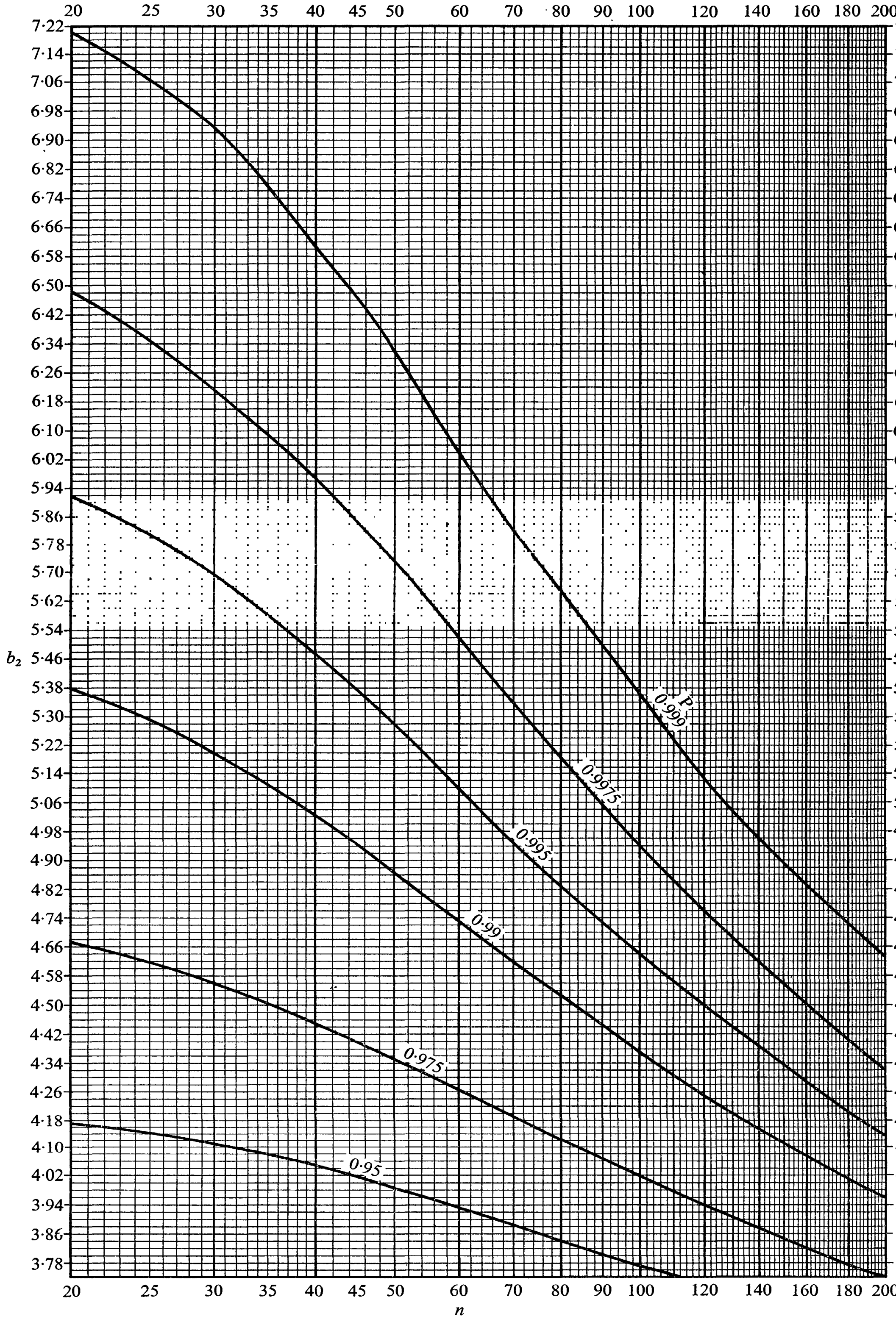

Fig. 2. Empirical probability integral, P, of the distribution of b_2 in samples of n observations from a normal population.

(*f*) Note, as shown in Table 1, that for $n = 30$, 100 and 200 there were no simulations in the central portions of the cumulative curves, because in the last two cases the amount of machine time that would have been involved seemed prohibitive. But as the Pearson curves fitted the upper tail simulation points excellently and slid smoothly into the curves indicated by the lower tail simulations, it was assumed that they might be used in the central portion.

Table 1. *Summary table of simulations for the distribution of* b_2

Lower tail			Central portion	Upper tail	
Sample size	Range covered by simulation	No. of samples	No. of samples	Range covered by simulation	No. of samples
20	$P \leqslant 0{\cdot}0176$	110,000	20,000	$P \geqslant 0{\cdot}9967$	110,000
25	$P \leqslant 0{\cdot}0120$	55,000	20,000	$P \geqslant 0{\cdot}9965$	55,000
30	$P \leqslant 0{\cdot}0072$	55,000	None	$P \geqslant 0{\cdot}9956$	55,000 or more
35	$P \leqslant 0{\cdot}00415$	40,000	10,000	$P \geqslant 0{\cdot}9978$	40,000
40	$P \leqslant 0{\cdot}01$	50,000	10,000	$P \geqslant 0{\cdot}90$	50,000
45	$P \leqslant 0{\cdot}0106$	40,000	10,000	$P \geqslant 0{\cdot}9966$	40,000
50	$P \leqslant 0{\cdot}0137$	52,000	24,000	$P \geqslant 0{\cdot}9916$	39,000
100	$P \leqslant 0{\cdot}0181$	40,000	None	$P \geqslant 0{\cdot}9896$	40,000
200	$P \leqslant 0{\cdot}0127$	16,000	None	$P \geqslant 0{\cdot}9857$	16,000

3. Method of computation

The computer used was the IBM 360, model 50, at Boston University. The simulation required the generation of normal deviates. This was accomplished by first generating a uniform deviate using the IBM subroutine RANDU and then transforming this to normality using an approximation due to Hastings [Abramowitz & Stegun, 1964, p. 933, formula (26·2·23)]. It is believed that this method should be more than adequate to provide the normal deviates required.

Where for b_2, and later for $\sqrt{b_1}$, the simulation frequencies in the central portion of the distributions have not been actually recorded, the complete simulation was nevertheless carried out. Where the samples were large, only the observations in the extreme tails of the b_2 and $\sqrt{b_1}$ distributions were ordered in most cases. This resulted in a great saving in computer time since it meant that only a small fraction of the sampling distribution had to be manipulated.

4. Derivation of cumulative probability curves

From the nine cumulative probability curves, each for a different value of n, values of b_2 were read off corresponding to $P = 0{\cdot}001$, 0·0025, 0·005, 0·01, 0·025, 0·05, 0·10, 0·25, 0·50, 0·75, 0·90, 0·95, 0·975, 0·99, 0·995, 0·9975 and 0·999. As a check on smoothness, first differences of b_2 were taken in the n direction, for $n = 20(5)50$, at each of these 17 P-values. A few irregularities were then removed where this seemed justifiable on re-examination of the original $P = F(b_2|n)$ curves.

To fill in gaps it was found easiest to plot b_2 against $\log n$, and the abscissa-scale in Figs. 1 and 2 is a single-cycle logarithmic one. There was still a long gap between $n = 50$ and 100, and between $n = 100$ and 200; to avoid bridging this entirely by splining we proceeded as follows. Since Pearson curves with the correct first four moments appeared to give a good fit at $n = 50$, 100 and 200 in the range $0 \cdot 02 < P < 0 \cdot 98$, the seventeen percentage points were also calculated at:

n	Mean	S.D.	$\sqrt{\beta_1}$	β_2
75	2·92105	0·512340	1·409913	7·493339
150	2·96026	0·380586	1·091706	5·825808

With an adjustment at the tails suggested by the curves with neighbouring values of n, the resulting points (b_2, P) lay almost exactly on the curves which had already been obtained using the spline. This result seemed to justify a very reasonable confidence in the adequacy of the procedure on which Figs. 1 and 2 were built, although we make no claim to perfection. The most likely errors are in the very high P curves, but here the curves are so far separated that errors become of less importance.

Finally, it will be noted that we have inserted curves on the charts for $P = 0 \cdot 15$, 0·20, 0·30, 0·35, 0·40, 0·45, 0·55, 0·60, 0·65, 0·70, 0·80 and 0·85. The appropriate values were obtained by interpolation between the 17 standard values, using the 5-point Lagrangian tables referred to above.

To judge the significance of a single b_2 value the curves are clearly more than adequate, but it is also possible to obtain from them reasonable approximations to the probability integral, P, given b_2. The following examples illustrate this point.

Example 1. Suppose that $n = 82$ and $b_2 = 2 \cdot 17$, what is $\mathrm{pr}\,(b_2 \leqslant 2 \cdot 17 | n = 82)$?

Put a rule vertically along the grid line in Fig. 1 for $n = 82$. Then use of a horizontal rule, if possible transparent, shows that this vertical cuts four P-curves as follows:

P	0·005	0·01	0·025	0·05
b_2	2·07	2·13	2·21	2·30

Plot b_2 as ordinate against P as abscissa on finely squared paper. Join the four points with the help of a 'French curve'; it is found that $\mathrm{pr}\,(b_2 \leqslant 2 \cdot 17) \doteq 0 \cdot 0160$. Any adequate table of the normal probability integral shows that $X(b_2) \doteq -2 \cdot 14$.

Example 2. Again, suppose $n = 82$ and $b_2 = 4 \cdot 621$. What is $\mathrm{pr}\,(b_2 \geqslant 4 \cdot 621 | n = 82)$?

From Fig. 2 we obtain the following:

P	0·975	0·990	0·995	0·9975
b_2	4·11	4·51	4·81	5·15

Again, after plotting and joining the four points with a smooth curve, we find

$$\mathrm{pr}\,(b_2 \leqslant 4 \cdot 621) \doteq 0 \cdot 992, \quad \mathrm{pr}\,(b_2 \geqslant 4 \cdot 621) \simeq 0 \cdot 008.$$

Hence from tables of the normal probability integral the equivalent normal deviate is $X(b_2) \doteq 2 \cdot 41$.

Of course if the (b_2, n) point falls in one of the extreme regions of the figures, i.e. between the 0·001 and 0·0025 or between the 0·9975 and 0·999 curves, this graphical interpolation will be less accurate, but in any case the extreme (P, b_2) curves are not here too reliable.

5. The probability integral of $\sqrt{b_1}$

We recapitulate here an approximation using a symmetrical Johnson S_U curve which was first suggested by one of us (Pearson, 1965) and more fully investigated by the other (D'Agostino, 1970). In the latter paper comparisons were made between the fraction of values of $|\sqrt{b_1}|$ in 20,000 simulated normal samples falling beyond the two-tailed 0·01, 0·02, 0·05, 0·10 and 0·20 points derived from the S_U distribution having the correct first four moments. The sample sizes taken were $n = 8$, 10, 15 and 20. The agreement was good, the fractions, α, never differing by more than 0·002 at $n = 20$.

To reinforce this work, further simulations were undertaken. Table 2 shows comparisons at the 1 % point and further out. It will be seen that except at the 0·1 % point the S_U and simulation figures never differ by more than 0·01 in terms of $\sqrt{b_1}$ as unit. In Table 3 we present results covering the full range of the distribution for samples with $n = 100$. Here the body of

Table 2. *Comparison of simulation and S_U approximation percentage points of $\sqrt{b_1}$ in the tails of its sampling distribution*

No. of samples ...	11,250		21,000		15,000		15,000		23,000		10,000 (8,000)	
n	20		25		30		35		50		100	
α	Sim	S_U	Sim	S_U	Sim	S_U	Sim	S_U	Sim	S_U	Sim	S_U
0·010	1·15	1·15	1·07	1·06	0·98	0·98	0·93	0·92	0·78	0·79	0·56	0·57
0·005	1·30	1·30	1·22	1·20	1·11	1·12	1·04	1·04	0·88	0·89	0·65	0·64
0·0025	1·44	1·45	1·35	1·34	1·23	1·24	1·17	1·16	0·98	0·99	0·71	0·70
0·0010	1·61	1·65	1·47	1·52	1·41	1·41	1·33	1·32	1·09	1·12	0·80	0·79

The necessarily symmetrical distributions have been 'doubled over' so that the frequencies so obtained are divisible by twice the number of samples. As the sampling frequencies were recorded at values of $\sqrt{b_1}$ increasing by units of 0·01, an interpolation process was needed to estimate the position of the percentage points.

Table 3. *Comparison of simulation and S_U approximation for upper tail probabilities, Q, of $\sqrt{b_1}$ in samples of $n = 100$*

		$Q = 1-P$				$Q = 1-P$				$Q = 1-P$	
$\sqrt{b_1}$	N	Sim	S_U	$\sqrt{b_1}$	N	Sim	S_U	$\sqrt{b_1}$	N	Sim	S_U
0·02		0·460	0·465	0·30	8,000	0·1007	0·1001	0·58	10,000	0·0085	0·0088
0·06		0·398	0·397	0·34	8,000	0·0733	0·0743	0·62	21,000	0·0060	0·0059
0·10		0·334	0·332	0·38		0·0533	0·0541	0·66	21,000	0·0044	0·0039
0·14	8,000	0·274	0·272	0·42		0·0384	0·0388	0·70		0·0025	0·0026
0·18		0·221	0·219	0·46	10,000	0·0266	0·0273	0·74		0·0017	0·0017
0·22		0·172	0·172	0·50		0·0193	0·0190	0·78	19,000	0·0012	0·0011
0·26		0·134	0·132	0·54		0·0128	0·0130	0·82		0·0008	0·0007

N = number of samples of $n = 100$ drawn from a normal population. The observed numbers of samples giving $\sqrt{b_1}$ below a given negative value and above the corresponding positive value were added and the sum divided by $2N$ to obtain the simulation estimate of Q.

the table gives the values of $Q = 1-P$, where P is the probability integral, for simulation samples and for the S_U distribution, with $\sqrt{b_1} = 0{\cdot}02(0{\cdot}04)0{\cdot}82$. Remembering that ($a$) the standard errors of the simulation estimates are such that true differences may be obscured, and (b) for a given sample size, the simulation values are to some extent correlated, we consider that these results are such as to justify the use of the S_U approximation.

The approximation is represented by the transformation

$$X(\sqrt{b_1}) = \delta \sinh^{-1}(\sqrt{b_1}/\lambda), \tag{1}$$

where X is a standardized normal deviate, with zero mean and unit standard deviation. The parameters δ and $1/\lambda$ are functions of the variance and $\beta_2(\sqrt{b_1})$ in samples from a normal population which, in turn, are simply functions of the sample size n. To avoid reference to tables of $\sinh^{-1} x$, D'Agostino (1970) presented the parameters in a form involving a natural logarithm. However, in our Table 4 we provide what we believe to be an adequate table of δ and $1/\lambda$ for $n = 8(1)50(2)100(5)250(10)500(20)1000$.

6. Two 'omnibus' tests for departure from normality based on both $\sqrt{b_1}$ and b_2, $20 \leqslant n \leqslant 200$

Two 'omnibus' tests suggest themselves, both depending on the fact that, under the null hypothesis, $\sqrt{b_1}$ and b_2 are independent and also that the charts and table given in §§ 4 and 5 above enable us to obtain approximately either

(a) standardized normal equivalent deviates, $X(\sqrt{b_1})$ and $X(b_2)$ or

(b) the probability integrals $P(\sqrt{b_1})$ and $P(b_2)$,

corresponding on the null hypothesis to any observed couplet $\sqrt{b_1}$, b_2.

First test. Having found $P(b_2|n)$ by the procedure described in § 4, the corresponding standardized normal deviate is obtainable from any appropriate tables of the normal function, i.e. $X(\sqrt{b_1})$ is found from equation (1) and Table 4, if necessary applying the small correction suggested by Table 2 for $n < 30$.

Then if the hypothesis of population normality is true

$$\chi^2(\nu=2) = X^2(\sqrt{b_1}) + X^2(b_2) \tag{2}$$

will be distributed as a χ^2 having 2 degrees of freedom, and significance may be examined in the usual way.

Second test. This was first suggested by one of us (Pearson, 1938, p. 137) and further illustrated by Neyman & Pearson (1938, pp. 51–7). If u and v are two independent random variates having densities under the null hypothesis denoted by $p(u|H_0)$ and $p(v|H_0)$, we put

$$q_1 = \begin{cases} 2\int_U^\infty p(u|H_0)\,du & \text{if } U \text{ is above its median,} \\ 2\int_{-\infty}^U p(u|H_0)\,du & \text{if } U \text{ is below its median,} \end{cases}$$

where U is the observed value of u. Also we use a similar definition for v and q_2; then under the null hypothesis

$$\chi^2(\nu=4) = -2\log(q_1 q_2) \tag{3}$$

will be distributed as a χ^2 having 4 degrees of freedom.

If $u = \sqrt{b_1}$ and $v = b_2$, then q_2 is found as in § 4 and q_1 from the $X(\sqrt{b_1})$ provided by Table 4. It follows that equation (3) provides an alternative test based jointly on the two beta criteria. These two tests are bound to produce very similar answers. Neyman & Pearson (1938, p. 54, fig. 4) showed how close together the critical contours were in the normalized $X(\sqrt{b_1})$, $X(b_2)$ space, when the test level was $\alpha = 0{\cdot}05$. We have found a very similar position for $\alpha = 0{\cdot}01$ and $0{\cdot}20$.

Any further comparative discussion of these tests and of others, both 'omnibus' and

Table 4. *Coefficients in normalizing transformation for* $\sqrt{b_1}$

n	δ	$1/\lambda$
8	5·563	0·3030
9	4·260	0·4080
10	3·734	0·4794
11	3·447	0·5339
12	3·270	0·5781
13	3·151	0·6153
14	3·069	0·6473
15	3·010	0·6753
16	2·968	0·7001
17	2·937	0·7224
18	2·915	0·7426
19	2·900	0·7610
20	2·890	0·7779
21	2·884	0·7934
22	2·882	0·8078
23	2·882	0·8211
24	2·884	0·8336
25	2·889	0·8452
26	2·895	0·8561
27	2·902	0·8664
28	2·910	0·8760
29	2·920	0·8851
30	2·930	0·8938
31	2·941	0·9020
32	2·952	0·9097
33	2·964	0·9171
34	2·977	0·9241
35	2·990	0·9308
36	3·003	0·9372
37	3·016	0·9433
38	3·030	0·9492
39	3·044	0·9548
40	3·058	0·9601
41	3·073	0·9653
42	3·087	0·9702
43	3·102	0·9750
44	3·117	0·9795
45	3·131	0·9840
46	3·146	0·9882
47	3·161	0·9923
48	3·176	0·9963
49	3·192	1·0001
50	3·207	1·0038
52	3·237	1·0108
54	3·268	1·0174
56	3·298	1·0235
58	3·329	1.0293
60	3·359	1·0348

n	δ	$1/\lambda$
62	3·389	1·0400
64	3·420	1·0449
66	3·450	1·0495
68	3·480	1·0540
70	3·510	1·0581
72	3·540	1·0621
74	3·569	1·0659
76	3·599	1·0695
78	3·628	1·0730
80	3·657	1·0763
82	3·686	1·0795
84	3·715	1·0825
86	3·744	1·0854
88	3·772	1·0882
90	3·801	1·0909
92	3·829	1·0934
94	3·857	1·0959
96	3·885	1·0983
98	3·913	1·1006
100	3·940	1·1028
105	4·009	1·1080
110	4·076	1·1128
115	4·142	1·1172
120	4·207	1·1212
125	4·272	1·1250
130	4·336	1·1285
135	4·398	1·1318
140	4·460	1·1348
145	4·521	1·1377
150	4·582	1·1403
155	4·641	1·1428
160	4·700	1·1452
165	4·758	1·1474
170	4·816	1·1496
175	4·873	1·1516
180	4·929	1·1535
185	4·985	1·1553
190	5·040	1·1570
195	5·094	1·1586
200	5·148	1·1602
205	5·202	1·1616
210	5·255	1·1631
215	5·307	1·1644
220	5·359	1·1657
225	5·410	1·1669
230	5·461	1·1681
235	5·511	1·1693
240	5·561	1·1704
245	5·611	1·1714
250	5·660	1·1724

n	δ	$1/\lambda$
260	5·757	1·1744
270	5·853	1·1761
280	5·946	1·1779
290	6·039	1·1793
300	6·130	1·1808
310	6·220	1·1821
320	6·308	1·1834
330	6·396	1·1846
340	6·482	1·1858
350	6·567	1·1868
360	6·651	1·1879
370	6·733	1·1888
380	6·815	1·1897
390	6·896	1·1906
400	6·976	1·1914
410	7·056	1·1922
420	7·134	1·1929
430	7·211	1·1937
440	7·288	1·1943
450	7·363	1·1950
460	7·438	1·1956
470	7·513	1·1962
480	7·586	1·1968
490	7·659	1·1974
500	7·731	1·1979
520	7·873	1·1989
540	8·013	1·1998
560	8·151	1·2007
580	8·286	1·2015
600	8·419	1·2023
620	8·550	1·2030
640	8·679	1·2036
660	8·806	1·2043
680	8·931	1·2049
700	9·054	1·2054
720	9·176	1·2060
740	9·297	1·2065
760	9·415	1·2069
780	9·533	1·2073
800	9·649	1·2078
820	9·763	1·2082
840	9·876	1·2086
860	9·988	1·2089
880	10·098	1·2093
900	10·208	1·2096
920	10·316	1·2100
940	10·423	1·2103
960	10·529	1·2106
980	10·634	1·2109
1000	10·738	1·2111

$$X(\sqrt{b_1}) = \delta \sinh^{-1}(\sqrt{b_1}/\lambda)$$

'directional' will be left to a second communication. We shall here merely give a brief example.

Example 3. Suppose that in a sample of $n = 82$ observations it is found that $\sqrt{b_1} = 0{\cdot}506$ and $b_2 = 4{\cdot}621$. Is there evidence of departure from normality in the population sampled?

Proceeding as suggested, we find, with P denoting the probability integral, for

$$u = \sqrt{b_1}, \quad P = 0{\cdot}0271, \quad X = 1{\cdot}9247, \quad q_1 = 0{\cdot}0542 = 2P(\sqrt{b_1});$$

for

$$v = b_2, \quad P = 0{\cdot}992, \quad X = 2{\cdot}41, \quad q_2 = 0{\cdot}0165 = 2\{1-P(b_2)\};$$

$$\chi^2(\nu=2) = 9{\cdot}51, \quad \mathrm{pr}\,(\chi^2 \geqslant 9{\cdot}51|\nu = 2) = 0{\cdot}0086, \quad \text{from (2)};$$

$$\chi^2(\nu=4) = 14{\cdot}10, \quad \mathrm{pr}\,(\chi^2 \geqslant 14{\cdot}10|\nu = 4) = 0{\cdot}0070 \quad \text{from (3)}.$$

The values corresponding to b_2 are obtained from Example 2.

The observed point $(\sqrt{b_1}, b_2)$ is therefore just outside the 1 % critical contour of both 'omnibus' tests and the difference in significance probabilities is consistent with the close proximity and crossing over of the critical contours in the $X(\sqrt{b_1})$, $X(b_2)$ plane.

REFERENCES

ABRAMOWITZ, M. & STEGUN, I. A. (Eds.) (1964). *Handbook of Mathematical Functions*. Washington: National Bureau of Standards.

CHEN, E. H. (1971). The power of the Shapiro–Wilk W test for normality in samples from contaminated normal distributions. *J. Am. Statist. Ass.* **66**, 760–2.

D'AGOSTINO, R. B. (1970). Transformation to normality of the null distribution of g_1. *Biometrika* **57**, 679–81.

D'AGOSTINO, R. B. & TIETJEN, G. L. (1971). Simulation probability points of b_2 for small samples. *Biometrika* **58**, 669–72.

GEARY, R. C. (1947*a*). Frequency distribution of $\sqrt{b_1}$. *Biometrika* **34**, 68–97.

GEARY, R. C. (1947*b*). Testing for normality. *Biometrika* **34**, 209–42.

HARTER, H. L. (1961). Expected values of normal order statistics. *Biometrika* **48**, 151–65.

JOHNSON, N. L. (1965). Tables to facilitate fitting S_U frequency curves. *Biometrika* **52**, 547–58.

NEYMAN, J. & PEARSON, E. S. (1938). Contributions to the theory of testing statistical hypotheses. *Statist. Res. Mem.* **2**, 265–97.

PEARSON, E. S. (1938). The probability integral transformation for testing goodness of fit and combining independent tests of significance. *Biometrika* **30**, 134–48.

PEARSON, E. S. (1965). Tables for percentage points of $\sqrt{b_1}$ and b_2 in normal samples; a rounding off. *Biometrika* **52**, 282–5.

PEARSON, E. S. & HARTLEY, H. O. (Eds.) (1972). *Biometrika Tables for Statisticians*, vol. 2. Cambridge University Press.

SHAPIRO, S. S. & WILK, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika* **52**, 591–611.

SHAPIRO, S. S., WILK, M. B. & CHEN, H. J. (1968). A comparative study of various tests for normality. *J. Am. Statist. Ass.* **63**, 1343–72.

SHAPIRO, S. S. & FRANCIA, R. S. (1972). Approximate analysis of variance test for normality. *J. Am. Statist. Ass.* **67**, 215–6.

[*Received February* 1973. *Revised April* 1973]