ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 48, 1985

Wiesław Wagner *

TESTS OF UNIVARIATE NORMALITY

1. Introduction

A large class of goodness of fit tests in the theory of statistical inference are tests of univariate normality. Tests of normality enable testing the goodness of fit of a sampled distribution function by the normal distribution function of a given random variable.

Historically, the tests $\sqrt{b_1}$ and b_2, built in the thirties, are the first tests for normality. Other tests were developed by Geary (test g), David et al. (test n), Kolmogorov (test D), Cramer and von Mises (test CM), Anderson and Darling (test A^2), Shapiro and Wilk (test W), D'Agostino (test Y) and others. Beside the above mentioned tests there are many modifications which are widely discussed in the paper.

Let a random variable X of a continuous type be distributed in the way determined by a distribution function F(x) and with distribution parameters $\mu = E(X)$ and $\sigma^2 = D^2(X)$. The fact that the variable X is normally distributed with parameters μ and σ^2 is denoted as $X \sim N(\mu, \sigma^2)$. Let the sequence $x_1, \ldots, x_n$ $(\{x_i\})$ denote a sample consisting of n independent observation of variable X, and $\bar{x}$ and arithmetic mean, S^2 - the sum of squared deviations, s^2 - variance, and s - standard deviation from the sample. Sample $\{y_i\}$ denotes non-decreasing ordered observations of sample $\{x_i\}$, so that $y_1 \leq \ldots \leq y_n$. Function $F_n(x)$

* Lecturer at the Department of Mathematical and Statistical Methods, Academy of Agriculture, Poznań.

is an empirical distribution function, where x is an arbitrary real number, i.e. the observation function $\leq x$, while $\Phi(u)$ and $\Phi^{-1}(p)$ are the normal distribution function and quantile of the p-th order of distribution N(0,1), respectively. For a determined x the value of $F_n(x)$ is a random variable 0, 1 (n,2) n, ..., 1.

The order statistic is a random variable y_k being the k-th variable in the sample $\{y_i\}$. The sample $\{y_i\}$ is called the - order statistic. When $X \sim N(\mu, \sigma^2)$ then y_k is a normal order statistic, $\{y_i\}$ - a sequence of normal order statistics, and $\{u_i\}$ a sequence of N(0, 1) - order statistics, where $u_i = (y_i - \mu)/\sigma$. The order statistic u_i has distribution parameters: $E(u_i) = m_i$, $D^2(u_i) = v_{ii}$, $Cov(u_i, u_j) = v_{ij}$; i, j = 1, ..., n, then for $N(\mu, \sigma)$ - order statistics we have $E(y_i) = \mu + \sigma u_i$, $D^2(y_i) = \sigma^2 v_{ii}$ and $Cov(y_i, y_j) = \sigma^2 v_{ij}$. The values of $m_i = m_{i,n}$, i = 1, ..., [n/2] are tabelarized for various ns, while $v_{ii} = v_{ii,n}$ and $v_{ij} = v_{ij,n}$ for i,j = 1,..., [n/2]; $i \neq j$ only for n = 1, ..., 20 (cf. [20], tables 9 and 10).

2. The Hypothesis of Goodness of Fit of the Empirical Distribution with the Normal Distribution

The statistical hypothesis of the empirical and normal distributions is formulated as follows. Let F_n denote a class of normal distribution function and G - a class of distribution function of random variables having the third moment (μ_3) other than zero and finite fourth central moment (μ_4), at $F_N \cap G = \emptyset$. Let us put a hypothesis $H_0 : F \in F_N$ and $H_1 : F \in G$, that the function F(x) belongs to the class of distribution functions F_N and G. The hypotheses H_0 or H_1 will be simple if the distributions belonging to classes F_N or G have the known distribution parameters.

The hypothesis H_0 against H_1 is verified using one of the tests for normality. Generally, they are divided, according to the structure of test statistics into

1) Pearson's χ^2 tests for goodness of fit,

2) tests based on the comparison of empirical and normal distribution functions,

3) tests using sample moments,

4) tests based on order statistics.

Pearson's χ^2 test will not be discussed here, since it is well-known. It is used mainly for large samples, especially when the observations are in the form of grouped data.

Next we shall consider the tests within each group and finally a general discussion will be presented.

3. Tests Based on the Comparison of Empirical and Normal Distribution Functions

These tests are based on the distance of distribution functions $\hat{F}(x) \in F_N$ and $F_n(x)$ being the estimators $F(x)$ and from the sample under the hypothesis H_0 [13]. Under the simple hypothesis $F(x)$ will be replaced by $F(x)$.

Let $F_n(y_i)$ denote the value of distribution function $F_n(x)$ in point y_i and $z_i = \Phi(u_i)$, where

$$u_i = \begin{cases} (y_i - \mu)/\sigma, & \mu, \sigma \text{ - known} & (a) \\ (y_i - \bar{y})/s, & \mu, \sigma \text{ - unknown} & (b) \\ (y_i - \mu)/\tilde{s}, & \mu \text{ - known}, \sigma \text{ - unknown} & (c) \\ (y_i - \bar{y})/\sigma, & \mu \text{ - unknown}, \sigma \text{ - known} & (d) \end{cases}$$

at $\tilde{s}^2 = \sum_{i=1}^{n}(y_i - \mu)^2/n$. The case (a) refers to the class F_N of normal distributions determined completely (a simple hypothesis), while other cases do not determine the distribution function $F(x)$ wholly (a complex hypothesis). The following tests belong to the above mentioned group:

- Kolmogorov-Smirnov

$$K = \sup_{-\infty < x < \infty} \left| F_n(x) - F(x) \right|,$$

$$KS = \sup_{1 \leq i \leq n} \left| \frac{i}{n} - z_i \right|,$$

- Kolmogorov

$$D^+ = \max_{1 \leq i \leq n} (i/n - z_i), \quad D^- = \max_{1 \leq i \leq n} (z_i - (i-1)/n)$$

$$D = \max (D^+, D^-),$$

- Kuiper

$$V = D^+ + D^-,$$

- Cramer-von Mises

$$W^2 = 1/(12n) + \sum_{i=1}^{n} (z_i - (2i-1)/2n)^2$$

$$= \sum_{i=1}^{n} z_i^2 - (1/n) \sum_{i=1}^{n} (2i-1)z_i + n/3,$$

$$W_0 = \sum_{i=1}^{n} (z_i - (2i-1)/2n)^2$$

$$W_{11} = \sum_{i=1}^{n} (z_i - i/(n+1))^2,$$

$$W_{21} = \sum_{i=1}^{n} (z_i - (2i-1)/(2(n+1)))^2,$$

- Watson

$$U^2 = W^2 - n(\overline{z} - 0.5)^2, \quad \overline{z} = \sum_{i=1}^{n} z_i/n,$$

- Anderson-Darling

$$A^2 = -n - (1/n) \sum_{i=1}^{n} (2i-1) \left[\ln z_i + \ln (1 - z_{n-i+1})\right],$$

$$A_{21} = -n - n/(n+1)^2 \sum_{i=1}^{n} \left[(2i-1) \ln z_i + \right.$$

$$\left. + (2i+1) \ln (1-z_{n-i+1})\right] - \left[(2n+1) \ln z_n - \ln(1 - z_n)\right],$$

$$A_{12} = -(n+1) - 1/(n+1) \sum_{i=1}^{n} i \left[\ln z_i + \ln(1 - z_{n-i+1})\right].$$

In the above test statistics z_i assumes the values according to the case (a)-(d). The best known tests for the simple hypothesis H_0 are D, W^2 and A^2. The above mentioned tests are given in their summation form though originally they were presented in an integral form, which we shall mention later. The tests belonging to this group have known distributions in case (a). For some tests (D, V, u, W^2, A^2) modifications are given for cases (b)-(d) [25], whose critical values do not depend on the sample size but only on the significance level α. For the KS test in case a the critical value was given, among others, by F i s z [10] (table VIII) and for (b) by L i l l i e f o r s [16].

The tests originating from Cramer-von Mises and Anderson--Darling tests were generated from general integral forms, respectively

$$W = n \int_0^1 (y - S_i)^2 dy$$

and

$$A^2 = n \int_0^1 \{(y - S_i)^2 / y(1-y)\} dy,$$

where S_i is a certain function of i and n.

The statistics W_{11}, W_{21}, A_{21} and A_{12} were obtained replacing i/n by $i/(n+1)$ (the first index) or $(i + 0.5)/(n + 1)$ (the second index). G r e e n and H e g a z y [13] introduced tests for normality of the above statistics, and proved their predominance as far as their power was concerned, over the tests W^2 and A^2 given in a summation form.

4. Tests Using Sample Moments

For the sample of n observations $\{x_i\}$ we determine a central moment of the k-th order

$$m_k = (1/n)\sum_{i=1}^{n}(x_i - \bar{x})^k, \quad k = 2, 3, \ldots$$

Statistics $\sqrt{b_1} = m_3/\sqrt{m_2^3}$ and $b_2 = m_4/m_2^2$ are unbiased estimators of parameters $\sqrt{\beta_1} = \mu_3/\sigma^3$ and $\beta_2 = \mu_4/\sigma^4$. If $X \sim N(\mu, \sigma^2)$, then $\mu_3 = 0$ and $\mu_4 = 3\sigma^4$, hence $\sqrt{\beta_1} = 0$ and $\beta_2 = 3$. This means that $\sqrt{\beta_1}$ and β_2 are equal to 0 and 3, respectively, for normally distributed random variables. Therefore distributions of variables for which parameters $(\sqrt{\beta_1}, \beta_2)$ have values close to (0,3) are treated as "almost normal". For the construction of tests for normality being discussed Slutsky's theorem is used [10]. It follows from this theorem that $\sqrt{b_1}$ and b_2 are converging to $\sqrt{\beta_1}$ and β_2 when $n \to \infty$. The parameters of distribution of these variables under H_0 are [2]

$$E(\sqrt{b_1}) = 0,$$

$$E(b_2) = \begin{cases} 3(n-1)/(n+1), \\ 0, \text{ when } n \to \infty, \end{cases}$$

$$D^2(\sqrt{b_1}) = \begin{cases} 6(n-2)/(n+1)(n+3), \\ 6/n, \text{ when } n \to \infty, \end{cases}$$

$$D^2(b_2) = \begin{cases} 24n(n-2)(n-3)/(n+1)^2(n+3)(n+5), \\ 24/n, \text{ when } n \to \infty. \end{cases}$$

To the discussed group the following tests belong:

- standardized third sample moment

$$\sqrt{b_1} = m_3/m_2^{3/2},$$

used against the hypothesis stating that the G-class distributions are skew ($\sqrt{\beta_1} \neq 0$); the critical value for $n > 25$ was given by Pearson and Hartley [19] (table 34B) and for $n \leq 25$ by Muldholland [18];

- standardized fourth central sample moment

$$b_2 = m_4/m_2^2,$$

used against H_1 stating that significant points class distributions are symmetric; a critical value for $n \geq 50$ was given by Pearson and Hartley [19];

- D'Agostino-Pearson [5]. For a given value of the above determined statistic b_2 we establish probability P $(b_2 \leq b_2(p, n)) = p$, where $b_2(p, n)$ is a critical value of b_2 distribution for given p and n (p = 0.001, 0.0025, 0.05, 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, 0.99; n = 20, 21, ..., 200; b_2 = 1.54 (0.08) 7.22). Then we determine quartile $X(b_2) = \Phi^{-1}$ p of the p-th order distribution N(0,1) using adequate tables [29] (table 3). For $\sqrt{b_1}$ statistic we have the quantile

$$X(\sqrt{b_1}) = \delta \ln\left\{\sqrt{b_1}/\lambda + [(\sqrt{b_1}/\lambda)^2 + 1]^{1/2}\right\},$$

where constants δ and $1/\lambda$ are tabelarized at n = 8(1)50, 52(2)100, 105(5)250, 260(10)500, 520(20)1000. We determine K^2 test based on the statistic

$$K^2 = X^2(\sqrt{b_1}) + X^2(b_2),$$

which, under H_0, has χ_2^2 distribution. The advantage of test

K^2 is the possibility of testing departures from normality caused by skewness and kurtosis. Such a test is called the omnibus test. At $n > 200$ instead of K^2 the statistic

$$\tilde{K}^2 = (n/24)\left[4\,(\sqrt{b_1})^2 + (b_2 - 3)^2\right]$$

is used. The variable $\tilde{K}^2$ has an asymptotic χ^2_2 distribution.

- Bowman-Shenton [1]. The variable $\sqrt{b_1}$ is symmetrically distributed, while b_2 is asymmetrically distributed. The curves used for each of these variables are as follows: for $\sqrt{b_1}$ Pearson's curves of type VII or t-Students, for b_2 Pearson's curves of type VI or IV. In both cases Johnson's transformation [15] to S_U curves gives a satisfactory normal curve for $\sqrt{b_1}$ [3]; the system of S_U curves is, however, less satisfactory for b_2, S_U curves provide good consistency for $\sqrt{b_1}$ at $n \geq 8$ and for b_2 at $n \geq 25$. For small samples Johnson's S_B system is a sufficient approximation for b_2 (e.g. for $n = 20$ it is $\beta_1(b_2) = 3.019$. and $\beta_2(b_2) = 8.54$). Hence we have test statistics

$$X_S(\sqrt{b_1}) = \delta_1 \sinh^{-1}(\sqrt{b_1}/\lambda_1),$$

$$X_S(b_2) = \begin{cases} \gamma_2 + \delta_2 \sinh^{-1}(b_2 - \xi/\lambda_2), & n \geq 25 \\ \gamma_2 + \delta_2 \ln \dfrac{b_2 - \xi}{\xi + \lambda_2 - b_2}, & n < 25 \end{cases}$$

where constants $\delta_1, \gamma_2, \delta_2, \lambda_1, \lambda_2$ are determined using the method presented, among others, by Pearson and Hartley [20]. We determine the statistic

$$Y_S^2 = X_S^2(\sqrt{b_1}) + X_S^2(b_2),$$

which under H_0 has approximately χ^2_2 distribution. Critical values generated by the Monte-Carlo method for $n = 20, 25, 50, 100, 150, 200, 300, 500, 1000$ were given by Bowman and Shenton [1]. The Y_S^2 test is also an omnibus test.

- Pearson-D'Agostino-Bowman [21]. Let

$\sqrt{b_1}(\alpha')$ be the lower and upper $100\alpha'$%-th percentile of $\sqrt{b_1}$ distribution, and let ${}_2b_2(\alpha')$ and ${}_1b_2(\alpha')$ be lower and upper critical points of b_2 distribution. Four points with coordinates $\{-\sqrt{b_1}(\alpha'), {}_2b_2(\alpha')\}, \{\sqrt{b_1}(\alpha'), {}_2b_2(\alpha')\}, \{-\sqrt{b_1}(\alpha'), {}_1b_2(\alpha')\}$, $\nu \{\sqrt{b_1}(\alpha'), {}_1b_2(\alpha')\}$ form a rectangle. When variables $\sqrt{b_1}$ and b_2 are independent then their values determined from the sample x_i are outside the rectangle with probability $\alpha' = 0.5 \{1 - (1-\alpha)^{1/2}\}$. The R test determines the fraction of points $(\sqrt{b_1}, b_2)$ which should be inside the rectangle.

5. Tests Based on Order Statistics

Under H_0 order statistics $\{y_i\}$ have expected values and variances-covariances denoted by known linear functions of parameters μ and σ. This allows us to apply the least squares method to the estimation of these parameters [17]. The best unbiased linear estimator $\hat{\sigma}$ of the parameter σ can be generally written in the form:

$$\hat{\sigma} = \sum_{i=1}^{h} d_{n-i+1} t_i,$$

where $h = [n/2]$, $t_i = y_{n-i+1} - y_i$, while $\{d_{n-i+1\,n}\}$ are constants satisfying certain conditions (e.g. their sum is equal to zero for each n). On the other hand the unbiased sample estimator of the parameter σ^2 is expressed by the variance from the sample s^2. The ratio of $\hat{\sigma}^2/s^2$ without a constant is close to 1. Tests built on the above ratio have left-hand-side critical regions, and the values of statistics are $\leqslant 1$. The following tests are included into the discussed group:

- G e a r y [11]

$$g = \sum_{i=1}^{n} |y_i - \bar{y}| / (nS^2)^{1/2},$$

critical values are given in tables prepared by P e a r s o n and H a r t l e y [19] (table 34A);

- D a v i d-H a r t l e y-P e a r s o n [6]

$$u = (y_n - y_1)/s,$$

critical values are in the above mentioned tables [19] (table 29C);

- S p i e g e l h a l t e r [26]

$$T = [1/(c_n u)^m + 1/g^m]^{1/m},$$

where $c_n = (1/2n)(n!)^{1/m}$ and $m = n-1$, and u and g are given above;

- S h a p i r o-W i l k [23]

$$W = [\sum_{i=1}^{h} a_{n-i+1} t_i]^2/S^2,$$

where

$$\sum_{i=1}^{n} a_{i,n} = 0, \quad \sum_{i=1}^{n} a_{i,n}^2 = 1;$$

- S h a p i r o-F r a n c i a [22]

$$W' = [\sum_{i=1}^{h} c_{n-i+1} t_i]^2/(\tilde{c}_n S^2),$$

where

$$c_{n-i+1} = m_{n-i+1}/\sqrt{\tilde{c}_n}, \quad \tilde{c}_n = \sum_{i=1}^{n} m_{i,n}^2;$$

- W e i s b e r g-B i n g h a m [28]

$$\tilde{W} = [\sum_{i=1}^{h} \tilde{m}_{n-i+1,n} t_i] 2/(d_n S^2),$$

where

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

$$\tilde{d}_n = \sum_{i=1}^{n} \tilde{m}_{i,n}^2;$$

- D'Agostino [4]

$$D_A = \left[\sum_{i=1}^{n} i y_i - \frac{n(n+1)}{2}\bar{y}\right] / (n^3 S^2)^{1/2}$$

or

$$Y = \sqrt{n}\,(D_A - 0.282095)/0.029986$$

- Fillibren [9]

$$r = \sum_{i=1}^{h} M_{n-i+1}\, t_i/(\tilde{e}_n\, S^2)^{1/2},$$

where

$$M_{i,n} = \Phi^{-1}(m_{i,n}), \quad \tilde{e}_n = \sum_{i=1}^{n} M_{i,n}^2,$$

$$m_{n,n} = (0.5)^{1/n}, \quad m_{1,n} = 1 - m_{n,n},$$

$$m_{i,n} = (i - 0.3175)/(n + 0.365),\ i = 2, 3, \ldots, n-1.$$

Critical values for the above mentioned tests are given by their authors, and for the W test also by Domański [7] (table 11).

6. General Discussion of Tests for Normality

We reviewed various tests for normality. They are divided into three basic groups according to their structure. Many of these tests have the properties of the omnibus test. The latter can be adviseable when some a priori information on the departure from normality is given. From this group we discussed the tests K^2, Y^2, T, W, W', W, D_A, and r. All of them have been given in the last 15 years. So far the omnibus test has not been constructed in the group of tests based on the comparison of normal and empirical distribution functions. In this group of tests the Cramer-von Mises and Anderson-Darling tests with modifications should be mainly used.

Many normality tests reveal similar properties. Statistics on which the tests $\sqrt{b_1}$, b_2, u, g, W, r are based, are invariance due to the shift of scale and location. Therefore, they are suitable for testing complex H_0 hypothesis. The tests D, KS, W^2, U^2, V, A^2 have completely determined distributions under H_0 and are suitable for testing simple hypotheses of normality.

For the majority of tests the density function of their distribution statistics have not been found yet. The critical values required for them have been generated by means of the Monte-Carlo method. The tests W, W', $\tilde{W}$, and r have left-hand-sided critical regions. Some tests statistics need some constants for each n (e.g. tests W, W', $\tilde{W}$, r).

The problem of power of the normality tests is relatively well known and has been discussed among others by Shapiro et al. [24], Stephens [25], Giorgi and Cinci [12], Green and Hegazy [13] and Pearson et. al [21]. However, there are few general results which are complete and applicable, as it is the case in the theory of parametric tests. Mathematical difficulties connected with determining the power of tests are usually very big. It is also difficult to determine practically the G class distributions.

It is known that for the normal distribution $\sqrt{\beta_1} = 0$ and $\beta_2 = 3$. Similarly, other distributions can be characterized by giving a pair of values $(\sqrt{\beta_1}, \beta_2)$. Hence, from the G class the distributions from the below groups are chosen according to the values of $\sqrt{\beta_1}$ and β_2. Below we shall present particular groups giving for each of them the type of distribution and some distributions belonging to these groups:

Group 1: $|\sqrt{\beta_1}| > 0.3$, $\beta_2 > 3.0$; asymmetric distributions with long tails - χ^2, log-normal, non-central χ^2, exponential, Weibull's, Pareto's;

Group 2: $|\sqrt{\beta_1}|>0.3$, $\beta_2 < 3.0$; asymmetric distributions with short tails - beta, S_B Johnson's;

Group 3: $|\sqrt{\beta_1}| \leq 0.3$, $\beta_2 > 4.5$; symmetric distributions with long tails - double χ^2, uniform, Cauchy's, Laplace's, Tukey's, S_U Johnson's, logistic;

Group 4: $|\sqrt{\beta_1}| \leq 0.3$, $\beta_2 < 2.5$; symmetric distributions with short tails - beta, double χ^2, S_B Johnson's, Tukey's;

Group 5: $|\sqrt{\beta_1}| \leq 0.3$, $2.5 \leq \beta_2 \leq 4.5$; almost normal distributions - t-Student's with 10 degrees of freedom, S_B Johnson's with parameters $\gamma = 0$ and $\delta = 3$, logistic, Weibull's at $k = 2$.

We include the distributions into particular groups using the known values of $\sqrt{\beta_1}$ and β_2. They differ significantly within the same distribution determined at various values of parameters determining it. For instance, for χ^2_4 we have $\sqrt{\beta_1} = 1.41$ and $\beta_2 = 6.00$, while for χ^2_{10} - $\sqrt{\beta_1} = 0.89$ and $\beta_2 = 4.20$. Due to this the same distribution at the values of parameters determined in different ways, is included into various groups.

Table 1 presents power of some tests for normality, expressed in per cent, for $\alpha = 0{,}05$ and $n = 20$, taking into account alternative distributions: χ^2_m, $m = 1, 2, 4, 10$ degrees of freedom, log-normal $LN(\mu, \sigma)$ with parameters $\mu = 0$ and $\sigma = 1$, t-Student's t_2 with two degrees of freedom, Cauchy's t_1, beta $B(p,q)$ with parameters $p = 2$ and $q = 1$, uniform $B(1,1)$, and Laplace's.

The procedure for testing normality based on the W test has a greater power for almost all G class distributions than other tests. Especially, the W test is sensitive to asymmetry with a long tail. For instance, for the population with the distribution χ^2_{10}, χ^2_4 and $LN(0,1)$ the values of power are 29, 50 and 93%, respectively. In the group of tests for normality based on order statistics the W test should be assumed the best one as far as power is concerned, when $n \leq 50$. Similar power properties have the W' and r tests which can be proposed for large samples of sizes $50 \leq n \leq 100$.

The tests based on sample moments reveal big power at a defined type of G class distributions: $\sqrt{b_1}$ for asymmetric distributions with long tails (χ^2_1, χ^2_2, χ^2_4, $LN(0,1)$) and test b_2

Table 1

Empirical power of tests for normality in per cent, for $\alpha = 0.05$ and $n = 20$

G class distributions	$\sqrt{\beta_1}$	β_2	χ^2	D	W^2	V	U^2	A^2	g	b_1	b_2	K^2	R	W	W'	r	Y	Y*
χ^2_1	2.83	15.0	94	86	94	94	93	-	-	89	53	82	82	98	94	94	80	-
χ^2_2	2.00	9.0	33	59	74	71	70	82	-	74	34	60	61	84	82	82	52	-
χ^2_4	1.41	6.0	13	33	45	23	21	15	19	49	27	40	39	50	24	-	24	-
χ^2_{10}	0.89	4.2	7	18	23	24	14	-	14	29	19	27	25	29	-	-	16	-
LN(0,1)	6.19	113.9	95	78	88	84	85	91	49	89	58	82	81	93	94	94	77	-
B(1,1)	0	1.8	11	12	16	17	18	17	-	0	29	16	17	23	4	4	8	14
B(2,1)	-0.57	2.4	8	-	-	23	16	12	19	8	13	12	11	35	-	-	6	-
Laplace's	0	6.0	17	22	26	22	25	26	-	25	27	25	26	31	33	33	28	34
t_1	0	-	41	86	88	87	88	98	-	89	81	79	80	89	91	92	92	94
t_2	0	-	-	55	-	-	-	-	-	52	53	54	52	54	59	60	56	66

* Directional test.

\- Not tested.

for symmetric distributions with long tails (t_1, $B(1, t_2)$). An interesting power value ranging from the power of tests $\sqrt{b_1}$ to b_2 is in the case of K^2 and R tests. Hence, a conclusion can be drawn that from the above mentioned tests the tests b_1 and b_2 should be used at determined distributions and tests K^2 and R in the case when it is impossible to determine distributions from the G class.

The tests for normality based on the measure of consistency of empirical and normal distributions reveal similar powers although the best of them is the A^2 test and the worst the D test. Tests W^2, V and U^2 have similar powers as the A^2 test.

In the light of the above mentioned tests for normality the χ^2 test is much weaker at G class distributions given in the Table 1. Therefore, it should not be used in practice as a test for normality. Instead of the χ^2 test in the case of large samples $n > 50$ without constructing the disjoint series, D'Agostino's Y test or W' test should be used.

As was shown by D y e r [8] the tests for normality with unknown parameters μ and σ^2 have greater power than the tests with unknown σ^2 only. Besides, the power increases with the increase of n at both unknown parameters μ and σ^2.

Insignificant differences in the critical values of tests for normality based on the comparison of normal and empirical distribution functions occur when σ^2 is estimated by s^2 or $\hat{\sigma}^2$. The critical values are much lower when both parameters μ and σ^2 are to be estimated. The power of the normality test increases significantly, as shown by P e a r s o n et al. [21], when instead of the omnibus test a directional test is used. The directional tests for normality are used for determined G class distributions. For instance, D'Agostino Y test can be treated as an omnibus test for various G class distributions, however for some of them, when $\beta_2 < 3$ a left-hand-sided Y test can be used, while for others, when $\beta_2 > 3$ a right-hand-sided Y test is employed.

It should be noted that the power of test increases with the increase of the sample size n. For instance, for the W test at $\alpha = 0.05$ and distributions χ^2_4, χ^2_{10}, LN(0,1) the power is as follows:

n	χ^2_4	χ^2_{10}	LN(0,1)
10	24	11	60
20	50	29	93
30	71	35	99
40	87	48	100
50	95	56	100

That is why when testing the normality of given variable at least one sample of size $n \geq 30$ should be used. On the other hand, the power of test decreases with the decrease of significance level α. For instance, for tests Y and W at the distribution LN(0,1) we have the following powers:

α \ n	0.10		0.05		0.02		0.01	
	Y	W	Y	W	Y	W	Y	W
10	51	68	42	58	34	45	28	38
20	80	95	75	92	66	86	61	81
30	93	99	90	99	86	97	82	96
40	97	100	94	100	92	99	90	99
50	99	100	98	100	97	100	96	100

Hence, of great importance is an adequate choice of significance level α to verify the hypothesis H_0.

It can be suggested that in the above groups the best tests for normality from the point of view of their power at adequate sample sizes are the following tests:

Group 1 - W test, for $n \leq 50$,
W' test for $50 \leq n \leq 100$,
an arbitrary W^2, V, V^2 test, for $n > 100$;

Group 2 - W test for $n \leq 50$,
tests K^2, R, for $n > 50$;

Group 3 - r test, for $n \leq 50$,
one of the W' or r tests, for $50 < n \leq 100$,
Y test, for $n > 100$;

Group 4 - b_2 test, for $n < 20$,
K^2 test, for $20 \leq n \leq 200$;

Group 5 - W test, for $n \leq 20$,
b_2 test, for $20 < n < 50$,
one of the W' or r tests, for $50 \leq n \leq 100$,
one of the K^2 or Y tests, for $n > 100$.

Generally, the procedure of verification of H_0 with normal distribution of random variable X on the basis of a simple sample $\{x_i\}$ taken from a population according to a corresponding scheme of sampling, should be as follows. We determine the values of $\sqrt{b_1}$ and b_2 from sample $\{x_i\}$. They are the estimates $\sqrt{\beta_1}$ and β_2. Then we choose one of the above mentioned groups and a respective test for normality according to the sample size n. If the type of departure from normality (e.g. skewness) is known, we choose the test of the greatest power which would correspond to the determined hypothesis H_1 (e.g. the $\sqrt{b_1}$ test for asymmetric skew distributions) and if there is no such an a priori information one of the omnibus tests is used.

BIBLIOGRAPHY

[1] Bowman K. O., Shenton L. R. (1975): Omnibus Contours for Departures from Normality Based on $\sqrt{b_1}$ and b_2, "Biometrika" 62, p. 243-250.

[2] Cramer H. (1946), Mathematical Methods of Statistics, Princeton University Press (wyd. pol. Metody matematyczne w statystyce, PWN, Warszawa 1958).

[3] D'Agostino R. B. (1970), Transformations to Normality of the Null Distribution g_1, "Biometrika" 57, p. 679-681.

[4] D'Agostino R. B. (1971), An Omnibus Test Normality for Moderate and Large Size Samples, "Biometrika" 58, p. 341-348.

[5] D'Agostino R. B., Pearson E. S (1973), Tests for Departure from Normality, Empirical Results for the Distributions of b_2 and $\sqrt{b_1}$, "Biometrika" 60, p. 613-622.

[6] David H. A., Hartley H. O., Pearson E. S. (1954), The Distribution of the Ratio, in a Single Sample, of Range to Standard Deviation, "Biometrika" 41, p. 482-493.

[7] Domański C. (1979), Statystyczne testy nieparametryczne, PWE Warszawa.

[8] Dyer A. R. (1974), Comparisons of Test for Normality with a Cautionary Note, "Biometrika" 61, p. 185-189.

[9] Fillibren J. L. (1975), The Probability Plot Correlation Coefficient Tests for Normality, "Technometrics" 17, p. 111-117.

[10] Fisz M. (1967), Rachunek prawdopodobieństwa i statystyka matematyczna, PWN, Warszawa.

[11] Geary R. C. (1947), Testing for Normality, "Biometrika" 34, p. 209-242.

[12] Giorgi G. M., Cinci S. (1975), Sulle funzioni discriminatorie di alcuni testes univariate di normalita, Stud. di Econ. 6, p. 100-120.

[13] Green J. R., Hegazy Y. A. S. (1976), Powerful Modified-EDF Goodness of Fit Tests, JASA 71, p. 204-209.

[14] Hegazy Y. A. S., Green J. R. (1975), Some New Goodness of Fit Tests using Order Statistics, Appl. Statist., 24, p. 299-308.

[15] Johnson J., Kotz S. (1970), Continuous Univariate Distribution, Vol. 2, Boston, Houghton-Mifflin.

[16] Lilliefors H. W. (1967), On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown, JASA 62, p. 399-402.

[17] Lloyd E. H. (1952), Least Squares Estimation of Location and Scale Parameters using Order Statistics, "Biometrika" 39, p. 88-95.

[18] Muldholland H. P. (1977), On the Distribution of $\sqrt{b_1}$ for Samples of Size at Most 25, with Tables, "Biometrika" 64, p. 401-409.

[19] Pearson E. S., Hartley H. O. (1966), Biometrika Tables for Statisticians, Vol. 1, Cambridge University Press.

[20] Pearson E. S. (1972), Biometrika Tables for Statisticians, Vol. 2, Cambridge University Press.

[21] Pearson E. S., D'Agostino R. B., Bowman K. O. (1977), Tests for Departure from Normality: Comparison of Powers, "Biometrika" 64, p. 231-246.

[22] Shapiro S. S., Francia R. S. (1972), Approximate Analysis of Variance Test for Normality, JASA 67, p. 215-216.

[23] Shapiro S. S., Wilk M. B. (1965), An Analysis of Variance Test for Normality (Complete Samples), "Biometrika" 52, p. 591-611.

[24] Shapiro S. S., Wilk M. B., Chen H. J. (1968), A Comparative Study of Variance Test for Normality, JASA 63, p. 1343-1372.

[25] Stephens M. A. (1974), EDF Statistics for Goodness of Fit and Some Comparisons, JASA 69, p. 730-737.

[26] Spiegelhalter D. J. (1977), A Test for Normality Against Symmetric Alternatives, "Biometrika" 64, p. 415-418.

[27] Van Soest J. (1967), Some Experimental Results Concerning Tests of Normality, Statist. Neerland. 21, p. 91--97.

[28] Weisberg S., Bingham C. (1975), An Approximate Analysis of Variance Test for Non-Normality Suitable for Machine Calculation, "Technometrics" 17, p. 133--134.

[29] Zieliński R. (1972), Tablice statystyczne, PWN, Warszawa.

Wiesław Wagner

TESTY JEDNOWYMIAROWEJ NORMALNOŚCI

W artykule prezentowane są testy jednowymiarowej normalności w podziale na testy oparte na porównaniu dystrybuant rozkładu empirycznego i normalnego, testy wykorzystujące momenty z próby oraz testy oparte na statystykach pozycyjnych. Omówione zostały podstawowe własności poszczególnych testów, w szczególności z punktu widzenia odstępstwa od normalności. Przedyskutowana została również moc przedstawionych testów i ich przydatność do zastosowań z punktu widzenia liczebności próby i odstępstwa od normalności.