ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 123, 1992

Czesław Domański, Wiesław Wagner***

ON MULTIVARIATE GOODNESS-OF-FIT TESTS

1. INTRODUCTION

Widely conducted research proves that for univariate distributions Shapiro-Wilk test is an optimal one as the test power is concerned.

In this paper properties of the generalized Shapiro-Wilk test are presented. The test power is tested against generalized goodness-of-fit Kolmogorov-Smirnov, Chi-square and Hellwig tests. A comparative analysing using Monte Carlo methods for n = 10, 20, 30, 50, p = 2, 3, 5, α = 0,05 is also carried out.

2. SOME MULTIVARIATE NORMALITY TESTS

In comparative analysis of power of multivariate normality tests four multivariate tests i.e. χ^2-Pearson, Kolmogorov-Smirnov, Shapiro-Wilk and Hellwig tests were taken into account. Normal, uniform, exponential and Pareto distributions were used as alternative distributions.

Let $\underline{X}_1, \ldots, \underline{X}_n$ be a sequence of n independent p-variate vectors of observations, and let it be a sample of p-variate $\underline{U} = (\underline{X}_1, \ldots, X_n)$ observations.

Let $\underline{X}$ and $\underline{S}$ denote vector of means and variance-covariance

* Associate Professor at the University of Łódź.
** Lecturer at the Academy of Agriculture, Poznań.

matrix for U sample, respectively. It is assumed that $\underline{S} = 0$ therefore $\underline{S}^{-1}$ and $\underline{S}^{-1/2}$ exist. Mahalonobis distance between $\underline{X}_j$ and $\underline{X}$) vectors is expressed by $Y_j = (\underline{X}_j - \underline{X}) \underline{S}_j^{-1} (\underline{X}_j - \underline{X})$, $j = 1, \ldots, n$.

Lastly, HCM denotes the null hypothesis that a sequence of p-variate independent $\underline{X}_1, \ldots, \underline{X}_n$ vectors belongs to a population having with $N_p(\underline{\mu}, \underline{\Sigma})$ distribution when μ and $\underline{\Sigma}$ parameters are unknown (D o m a ń s k i, G a d e c k i, W a g n e r, 1984).

2.1. Multivariate χ^2 - Pearson test

When constructing test statistic of the multivariate normality χ^2 - Pearson test we use generalized sample class intervals which are defined as follows: $E_k(\underline{\bar{X}}, \underline{S}) = \underline{X}_j \in R^p \; c_{k-1} \le Y_j < c_k$ for $k = 1, \ldots, q$ with division $0 = c_0 < c_1 < \ldots < c_q < c_{q+1} = \infty$ pre-assigned.

The construction of $\chi_p^2(q)$ test statistic follows 6 stages:

a) q number is defined;

b) a sequence of numbers $c_0, c_1, \ldots, c_q$ defined by a sequence of inequalities $0 = c_0 < c_1 < \ldots < c_q < c_{q+1} = \infty$;

c) generalized, class intervals and their empirical sizes

$n_k = E_k(\underline{\bar{X}}, \underline{S}) = |\{\underline{\bar{X}}_j \in R^p: c_{k-1} \le Y_j < c_k\}|$,

$k = 1, \ldots, q$, with $n = n_1 + \ldots + n_q$, where $|B|$ denotes power of B set, are determined;

d) for a generalized class interval $E_k(\underline{\bar{X}}, \underline{S})$ probabilities $\hat{p}_k = P(c_{k-1} \le Y_j < c_k)) = F_{\chi_p^2}(c_k) - F_{\chi_p^2}(c_{k-1})$ are determined, where $F_{\chi_p^2}(.)$ denotes distribution function of χ^2 with p degrees of freedom;

e) expected (theoretical) sizes $\hat{p}_k = n\hat{p}_k$ for generalized sample class intervals $E_k(\underline{\bar{X}}, \underline{S})$, where $\hat{n}_1 + \ldots + \hat{n}_q = n$ are accepted;

f) test statistic

$$\chi_p^2(q) = \sum_{k=1}^{q} \left\{ \frac{n_k - \hat{n}_k}{\sqrt{\hat{n}_k}} \right\}^2 = \frac{1}{n} \sum_{k=1}^{q} \frac{n_k^2}{\hat{p}_k} - n,$$

is calculated. Asymptotic distribution of $\chi_p^2(q)$ is $\chi_{q-2}^{\ \ 2} + \lambda\chi_1^2$, $0 < \lambda < 1$ (M o o r e and S t u b b e b i n s, 1981). In practice, our analysis can be limited to quantiles of χ_{q-2}^2 and χ_{q-1}^2. Hence, HCM hypothesis is rejected on the level of significance α when $\chi_p^2(q) \notin (\chi_{\alpha,q-2}^2, \chi_{\alpha,q-1}^2)$.

The research conducted on the power of this test was limited to the case when $q = 5, 6, 8$ and to the following intervals:

$q = 5$, $0 = c_o < c_1 < c_2 < c_3 < c_4 < c_5 < c_6 = \infty$, where $c_1 = 2$
$c_2 = 5$, $c_3 = 8$, $c_4 = 12$, $c_5 = 17$;

$q = 6$, $0 = c_o < c_1 < c_2 < c_3 < c_4 < c_5 < c_6 < c_7 = \infty$, where
$c_1 = 2$, $c_2 = 5$, $c_3 = 8$, $c_4 = 11$, $c_5 = 15$, $c_6 = 19$;

$q = 8$, $0 = c_o < c_1 < c_2 < c_3 < c_4 < c_5 < c_6 < c_7 < c_8 < c_9 = \infty$,
where $c_1 = 2$, $c_2 = 5$, $c_3 = 8$, $c_4 = 11$, $c_5 = 14$,
$c_6 = 17$, $c_7 = 20$, $c_8 = 22$.

2.2. Multivariate Kolgomorov-Smirnov test

Construction of KS_p test statistics of multivariate normality Kolmogorov-Smirnov test includes the following 3 stages:

a) Y_j is ordered into non-decreasing sequence $Y_{(1)} \le Y_{(2)} \le$
$\le \ldots \le Y_{(n)}$;

b) values of χ_p^2 $F(Y_{(j)})$ for $Y_{(j)}$, $1 = 1, \ldots, n$; quantiles are calculated;

c) values of the test statistic

$$KS_p = \max_{1 \le j \le n} | F_{\chi_p^2} (Y_{(y)} - j/n)|$$

are determined.

HCM hypothesis is rejected when $KS_p > KS_p(\alpha, n)$, and where $KS_p(\alpha, n)_i$ is a criterial value of KS_p distribution for α, p and n pre-assigned.

2.3. Multivariate Shapiro-Wilk test

Construction of W_p test statistic of multivariate Shapiro-Wilk test follows 6 stages:

a) $Z_j = Y_{j/n}$, $j = 1, 2, \ldots, n$ is calculated;

b) such Z_m are chosen for which $Z_m = \max\{Z_1, \ldots, Z_n\}$;

c) bilinear forms are calculated

$\tilde{Z}_j = (\underline{X}_m - \underline{\bar{X}})S^{-1}(\underline{X}_j - \bar{X})/n, \quad j = 1, \ldots, n;$

d) $\{\tilde{Z}_j\}$ is ordered into a non-decreasing sequence

$\tilde{Z}_{(1)} \leq \tilde{Z}_{(2)} \leq \ldots \leq \tilde{Z}_{(n)}$;

e) a linear combination

$$Q = \sum_{l=1}^{h} a_{n-l+1,n}(\tilde{Z}_{(n-l+1)} - \tilde{Z}_{(l)}),$$

where $h = \begin{cases} n/2, & \text{when } n \text{ - even,} \\ (n-1)/2, & \text{when } n \text{ - uneven,} \end{cases}$ is calculated;

f) test statistic

$W_p = Q^2/\tilde{Z}_m$,

is determined.

HCM hypothesis is rejected when $W_p < W_p(\alpha, p, n)$, where $W_p(\alpha, p, n)$ is a critical value of W_p distribution with α, p and n pre-assigned. When determining Q at point (e) constant values $a_{n-l+1,n}$ $l = 1, \ldots, h$ are used. For small n ($n \leq 50$) they were given by S h a p i r o and W i l k (1965), while for large they are determined by R o y s t o n (1982) approximate formuale. PGENNW program of calculating constant values $a_{n-l+1,n}$, $n \in \langle 3, 2000)$ was given by D o m a ń s k i, G a d e c k i and W a g n e r (1986).

2.4. Multivariate Hellwig test

In case of Hellwig test the construction of the test statistic is done is the following way:

a) for the pre-assigned sample matrix of covariance $\underline{S}$ such $S^{1/2}$ is determined that $\underline{S}^{1/2}(\underline{S}^{1/2}) = \underline{S}$, and next $S^{-1/2}$ is settled;

b) vectors of scaled residuals are determined

$\underline{Z}_j = \underline{S}^{-1/2}(\underline{X}_j - \overline{\underline{X}})$, $j = 1, 2, \ldots, n$;

c) from matrix $\underline{Z} = (\underline{Z}_1, \ldots, \underline{Z}_n)$, treated as a system of n points of R^p, where $\underline{Z}_j = (z_{1j}, \ldots, z_{pj})$, Euclidean distance is calculated

$$d(\underline{Z}_j, \underline{Z}_{j'}) = \left[\sum_{i=1}^{p} (z_{ij} - z_{ij'})^2\right]^{1/2}$$

when $j, j' = 1, \ldots, n$; $j < j'$; since $d(\underline{Z}_j, \underline{Z}_{j'}) = d(\underline{Z}_{j'}, \underline{Z}_j)$;

d) for subsequent vectors $\underline{Z}_j$, $j = 1, \ldots, n$ their smallest distances from the remaining vectors are defined i.e.

$$c_{j'} = \min_{\substack{1 \le j \le n \\ j \ne j'}} d(\underline{Z}_j, \underline{Z}_{j'});$$

e) distances $c_1, \ldots, c_n$ calculated at point (d) are ordered into a non-decreasing sequence $c_{(1)} \le \ldots \le c_{(n)}$;

f) empirical distribution function is defined

$F_e(c) = \max\{j/n\}$, $j = 1, \ldots, n$;

g) new $(p \times n)$ - variate matrix $\underline{U}^*$ is created and independent realizations of $N_p(\underline{O}, \underline{I})$ distribution are its elements; thus np random normal numbers N(0, 1) are generated (cf. e.g. Zieliński, 1972);

h) actions mentioned in points (a)-(f) are repeated on the elements of matrix $\underline{U}^*$ which gives a new sequence $c^*_{(1)} \le c^*_{(2)} \le \ldots \le c^*_{(n)}$ and distribution function:

$F_g(c) = \max\{j/n\}$, $c_j < c$; $j = 1, 2, \ldots, n$;

i) finally Kolmogorov-Smirnov test statistic for two samples is defined

$$H = \max_c |F_e(c) - F_g(c)|;$$

if $H > H_{\alpha,p,n}$ HCM hypothesis is rejected.

2.5. Defining distribution function of χ^2_p

Kolmogorov-Smirnov test and χ^2 - Pearson test demand defining distribution function of the central χ^2 distribution with p degrees of freedom. We give the appropriate formulae for defining $F_{\chi^2_p}(x) \equiv Q(x|p)$ distribution function after Abramowitz and Stegun, (1964):

a) p - uneven

$$Q(x|p) = 2Q(z) + 2\varphi(z) \sum_{r=1}^{(p-1)/2} \frac{z^{2r-1}}{1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2r-1)},$$

b) p - even

$$Q(x|p) = \sqrt{2\pi}\,\varphi(z) \left\{1 + \sum_{r=1}^{(p-2)/2} \frac{z^{2r}}{2 \cdot 4 \cdot \ldots \cdot (2r)}\right\}$$

where: $z = \sqrt{x}$ and

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

$$Q(z) = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\,dt = 0.5(1 + c_1 z + c_2 z^2 + c_3 z^3 + c_4 z^4)^{-4}$$

with constant values

$c_1 = 0.196854,$ $c_2 = 0.115194$

$c_3 = 0.000344$ $c_4 = 0.019527.$

3. CHOSEN ALTERNATIVE DISTRIBUTION

We tested power of the given above multivariate normality tests using 4 alternative distributions: normal, uniform, exponential and Pareto.

Let $\alpha \in (0, 1)$ bea random number with uniform distribution within (0, 1) interval. Random numbers of N(0, 1) normal distributions were determined according to the rule of "dozen" i.e. y is a random normal number from standardised normal distribution when

$$y = \alpha_1 + \ldots + \alpha_6 - 6,$$

where $\alpha_i \in (0, 1)$, $i = 1, \ldots, 6$.

Random numbers for the remaining distributions were defined according to the following formulae:

a) exponential distribution $y = -\ln\alpha$

b) Pareto distribution

$y = 1/\alpha^{1/k}$ for parameter $k = 2,4$.

Matrices U for alternative distributions were created for a system n of p-variate vectors of independent components.

4. ANALYSIS OF THE TESTS' POWER

The results contained in Table 1-3 make a good starting point for the analysis of power of multivariate normality tests. Due to high costs of Monte-Carlo experiment[1] we did not take into account a large clase of alternative distributions. On the other hand, results obtained by other researchers are difficult to compare. Nevertheless, on the basis of the results presented below, the following conclusions can be drawn:

1) generally, the power of some tests increase to 1 even for small n's;

2) the power of some tests increases to 1 even for small n's;

3) Shapiro-Wilk test is characterized by the highest power;

4) tests based on Chi-square statistic (differing in the number of degrees offreedom) are characterized by the similar level of power;

5) power of tests for large n's increase to 1, mainly because the boundary theorems (for large samples) are used for testing distributions.

---

[1] The analysis took into account for each n, i alternative distributions q = 1000 repetitions (repetitions of sample from a defined generator) while for normal distribution q = 2500 in order to define quantiles for the considered tests.

Table 1

Empirical power of multivariate normality tests
for alternative uniform distribution

Test	p	Significance level							
		α = 0.10				α = 0.05			
		n = 50	n = 30	n = 20	n = 10	n = 50	n = 30	n = 20	n = 10
χ_5^2	2	988	827	587	480	809	747	576	432
	3	916	850	657	535	822	774	600	380
χ_6^2	2	978	824	573	309	889	782	499	726
	3	980	849	657	274	892	751	502	277
χ_8^2	2	958	827	639	410	889	476	599	327
	3	959	808	625	449	892	768	613	410
Kolmogorov--Smirnov	2	936	814	611	279	825	730	568	287
	3	964	808	610	282	867	749	563	282
Shapiro--Wilk	2	1 000	1 000	868	778	1 000	932	729	422
	3	1 000	1 000	921	822	1 000	942	741	512
Hellwig	2	581	477	425	326	520	376	354	279
	3	591	497	425	335	531	565	374	274

Source: The author's calculations.

Table 2

Empirical power of multivariate normality tests for alternative exponential distribution

Test	p	Significance level					
		α = 0.10			α = 0.05		
		n = 30	n = 20	n = 10	n = 30	n = 20	n = 10
χ^2_5	2	755	521	382	769	568	361
	3	848	656	395	937	554	384
	5	922	788	599	971	740	584
χ^2_8	2	765	570	325	769	504	278
	3	931	658	356	837	551	304
	5	987	783	580	970	751	537
Kolmogorov--Smirnov	2	717	698	593	829	611	409
	3	800	734	683	878	681	565
	5	971	813	736	994	712	620
Shapiro--Wilk	2	969	890	889	929	881	817
	3	957	895	957	923	892	886
	5	1 000	1 000	1 000	1 000	1 000	905
Hellwig	2	636	597	350	537	446	284
	3	726	605	406	544	505	436
	5	781	705	545	688	519	474

Source: The author's calculations.

T a b l e 3

Empirical power of multivariate normality tests for alternative, Pareto distribution

Test	p	Significance level					
		$\alpha = 0.10$			$\alpha = 0.05$		
		n = 30	n = 20	n = 10	n = 30	n = 20	n = 10
χ_5^2	2	456	350	303	465	295	299
	3	592	396	373	536	243	245
	5	672	417	268	689	261	243
χ_8^2	2	468	249	280	365	290	256
	3	588	264	216	536	283	225
	5	671	492	308	681	222	217
Kolmogorov--Smirnov	2	452	398	214	546	336	227
	3	708	462	307	605	406	311
	5	915	670	406	954	505	408
Shapiro--Wilk	2	1 000	993	801	944	986	744
	3	1 000	993	799	1 000	985	799
	5	1 000	984	727	1 000	967	628
Hellwig	2	630	287	256	510	256	221
	3	637	381	252	547	348	229
	5	727	475	310	676	306	298

S o u r c e: The author's calculations.

REFERENCES

A b r a m o w i t z M., S t e g u n I. A. (1964), *Handbook of Mathematical Functions*, National Bureau of Standards - Applied Mathematics.

D o m a ń s k i C., G a d e c k i H., W a g n e r W. (1986), *Budowa tablic współczynników i wartości krytycznych testu W Shapiro-Wilka dla n > 50*, Research work R.III.9.4.1.

Moore D. S., Stubbebins J. B. (1981), *Chi-square Tests for Multivariate Normality with Applications to Common Stock Processes*, "Comunications in Statistics. Theory and Methods", A. 10, p. 713-739.

Royston J. P. (1982), *An Extension of Shapiro and Wilk's W Test for Normality to Large Samples*, "Journal of the Royal Statistical Society", Ser. C: Applied Statistics, No. 31, p. 115-124.

Shapiro S. S., Wilk M. B. (1965), *An Analysis of Variance Test for Normality (Complete Samples)*, "Biometrica", No. 65, p. 591-611.

Zieliński R. (1972), *Generatory liczb losowych. Programowanie i testowanie na maszynach cyfrowych*, WN-T, Warszawa.

Czesław Domański, Wiesław Wagner

O TESTACH WIELOWYMIAROWEJ ZGODNOŚCI

W teorii wnioskowania statystycznego obszerną klasę stanowią testy zgodności, pozwalające sprowadzić hipotezę o zgodności rozkładu hipotetycznego z rozkładem badanej zmiennej losowej.

W artykule prezentujemy testy wielowymiarowej zgodności, w szczególności testy normalności. Wyróżniamy tutaj uogólniony test W Shapiro-Wilka, test chi-kwadrat, test Kołmogorowa-Smirnowa oraz Cramera von Misesa. Podane kwantyle i uwagi dotyczące mocy tych testów dają podstawę do szerszego ich wykorzystywania w praktyce statystycznej.