

ANALÝZA SÍLY VYBRANÝCH KLASICKÝCH A ROBUSTNÍCH TESTŮ NORMALITY PROTI BIMODÁLNÍMU ROZDĚLENÍ

L. Střelec

Došlo: 26. června 2009

Abstract

STŘELEČ, L.: *Analysis of power of the classical and robust normality tests against bimodal distribution*. Acta univ. agric. et silvic. Mendel. Brun., 2009, LVII, No. 6, pp. 253–260

The aim of this paper is to compare the power of selected normality tests to detect a bimodal distribution. We use some classical normality tests (the Shapiro-Wilk test, the Lilliefors test, the Anderson-Darling test, the classical Jarque-Bera test and the Jarque-Bera-Urzua test), some robust normality tests (the robust Jarque-Bera test and the Medcouple test) and the modified Jarque-Bera tests, where the median instead of the mean is used in the classical Jarque-Bera test statistic. The results of simulation study show that the Anderson-Darling and the Shapiro-Wilk tests outperform the others, especially in small sample sizes. On the other hand the classical Jarque-Bera, the Jarque-Bera-Urzua and robust Jarque-Bera tests are biased, especially in small sample sizes again. Finally, the modification of the Jarque-Bera test leads to increase of power against bimodal distribution.

tests of normality, robust tests, Monte Carlo simulation, power comparison, bimodal distribution

V současné době existuje velké množství odborné literatury, která se věnuje testům normality a jejich statistickým vlastnostem (např. Thode, 2002). Mimo klasických formálních testovacích procedur určených k testování normality existují rovněž metody založené na grafickém posouzení či testy testující odlehle hodnoty či jiné odchylky od normality. Přesto však nejvíce používanými metodami jsou formální testovací procedury, které doplňují prvotní posouzení normality vycházející z grafických metod. Všeobecně nejběžněji využívaným omnibus testem je Shapiro-Wilkův test (Shapiro a Wilk, 1965), v ekonomii či financích je naopak nejvíce využíván Jarque-Bera test (Jarque a Bera, 1980). V poslední době se mimo klasických testů objevují také nové testy normality, které jsou založeny na robustních charakteristikách a které dokážou částečně eliminovat problémy s odlehlými hodnotami. Jako příklad lze uvést robustní Jarque-Bera test (viz Gelová a Gastwirth, 2009), který vykazuje vyšší sílu testu proti těžkochvostým alternativám než klasický Jarque-Bera test, dále pak testy normality založené na medcouple charakteristice jako robustní charakteristice šikmosti (viz Brys et al., 2004 a 2008) či třídu robustních RT testů, kde speciálními případy jsou

jak klasický Jarque-Bera test, tak i robustní Jarque-Bera test (viz Střelec a Stehlík, 2009a).

Cílem práce je komparace a analýza síly vybraných testů normality (zahrnující jak klasické, tak i robustní testy normality) proti bimodálním alternativám, především se zaměřením na vylepšení síly klasického Jarque-Bera testu, který proti těmto alternativám a pro malé rozsahy souborů vykazuje nulovou sílu. Práce je rozdělena do čtyř hlavních částí. V první jsou charakterizovány modifikace Jarque-Bera testu normality a směsi normálních rozdělení. Ve druhé části jsou prezentovány výsledky provedené komparační analýzy síly testů. Ve třetí části jsou diskutovány dosažené výsledky především s ohledem na porovnání síly klasických, robustních a vybraných testů pocházejících z RT třídy testů normality. Čtvrtá část obsahuje souhrn nejdůležitějších výsledků.

MATERIÁL A METODY

Pro potřeby komparační analýzy bude sledována síla vybraných klasických testů normality, konkrétně Anderson-Darlingova testu (AD), Shapiro-Wilkova testu (SW), Lilliefors (Kolmogorov-Smirnovova)

testu (LT), Jarque-Bera testu (JB), vybraných robustních testů normality, konkrétně robustního Jarque-Bera testu (RJB) a medcouple testu (MC-LR), a modifikovaných Jarque-Bera testů, jako speciálních případů třídy RT testů definované v [Střelec a Stehlík, 2009a, b]. Modifikované Jarque-Bera testy jsou testy využívající mediánu jako odhadu střední hodnoty místo aritmetického průměru v klasické testové statistice Jarque-Bera testu. Necht $\mu_k = E(X - E(X))^k$ je k -tý teoretický moment rozdělení pro všechny pozitivní hodnoty k , necht $\hat{\mu}_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k$ je jeho empirická varianta a necht $\hat{m}_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{M}_n)^k$ je jeho robustní varianta, kde X_i pro $i = 1, \dots, n$ je výběr z iid náhodné proměnné, $\bar{X}$ je výběrový odhad prvního obecného momentu (aritmetický průměr) a $\bar{M}_n$ je výběrový medián. Obecně pak můžeme zavést odhad empirického momentu $M_{i,k} = \frac{1}{n} \sum_{m=1}^n (X_m - M_{(i)})^k$, kde $i \in \{0, 1\}$ značí buď aritmetický průměr $M_{(0)} = \bar{X}$ nebo medián $M_{(1)} = \bar{M}_n$. Třidu modifikovaných JB testů lze pak zapsat jako

$$MJB_{(i_1, i_2, i_3, i_4)} = \frac{n}{C_1} \left(\frac{M_{i_1, 3}}{M_{i_1, 3}^{3/2}} \right)^2 + \frac{n}{C_2} \left(\frac{M_{i_4, 4}}{M_{i_4, 4}^2} - 3 \right)^2. \quad (1)$$

C_1 a C_2 jsou konstanty, které mají jiné hodnoty než v případě klasického Jarque-Bera testu. Tyto konstanty lze získat na základě Monte Carlo simulací (viz níže). Blíže k uvedeným modifikovaným testům viz Střelec a Stehlík, 2009a, b.

Klasická Jarque-Bera testová statistika tak je speciálním případem této obecné třídy, kdy ji lze zapsat jako $MJB_{(0,0,0,0)}$. Testová statistika tak má tvar (Jarque a Bera, 1980)

$$JB = MJB_{(0,0,0,0)} = \frac{n}{6} \left(\frac{\hat{\mu}_3}{\hat{\mu}_2^{3/2}} \right)^2 + \frac{n}{24} \left(\frac{\hat{\mu}_4}{\hat{\mu}_2^2} - 3 \right)^2, \quad (2)$$

kde $\hat{\mu}_3/\hat{\mu}_2^{3/2}$ je výběrový odhad šikmosti rozdělení a $\hat{\mu}_4/\hat{\mu}_2^2$ je výběrový odhad špičatosti rozdělení.

Jak již bylo uvedeno výše, jednotlivé varianty modifikovaných Jarque-Bera testů se od sebe odlišují konstantami C_1 a C_2 . Konstanty pro všechny varianty

$MJB_{(i_1, i_2, i_3, i_4)}$ testů, kde $i \in \{0, 1\}$ značí buď využití aritmetického průměru $M_{(0)} = \bar{X}$, nebo mediánu $M_{(1)} = \bar{M}_n$, jsou uvedeny v Tab I. a II.

Z uvedených výsledků simulace konstanty C_1 je patrné, že její hodnota je závislá na použité charakteristice šikmosti. Naopak v případě konstanty C_2 jsou její asymptotické hodnoty shodné. Z výsledků lze rovněž posoudit rychlost konvergence k teoretické hodnotě, kdy v případě charakteristiky šikmosti nejrychleji konverguje výběrová charakteristika $\hat{\mu}_3/\hat{\mu}_2^{3/2}$ a v případě charakteristiky špičatosti výběrová charakteristika $\hat{m}_4/\hat{\mu}_2^2$.

Získané teoretické (asymptotické) hodnoty konstant C_1 a C_2 jsou pak využity při vlastní konstrukci modifikovaných testů normality MJB. Např. testovou statistiku $MJB_{(1,0,1,0)}$ lze zapsat

$$MJB_{(1,0,1,0)} = \frac{n}{18} \left(\frac{\hat{m}_3}{\hat{\mu}_2^{3/2}} \right)^2 + \frac{n}{24} \left(\frac{\hat{m}_4}{\hat{\mu}_2^2} - 3 \right)^2. \quad (3)$$

Všechny MJB testy mají stejně jako klasický Jarque-Bera test asymptoticky chí-kvadrát rozdělení se dvěma stupni volnosti (blíže viz Střelec a Stehlík, 2009a, b).

Pro všechny sledované testy bude využito exaktních kritických hodnot získaných na základě Monte Carlo simulací, které pro malé a středně velké rozsahy souborů v případě Jarque-Bera testu, robustního Jarque-Bera testu, modifikovaných Jarque-Bera testů a medcouple testů v porovnání s aproximací chí-kvadrát rozdělením věrohodněji zachycují rozdělení testových statistik těchto testů normality.

Jako alternativy jsou zvoleny vybrané směsi normálních rozdělení zajišťující bimodální rozdělení. Směs dvou normálních rozdělení lze obecně zapsat funkcí:

$$F = (1 - p)N(\mu_1, \sigma_1^2) + pN(\mu_2, \sigma_2^2), \quad (4)$$

kde $0 \leq p \leq 1$. Funkce F uvedená v (4) může být interpretována následovně: s pravděpodobností $(1 - p)$ pocházejí hodnoty z normálního rozdělení se střední hodnotou μ_1 a rozptylem σ_1^2 a s pravděpo-

I: Výsledky simulace konstanty C_1 pro různé rozsahy souborů

	25	50	75	100	200	500	1000	$n \rightarrow \infty$
$\hat{\mu}_3/\hat{\mu}_2^{3/2}$	4,71	5,40	5,38	5,69	5,91	6,00	5,98	6
$\hat{\mu}_3/\hat{m}_2^{3/2}$	4,32	5,15	5,21	5,55	5,83	5,97	5,97	6
$\hat{m}_3/\hat{\mu}_2^{3/2}$	15,66	16,36	16,28	16,67	16,87	17,17	17,23	18
$\hat{m}_3/\hat{m}_2^{3/2}$	13,51	15,18	15,44	16,01	16,53	17,03	17,16	18

Počet Monte Carlo simulací je 10 000.

II: Výsledky simulace konstanty C_2 pro různé rozsahy souborů

	25	50	75	100	200	500	1000	$n \rightarrow \infty$
$\hat{\mu}_4/\hat{\mu}_2^2$	13,02	18,46	19,29	21,08	23,07	23,91	23,32	24
$\hat{\mu}_4/\hat{m}_2^2$	12,92	18,27	19,25	20,95	23,02	23,84	23,31	24
$\hat{m}_4/\hat{\mu}_2^2$	19,09	22,05	21,62	23,18	24,10	24,42	23,54	24
$\hat{m}_4/\hat{m}_2^2$	15,28	19,98	20,29	22,04	23,55	24,15	23,42	24

Počet Monte Carlo simulací je 10 000.

dobností p pocházejí hodnoty z normálního rozdělení se střední hodnotou μ_2 a rozptylem σ_2^2 .

V komparační analýze budou konkrétně předpokládány následující hodnoty jednotlivých parametrů a to za účelem analýzy bimodálních rozdělení: $\mu_1 = 0$, $\sigma_1^2 = 1$, $\mu_2 \in \{3, 5\}$, $\sigma_2^2 = 1$ a $p = 0,50$. K porovnání síly sledovaných testů proti uvedeným alternativám bude využito Monte Carlo simulací (počet Monte Carlo simulací je 10000).

VÝSLEDKY

V Tab. III a na Obr. 1 je zachycena síla sledovaných klasických (AD, JB, JBU, LT, SW) a robustních (RJB a MC-LR) testů normality proti bimodálnímu rozdělení vycházející ze směsi dvou normálních rozdělení s funkcí $F = 0,5N(0; 1) + 0,5 N(3; 1)$. Z vý-

sledků je patrné, že nejsilnějším testem je Anderson-Darlingův test (AD) a to pro všechny sledované rozsahy souborů. Naopak nejslabšími testy pro malé a středně velké rozsahy souborů ($n \leq 100$) jsou klasický Jarque-Bera test (JB), Urzuova verze Jarque-Bera testu (JBU) a robustní Jarque-Bera test (RJB), které např. pro rozsah souboru $n = 24$ považují dané bimodální rozdělení za rozdělení normální v téměř 100% případů (síla testu je ve všech třech případech velmi blízká nule). Důvodem je, že dané rozdělení je symetrické, přičemž jeho špičatost pro $n \rightarrow \infty$ je 2,04 (viz Tab. IV), avšak pro malé rozsahy souboru není špičatost významně odlišná od špičatosti normálního rozdělení, což způsobuje, že uvedené testy mají prakticky nulovou sílu.

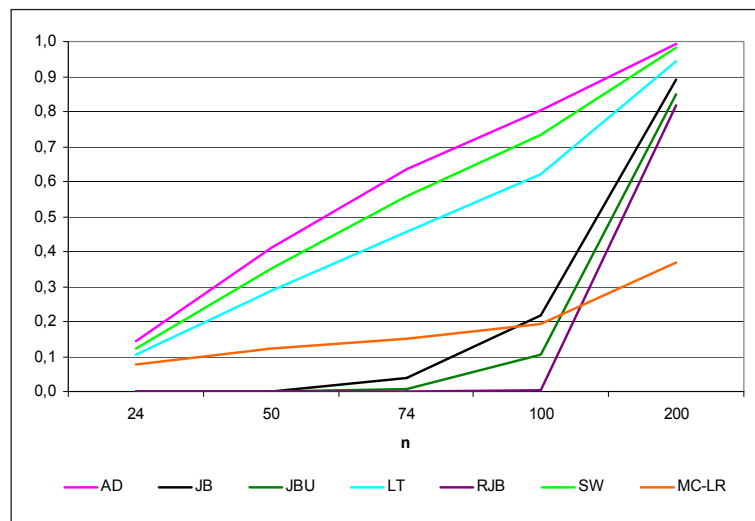
V případě vybraných modifikovaných Jarque-Bera testů je však patrné jisté zlepšení. Výsledky Monte

III: Síla sledovaných klasických a robustních testů normality pro různé rozsahy souborů proti alternativě bimodálního rozdělení s parametry $\mu_2 = 3$ a $\sigma_2^2 = 1$

	24	50	74	100	200
AD	0,1441	0,4113	0,6359	0,8024	0,9914
JB	0,0009	0,0012	0,0401	0,2182	0,8917
JBU	0,0008	0,0000	0,0086	0,1053	0,8492
LT	0,1051	0,2877	0,4555	0,6202	0,9456
MC-LR	0,0771	0,1218	0,1525	0,1935	0,3689
RJB	0,0003	0,0000	0,0000	0,0018	0,8190
SW	0,1219	0,3497	0,5590	0,7334	0,9837

Zvýrazněny jsou testy s nejvyšší silou.

Počet Monte Carlo simulací je 10000, hladina významnosti $\alpha = 0,05$.



1: Síla sledovaných klasických a robustních testů normality pro různé rozsahy proti alternativě bimodálního rozdělení s parametry $\mu_2 = 3$ a $\sigma_2^2 = 1$

Počet Monte Carlo simulací je 10000, hladina významnosti $\alpha = 0,05$.

IV: Charakteristiky šikmosti β_1 a špičatosti $\sqrt{\beta_2}$ pro vybrané směsi normálních rozdělení

		μ_2					
		0	1	2	3	4	5
$\sigma_2^2 = 1$	β_1	0,00	0,00	0,00	0,00	0,00	0,00
	$\sqrt{\beta_2}$	3,00	2,92	2,50	2,04	1,72	1,51

Carlo simulace síly modifikovaných Jarque-Bera testů jsou zachyceny v Tab. V a na Obr. 2. Nejvyšší sílu pro všechny sledované rozsahy souborů vykazuje $MJB_{(1,1,0,1)}$ test, který v porovnání s téměř nulovou silou klasického Jarque-Bera testu $MJB_{(0,0,0,0)}$ pro malé rozsahy souborů ($n < 75$) vykazuje sílu srovnatelnou s robustním medcouple MC-LR testem. Pro větší rozsahy souborů ($n \geq 100$) je síla uvedeného testu již vyšší než síla medcouple MC-LR testu a je srovnatelná dokonce s LT testem.

Zajímavé je i rozpětí síly modifikovaných Jarque-Bera testů, kdy např. pro středně velký rozsah souboru ($n = 100$) je jejich síla v rozmezí intervalu (0,04;

0,42), tj. lze nalézt test (konkrétně $MJB_{(0,0,1,0)}$), který má sílu přibližně srovnatelnou se zvolenou hladinou významnosti $\alpha = 0,05$, a rovněž test (konkrétně $MJB_{(1,1,0,1)}$), jehož síla je vyšší než síla klasického JB testu, JBU testu a obou sledovaných robustních testů normality (RJB, MC-LR), ale zároveň nižší než síla klasických testů normality (AD, LT a SW). Jedná se tedy o test, který představuje kompromis mezi robustností testu a jeho silou.

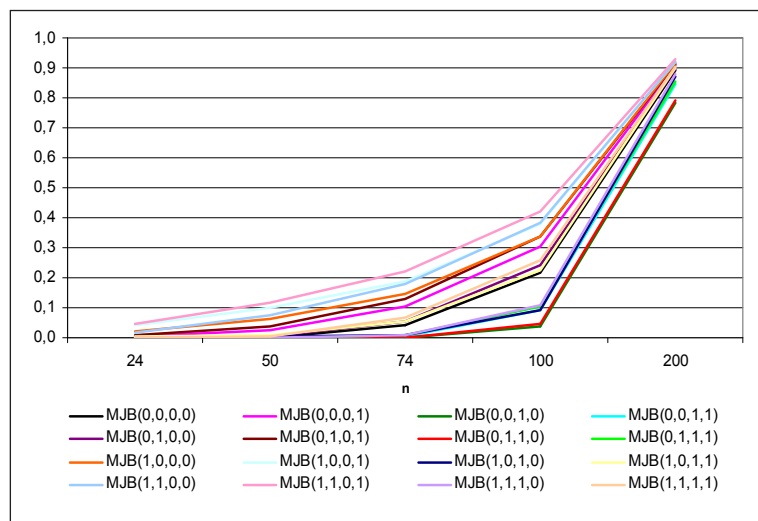
Obdobných výsledků je dosaženo i v případě síly sledovaných testů normality proti bimodálnímu rozdělení vycházející ze směsi dvou normálních rozdělení s funkcí $F = 0,5N(0; 1) + 0,5N(5; 1)$. Z vý-

V: Síla modifikovaných Jarque-Bera testů normality pro různé rozsahy souborů proti alternativě bimodálního rozdělení s parametry $\mu_2 = 3$ a $\sigma_2^2 = 1$

	24	50	74	100	200
$MJB_{(0,0,0,0)}$	0,0009	0,0012	0,0401	0,2182	0,8917
$MJB_{(0,0,0,1)}$	0,0043	0,0260	0,1059	0,3030	0,9094
$MJB_{(0,0,1,0)}$	0,0012	0,0002	0,0009	0,0368	0,7852
$MJB_{(0,0,1,1)}$	0,0012	0,0002	0,0062	0,0903	0,8475
$MJB_{(0,1,0,0)}$	0,0008	0,0024	0,0526	0,2409	0,8959
$MJB_{(0,1,0,1)}$	0,0069	0,0361	0,1294	0,3376	0,9138
$MJB_{(0,1,1,0)}$	0,0013	0,0002	0,0012	0,0439	0,7900
$MJB_{(0,1,1,1)}$	0,0011	0,0001	0,0081	0,1030	0,8529
$MJB_{(1,0,0,0)}$	0,0192	0,0618	0,1442	0,3395	0,9166
$MJB_{(1,0,0,1)}$	0,0400	0,0986	0,1860	0,3772	0,9225
$MJB_{(1,0,1,0)}$	0,0037	0,0028	0,0071	0,0928	0,8694
$MJB_{(1,0,1,1)}$	0,0037	0,0084	0,0526	0,2239	0,8980
$MJB_{(1,1,0,0)}$	0,0178	0,0754	0,1797	0,3824	0,9229
$MJB_{(1,1,0,1)}$	0,0474	0,1177	0,2210	0,4209	0,9291
$MJB_{(1,1,1,0)}$	0,0019	0,0010	0,0077	0,1104	0,8779
$MJB_{(1,1,1,1)}$	0,0010	0,0032	0,0651	0,2577	0,9058

Zvýrazněny jsou testy s nejvyšší silou.

Počet Monte Carlo simulací je 10000, hladina významnosti $\alpha = 0,05$.



2: Síla modifikovaných Jarque-Bera testů normality pro různé rozsahy souborů proti alternativě bimodálního rozdělení s parametry $\mu_2 = 3$ a $\sigma_2^2 = 1$

Počet Monte Carlo simulací je 10000, hladina významnosti $\alpha = 0,05$.

sledků zachycených v Tab. VI a Obr. 3 je patrné, že nejsilnějším testem je opět AD test, následován SW testem a LT testem. Naopak nulovou sílu pro malý rozsah souboru ($n = 24$) vykazuje klasický JB test, JBU test a robustní JB test. V případě klasického JB testu a JBU testu je jejich síla srovnatelná s ostatními testy až pro rozsah souboru $n \geq 74$, pro RJB test až pro rozsah souboru $n \geq 100$. Je tedy patrné, že tyto testy jsou, především pro malé rozsahy souborů, vychýlené, tj. nedokážou sledované bimodální rozdělení odlišit od rozdělení normálního. Důvodem je opět symetrie rozdělení a také pro malé rozsahy souborů nevýznamný rozdíl mezi hodnotou špičatosti daného rozdělení, která je 1,51 pro $n \rightarrow \infty$ (viz Tab. IV), a špičatostí normálního rozdělení.

Jednou z možností, jak zlepšit sílu Jarque-Bera testu, je opět využít jeho modifikace, kdy $MJB_{(1,1,0,1)}$ test vykazuje pro rozsah souboru $n = 24$ sílu 0,07 v porovnání s nulovou silou klasického JB testu (viz Tab. VII a Obr. 4). Nejvyššího rozpětí modifikovaných testů je dosaženo při rozsahu souboru $n = 50$, kdy nejsilnějším testem je $MJB_{(0,0,1,0)}$ test s téměř nulovou silou a nejsilnějším testem je $MJB_{(1,1,0,0)}$ test se silou 0,78. Pro středně velké a velké rozsahy souborů

($n \geq 100$) je již síla těchto modifikovaných testů srovnatelná s klasickými omnibus testy normality.

Využitím kombinace mediánu a aritmetického průměru v klasické Jarque-Bera statistice tak lze dosáhnout zvýšení síly testu proti bimodálnímu rozdělení, avšak uvedené zvýšení síly neplatí obecně. Naopak lze rozlišit skupinu modifikovaných testů, jejichž síla je prakticky shodná s klasickým Jarque-Bera testem, jenž je v Tab. VII označený jako $MJB_{(0,0,0,0)}$ – konkrétně se jedná o skupinu testů $MJB_{(0,0,1,0)}$, $MJB_{(0,0,1,1)}$, $MJB_{(0,1,0,0)}$, $MJB_{(0,1,1,0)}$, $MJB_{(0,1,1,1)}$, $MJB_{(1,0,1,0)}$, $MJB_{(1,1,1,0)}$ a $MJB_{(1,1,1,1)}$ – a skupinu modifikovaných testů, jejichž síla i pro malé rozsahy souborů je v porovnání s klasickým Jarque-Bera testem či robustním Jarque-Bera testem relativně vysoká, což umožňuje jejich praktické využití – konkrétně se jedná o skupinu testů $MJB_{(0,0,0,1)}$, $MJB_{(0,1,0,1)}$, $MJB_{(1,0,0,0)}$, $MJB_{(1,1,0,0)}$ a $MJB_{(1,1,0,1)}$, přičemž nejsilnější z nich je posledně jmenovaný, tj. test, jehož testová statistika má podobu

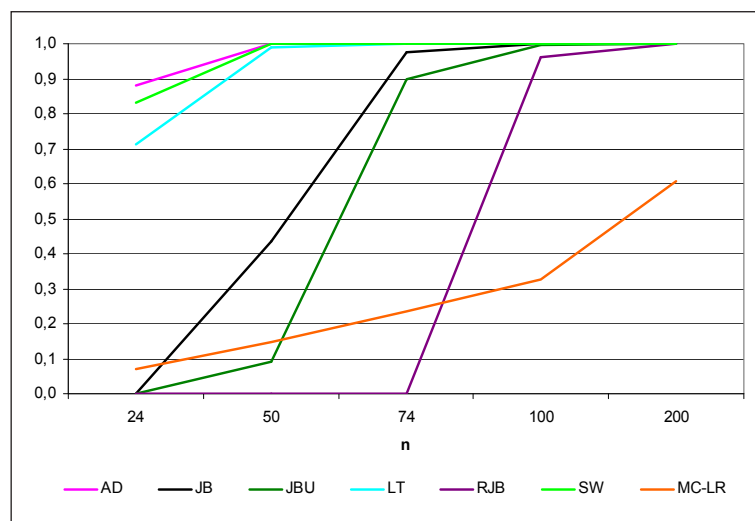
$$MJB_{(1,1,0,1)} = \frac{n}{18} \left(\frac{\hat{m}_3}{\hat{m}_2^{3/2}} \right)^2 + \frac{n}{24} \left(\frac{\hat{\mu}_4}{\hat{m}_2^2} - 3 \right)^2, \quad (5)$$

VI: Síla sledovaných klasických a robustních testů normality pro různé rozsahy souborů proti alternativě bimodálního rozdělení s parametry $\mu_2 = 5$ a $\sigma_2^2 = 1$

	24	50	74	100	200
AD	0,8813	0,9995	1,0000	1,0000	1,0000
JB	0,0000	0,4336	0,9755	0,9993	1,0000
JBU	0,0000	0,0923	0,8987	0,9982	1,0000
LT	0,7118	0,9905	1,0000	1,0000	1,0000
MC-LR	0,0694	0,1479	0,2338	0,3252	0,6061
RJB	0,0000	0,0000	0,0003	0,9627	1,0000
SW	0,8317	0,9983	1,0000	1,0000	1,0000

Zvýrazněny jsou testy s nejvyšší silou.

Počet Monte Carlo simulací je 10000, hladina významnosti $\alpha = 0,05$.



3: Síla sledovaných klasických a robustních testů normality pro různé rozsahy souborů proti alternativě bimodálního rozdělení s parametry $\mu_2 = 5$ a $\sigma_2^2 = 1$

Počet Monte Carlo simulací je 10000, hladina významnosti $\alpha = 0,05$.

a který má stejně jako klasický Jarque-Bera test asymptoticky chí-kvadrát rozdělení se dvěma stupni volnosti, tj. normalitu zamítá v případech, kdy $MJB_{(1,1,0,1)} > \chi^2_{1-\alpha}(2)$.

DISKUSE

Jak vyplývá z výše uvedených výsledků provedené simulační komparační analýzy, nejsilnějším testem z analyzovaných testů proti bimodálnímu rozdělení je Anderson-Darlingův test (Anderson a Darling, 1952) následovaný Shapiro-Wilkovým testem (Shapiro a Wilk, 1965) a Lillieforsovým testem (Lil-

iefors, 1967). Naopak testy založené na momentových charakteristikách, jako je klasický Jarque-Bera test, Urzuova verze Jarque-Bera testu (Urzua, 1996) či robustní Jarque-Bera test zkonstruovaný Gelovou a Gastwirthem, 2009, vykazují pro malé rozsahy souborů nulovou sílu, na což již v případě klasického Jarque-Bera testu upozornili Thadewald a Büning, 2007.

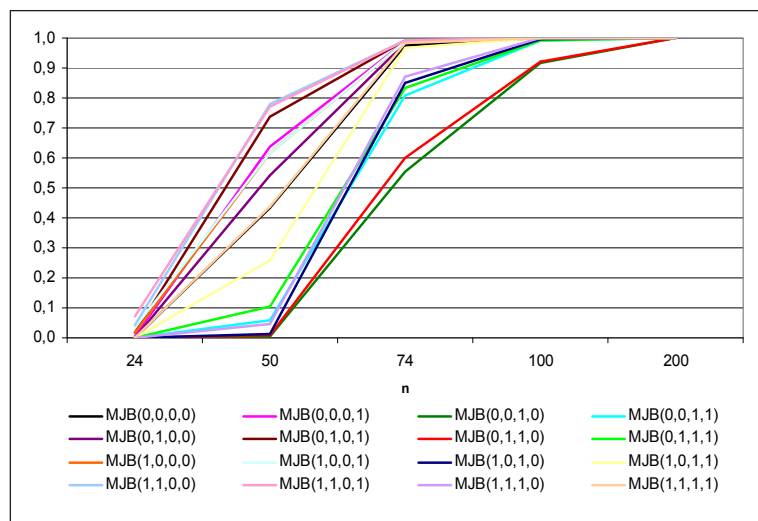
Uvedená nulová síla testu pro malé rozsahy souborů však neplatí pro všechny testy, které jsou založeny na momentových charakteristikách – např. D'Agostinův test založený na transformovaných charakteristikách šikmosti a špičatosti (blíže k testu

VII: Síla modifikovaných Jarque-Bera testů normality pro různé rozsahy souborů proti alternativě bimodálního rozdělení s parametry $\mu_2 = 5$ a $\sigma_2^2 = 1$

	24	50	74	100	200
MJB _(0,0,0,0)	0,0000	0,4336	0,9755	0,9993	1,0000
MJB _(0,0,0,1)	0,0044	0,6390	0,9892	0,9995	1,0000
MJB _(0,0,1,0)	0,0000	0,0024	0,5556	0,9163	0,9993
MJB _(0,0,1,1)	0,0000	0,0598	0,8068	0,9903	1,0000
MJB _(0,1,0,0)	0,0000	0,5412	0,9827	0,9995	1,0000
MJB _(0,1,0,1)	0,0166	0,7363	0,9927	0,9995	1,0000
MJB _(0,1,1,0)	0,0000	0,0065	0,5984	0,9229	0,9993
MJB _(0,1,1,1)	0,0000	0,1022	0,8337	0,9915	1,0000
MJB _(1,0,0,0)	0,0153	0,6174	0,9872	0,9994	1,0000
MJB _(1,0,0,1)	0,0333	0,6157	0,9877	0,9994	1,0000
MJB _(1,0,1,0)	0,0001	0,0118	0,8481	0,9979	1,0000
MJB _(1,0,1,1)	0,0005	0,2568	0,9672	0,9992	1,0000
MJB _(1,1,0,0)	0,0437	0,7803	0,9937	0,9995	1,0000
MJB _(1,1,0,1)	0,0693	0,7704	0,9933	0,9996	1,0000
MJB _(1,1,1,0)	0,0000	0,0449	0,8711	0,9982	1,0000
MJB _(1,1,1,1)	0,0000	0,4366	0,9822	0,9992	1,0000

Zvýrazněny jsou testy s nejvyšší silou.

Počet Monte Carlo simulací je 10000, hladina významnosti $\alpha = 0,05$.



4: Síla modifikovaných Jarque-Bera testů normality pro různé rozsahy proti alternativě bimodálního rozdělení s parametry $\mu_2 = 5$ a $\sigma_2^2 = 1$

Počet Monte Carlo simulací je 10000, hladina významnosti $\alpha = 0,05$.

viz D'Agostino a Stephens, 1986) vykazuje proti alternativě bimodálního rozdělení přibližně stejnou či dokonce nepatrně vyšší sílu než Anderson-Darlingův test (viz Střelec a Stehlík, 2009b). Dalším příkladem nenulové síly testu založeném na momentových charakteristikách proti alternativě bimodálního rozdělení jsou vybrané výše vymezené modifikované Jarque-Bera testy kombinující aritmetický průměr a medián v odhadech centrálních momentů klasické Jarque-Bera testové statistiky. Modifikované Jarque-Bera testy sice nedosahují síly Anderson-Darlingova testu, ale v porovnání s klasickým Jarque-Bera testem, Urzuovou verzí Jarque-Bera testu či robustním Jarque-Bera testem dosahují relativně vysoké síly testu. Konkrétním příkladem může být $MJB_{(1,1,0,1)}$ test.

Vybrané modifikované testy tak lze úspěšně využít při testování proti alternativám bimodálního rozdělení, neboť představují kompromis mezi robustností a silou testu (např. medcouple test (Brys et al., 2004 a 2008), který lze považovat za velmi robustní

test, vykazuje v porovnání s vybranými modifikovanými testy nižší sílu).

Modifikované testy lze použít i pro testování proti jiným alternativám než je bimodální rozdělení. Např. je lze využít pro testování proti alternativám těžko-chvostých rozdělení jako je Cauchyho, Laplaceovo či Studentovo t -rozdělení se třemi stupni volnosti či pro testování proti alternativám lehkochvostých rozdělení jako je exponenciální či lognormální rozdělení. Např. $MJB_{(1,1,1,1)}$ test vykazuje vyšší sílu než klasický Jarque-Bera test či robustní Jarque-Bera test proti alternativám zešikmeným s lehkými chvosty, avšak v případě těžkochvostých rozdělení vykazuje $MJB_{(1,1,1,1)}$ test v porovnání s klasickým Jarque-Bera testem a robustním Jarque-Bera testem stejnou či nižší sílu (viz Střelec, 2008). Modifikované $MJB_{(i_1,i_2,i_3,i_4)}$ testy tak poskytují velkou škálu testů, které je nutné podrobit dalšímu výzkumu a to především v porovnání s klasickými a robustními testy normality, neboť představují třídu testů, které využívají klasických i robustních momentových charakteristik.

SOUHRN

Článek se zabývá srovnáním síly vybraných klasických omnibus testů normality (Shapiro-Wilkova testu, Lilliefors (Kolmogorov-Smirnovova) testu, Anderson-Darlingova testu, klasického Jarque-Bera testu a Jarque-Bera-Urzuova testu), robustních testů normality (robustního Jarque-Bera testu, medcouple testu) a modifikovaných Jarque-Bera testů proti bimodálnímu rozdělení.

Na základě Monte Carlo simulací bylo zjištěno, že nejsilnějším testem proti alternativě bimodálního rozdělení je Anderson-Darlingův test a Shapiro-Wilkův test. Naopak klasický Jarque-Bera test, Urzuova verze Jarque-Bera testu a robustní Jarque-Bera test vykazují pro malé rozsahy souborů nulovou sílu. V těchto případech vykazují vyšší sílu vybrané modifikované Jarque-Bera testy, které kombinují medián a aritmetický průměr ve výběrových odhadech k -tých centrálních momentů. Nejsilnější z těchto modifikací je modifikovaný Jarque-Bera test $MJB_{(1,1,0,1)}$, jehož testová statistika má podobu

$$MJB_{(1,1,0,1)} = \frac{n}{18} \left(\frac{\hat{m}_3}{\hat{m}_2^{3/2}} \right)^2 + \frac{n}{24} \left(\frac{\hat{\mu}_4}{\hat{m}_2^2} - 3 \right)^2,$$

kde $\hat{\mu}_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k$ a $\hat{m}_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{M}_n)^k$ jsou výběrové odhady k -tého centrálního momentu pro všechny pozitivní hodnoty k , X_i pro $i = 1, \dots, n$ je výběr z iid náhodné proměnné, $\bar{X}$ je výběrový odhad prvního obecného momentu (aritmetický průměr), $\bar{M}_n$ je výběrový medián, $\hat{m}_3/\hat{m}_2^{3/2}$ je výběrový odhad šikmosti rozdělení a $\hat{\mu}_4/\hat{m}_2^2$ je výběrový odhad špičatosti rozdělení, a která má stejně jako klasická Jarque-Bera testová statistika asymptoticky chí-kvadrát rozdělení se dvěma stupni volnosti, tj. normalitu zamítá v případech, kdy $MJB_{(1,1,0,1)} > \chi^2_{1-\alpha}(2)$.

V porovnání s Anderson-Darlingovým a Shapiro-Wilkovým testem vykazuje $MJB_{(1,1,0,1)}$ test pro malé a středně velké rozsahy souboru nižší sílu, avšak v porovnání s nulovou silou klasického Jarque-Bera testu je zde patrné zlepšení – např. proti alternativě bimodálního rozdělení, které vznikne jako směs dvou normálních rozdělení s funkcí $F = 0,5N(0; 1) + 0,5N(5; 1)$ pro rozsah souboru $n = 24$ vykazuje $MJB_{(1,1,0,1)}$ test sílu 0,07, Jarque-Bera test sílu 0,00 a Anderson-Darlingův test sílu 0,88, a pro rozsah souboru $n = 50$ vykazuje $MJB_{(1,1,0,1)}$ test sílu 0,77, Jarque-Bera test sílu 0,43 a Anderson-Darlingův test sílu téměř 1,00.

testy normality, robustní testy, Monte Carlo simulace, komparace síly testů, bimodální rozdělení

Příspěvek vznikl za podpory programu AKTION Česká republika – Rakousko, spolupráce ve vědě a vzdělání, projektu č. 51p7 s názvem „Robust testing for the normality and its applications for the economy, sports and Basel III“.

SUMMARY

This paper deals with comparison of the power of selected classical omnibus normality tests (the Shapiro-Wilk test, the Lilliefors test, the Anderson-Darling test, the classical Jarque-Bera test and

the Jarque-Bera-Urzuza test), the robust normality tests (the robust Jarque-Bera test, the Medcouple test) and the modified Jarque-Bera normality tests to detect a bimodal distribution.

On the basis of the Monte Carlo simulations we can conclude that the Anderson-Darling and the Shapiro-Wilk tests outperform the other tests including the Lilliefors test and medcouple test for bimodal alternative distribution, especially in small sample sizes. On the other hand the classical Jarque-Bera, the Jarque-Bera-Urzuza and the robust Jarque-Bera tests are biased, especially in small sample sizes again. For these cases selected modified Jarque-Bera tests show a higher power than the classical Jarque-Bera test. These tests combine the median and the arithmetic mean for the k -th population central moment estimation. Consequently, the highest power show $MJB_{(1,1,0,1)}$ test with test statistic

$$MJB_{(1,1,0,1)} = \frac{n}{18} \left(\frac{\hat{m}_3}{\hat{m}_2^{3/2}} \right)^2 + \frac{n}{24} \left(\frac{\hat{\mu}_4}{\hat{m}_2^2} - 3 \right)^2,$$

where $\hat{\mu}_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k$ and $\hat{m}_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{M}_n)^k$ are sample estimates of the k -th population central moment for any positive integer k , X_i for $i = 1, \dots, n$ is a sample of independent and identically distributed random variables, $\bar{X}$ is the sample estimate of first population moment (sample mean), $\bar{M}_n$ is the sample median, $\hat{m}_3/\hat{m}_2^{3/2}$ is the sample estimate of the skewness and $\hat{\mu}_4/\hat{m}_2^2$ is the sample estimate of the kurtosis. $MJB_{(1,1,0,1)}$ test statistic follows the same asymptotically distribution as the classical Jarque-Bera test statistic, i.e. the chi-square with two degrees of freedom. Therefore, the implied one-sided rejection region is "reject H_0 : normality, if $MJB_{(1,1,0,1)} > \chi^2_{1-\alpha}(2)$."

The $MJB_{(1,1,0,1)}$ test shows a lower power than the Anderson-Darling and the Shapiro-Wilk tests, especially in small sample sizes. On the other hand the $MJB_{(1,1,0,1)}$ test shows a higher power than the classical Jarque-Bera test – e.g. $MJB_{(1,1,0,1)}$ test shows power 0.07 against the bimodal distribution with mixture of normal distributions function $F = 0.5N(0; 1) + 0.5N(5; 1)$ and sample size $n = 24$ (in contrast, the classical Jarque-Bera test is biased (the power is zero) and power of the Anderson-Darling test is 0.88) and for sample size $n = 50$ $MJB_{(1,1,0,1)}$ test shows power 0.77 (in contrast, power of the classical Jarque-Bera test is 0.43 and power of the Anderson-Darling test is nearly 1.00).

Research was supported by Project 51p7 AKTION Czech Republic-Austria: „Robust testing for the normality and its applications for the economy, sports and Basel II“.

LITERATURA

- ANDERSON, T. W., DARLING, D. A., 1952: Asymptotic theory of certain „goodness-of-fit“ criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23(2). p. 193–212.
- BRYŠ, G., HUBERT, M., STRUYF, A., 2004: A robustification of the Jarque-Bera test of normality. COMPSTAT 2004, *Proceedings in Computational Statistics*, ed. J. Antoch, Springer, Physica Verlag, p. 753–769.
- BRYŠ, G., HUBERT, M., STRUYF, A., 2008: Goodness-of-fit tests based on a robust measure of skewness. *Computational Statistics*, 23(3). p. 429–442.
- D'AGOSTINO, R., STEPHENS, M., 1986: *Goodness-of-fit Techniques*. Marcel Dekker, New York.
- GEL, Y. R., GASTWIRTH, J. L., 2009: A robust modification of the Jarque-Bera test of normality. *Economics Letters*, 99(1), p. 30–32.
- JARQUE, C. M., BERA, A. K., 1980: Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3). p. 255–259.
- LILLIEFORS, H., 1967: On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62(318). p. 399–402.
- SHAPIRO, S. S., WILK, M. B., 1965: An analysis of variance test for normality (complete samples). *Biometrika*, 52(3 and 4). p. 591–611.
- STŘELEČ, L., 2008: Comparison of power of modified Jarque-Bera normality tests and selected tests of normality. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis: Acta of Mendel University of Agriculture and Forestry*. sv. LVI, č. 6, s. 137–148. ISSN 1211-8516.
- STŘELEČ, L., STEHLÍK, M., 2009a: On robust testing for normality. *Proceedings of the 6th St. Petersburg Workshop on Simulation*. St. Petersburg: VVM com. Ltd., Ed. by S. M. Ermakov, V. B. Melas and A. N. Pepelyshev, 2009, p. 755–760.
- STŘELEČ, L., STEHLÍK, M., 2009b: Some properties of robust tests for normality and their applications. *IFAS research report*.
- URZUA, C. M., 1996: On the correct use of omnibus tests for normality. *Economics Letters*, 53(3), p. 247–251.
- THADEWALD, T., BÜNNING, H., 2007: Jarque-Bera test and its competitors for testing normality – a power comparison. *Journal of Applied Statistics*, 34(1). p. 87–105.
- THODE, H. C., 2002: *Testing for Normality*. Marcel Dekker, New York.

Adresa

Ing. Luboš Střelec, Ústav statistiky a operačního výzkumu, Mendelova zemědělská a lesnická univerzita v Brně, 613 00 Brno, Česká republika, xstřelec@node.mendelu.cz