

Czesław Domański

Uniwersytet Łódzki

TESTY ZGODNOŚCI DOTYCZĄCE ROZKŁADÓW WYKŁADNICZYCH TYPU KOŁMOGOROWA-SMIRNOWA

1. Wstęp

Często interesuje nas odpowiedź na pytanie, jaki rozkład ma badana zmienna w populacji generalnej. Hipoteza dotyczy bądź nieznanego kształtu rozkładu, bądź tego, że rozpatrywana zmienna losowa ma rozkład należący do określonej rodziny, a więc, że jej dystrybuanta należy do odpowiedniej rodziny dystrybuant. Przy zastosowaniu testów parametrycznych i metod estymacji najczęściej musimy znać kształt rozkładu badanej zmiennej. Czasami jednak mamy bardzo mało informacji o populacji, stawiamy wówczas hipotezę, że badana zmienna X ma taki a taki rozkład i na podstawie wylosowanej próby sprawdzamy, czy należy tę hipotezę odrzucić, czy też materiał empiryczny nie daje podstaw do tego, aby ją odrzucić. Testy takie stanowią ogólniejszą grupę, gdyż dotyczą nie poszczególnych parametrów, lecz całej funkcji rozkładu. Noszą one nazwę *testów zgodności*. Konstrukcja testów zgodności wymaga wprowadzenia pewnej miary odległości rozkładów. Istnieje kilka sposobów określania odległości między porównywanymi rozkładami. Najczęściej taka miara jest oparta na porównaniu dystrybuant rozkładu empirycznego $F_n(x)$ i rozkładu teoretycznego $F(x)$ następującej postaci:

$$\delta_1 = \sup_x |F_n(x) - F(x)|. \quad (1)$$

Rozpatruje się również dwa warianty tej miary:

$$\delta_2 = \sup_x [F_n(x) - F(x)] \quad (2)$$

oraz:

$$\delta_3 = \sup_x [F(x) - F_n(x)]. \quad (3)$$

Porównując z kolei gęstość rozkładów, otrzymujemy inną miarę:

$$\delta_4 = \int_{-\infty}^{\infty} [f_n(x) - f(x)]^2 dx \quad (4)$$

lub porównując funkcje prawdopodobieństwa, mamy:

$$\delta_5 = \sum_{i=1}^{\infty} [P_n(x_i) - P(x_i)]^2, \quad (5)$$

a w szczególnym przypadku, porównując liczebność empiryczną n_i z liczebnościami teoretycznymi (np_i), otrzymujemy:

$$\delta_6 = \sum_{i=1}^{\infty} (n_i - np_i)^2. \quad (6)$$

Ogólnie można zdefiniować odległość między dwoma rozkładami o gęstościach $f_n(x)$ oraz $f(x)$:

$$\delta = \sup_{x \in A} |f_n(x) - f(x)|, \quad (7)$$

gdzie A jest zbiorem, na którym są określone funkcje $f_n(x)$ oraz $f(x)$.

Zauważmy, że wszystkie niemal testy zgodności oparte na mutacjach jednej z wymienionych miar odległości.

Przedstawione miary odległości nie wyczerpują wszystkich możliwości określania odchyleń między rozkładami.

2. Sformułowanie problemu

Niech $X_1, \dots, X_n$ będzie próbą losową o ciągłym rozkładzie z dystrybuantą $F(x)$. Hipotezę o zgodności rozkładu empirycznego z rozkładem teoretycznym formułujemy w następujący sposób.

Niech $\tilde{F}$ oznacza klasę dystrybuant zmiennych losowych, $\tilde{F}_w$ klasę dystrybuant rozkładu wykładniczego, a $\tilde{G}$ klasę dystrybuant zmiennych losowych mających różny od zera trzeci moment centralny μ_3 i skończony czwarty moment centralny μ_4 , przy czym $\tilde{F} \cap \tilde{G} = \emptyset$.

Określamy hipotezę $H_0 : F(x) \in \tilde{F}$ oraz alternatywną hipotezę $H_1 : F(x) \in \tilde{G}$, że funkcja $F(x)$ należy do klasy dystrybuant $\tilde{F}$ lub $\tilde{G}$. Hipotezy H_0 lub H_1 będą nazwane *prostymi*, jeśli rozkłady należące do klasy $\tilde{F}$, $\tilde{F}_w$ lub $\tilde{G}$ mają znane

parametry rozkładu. W przypadku gdy parametry rozkładów zmiennych losowych szacujemy na podstawie próby, wówczas hipotezy H_0 oraz H_1 będą nazwane hipotezami *złożonymi*. Zgodnie z ogólnie przyjętą symboliką, wskaźnik w nawiasie oznacza numer odpowiedniego elementu w próbie uporządkowanej w ciąg nie malejący obserwacji $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$.

Testy zgodności, w zależności od konstrukcji funkcji testowych (sprawdzianu testu), możemy podzielić na 4 klasy:

- A – testy oparte na porównaniu gęstości,
- B – testy oparte na porównaniu dystrybuant,
- C – testy oparte na momentach,
- D – testy oparte na statystykach pozycyjnych.

Do pierwszej klasy zaliczamy test zgodności χ^2 Pearsona oraz test Davida–Hellwiga. Do klasy B należą wszystkie testy oparte na statystykach Kołmogorowa–Smirnowa lub ich funkcjach. Testy należące do klas C oraz D weryfikują hipotezy o normalności.

3. Testy statystyczne

Przypuśćmy, że $x_{(1)}, \dots, x_{(n)}$ jest uporządkowaną próbą n niezależnych obserwacji wylosowanych z populacji o dystrybuancie $F(x)$ i że chcemy testować hipotezę $H_0 : F(x) = F(x, \theta)$, gdzie $F(x, \theta)$ jest wyspecyfikowaną dystrybuantą ciągłą zależną od wektora $\theta = (\theta_1, \dots, \theta_p)$ p nieznanymi parametrów. Durbin [6] wykazał, że jeżeli próba pochodzi z populacji o dystrybuancie:

$$F(x) = 1 - \exp(-\theta x), \quad x > 0, \quad (8)$$

przy czym θ nie jest znane, to weryfikując $H_0 : F(x) \in \tilde{F}_w = 1 - \exp(\theta x)$ należy skorzystać z testu Kołmogorowa–Smirnowa, którego sprawdziany są następujące:

$$D^- = \max_{1 \leq i \leq n} \left(z_{(i)} - \frac{2i-1}{2n} \right) + \frac{1}{2n}, \quad (9)$$

$$D^+ = \max_{1 \leq i \leq n} \left(\frac{2i-1}{2n} - z_{(i)} \right) + \frac{1}{2n}, \quad (10)$$

$$D = \max\{D^+, D^-\},$$

gdzie:

$$z_{(i)} = 1 - \exp\left(-\frac{x_{(i)}}{\bar{x}}\right).$$

Hipotezę H_0 odrzucamy, gdy zachodzi jedna z nierówności $\sqrt{n}D > \sqrt{n}D_\alpha$ lub $\sqrt{n}D^+ > \sqrt{n}D_\alpha^+$, gdzie $\sqrt{n}D_\alpha, \sqrt{n}D_\alpha^+$, odczytane są z tab. 1 dla wybranego poziomu istotności α oraz n .

Tabela 1. Ocena kwantyli statystyki D_α^w

Statystyka	Poziom istotności α				
	0,15	0,10	0,05	0,025	0,01
D_α^w	0,926	0,990	1,094	1,190	1,308

Źródło: [5] na podstawie [9].

Stephens [10] zaproponował modyfikację sprawdzianów testu Kołmogorowa-Smirnowa dla testowania hipotezy H_0 , że rozkład jest wykładniczy o nieznanym parametrze θ . Zmodyfikowana postać jest następująca:

$$D^w = \left(D - \frac{0,2}{\sqrt{n}} \right) \left(\sqrt{n} + 0,26 + \frac{0,5}{\sqrt{n}} \right), \quad (11)$$

przy czym:

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

Hipotezę H_0 odrzucamy, jeśli dla wybranego poziomu istotności α zachodzi nierówność $D^w > D_\alpha^w$, przy czym D_α^w odczytujemy z tab. 1. Przybliżona postać statystyki D^w (por. (11)) całkowicie wystarcza do celów praktycznych, a do analiz teoretycznych dotyczących mocy testów należy posłużyć się dokładnymi kwantylami. Problemy własności testów wykładniczych są obecnie analizowane. W artykule tym zostaną przedstawione pewne testy tej klasy, które mają dobre własności.

Najbardziej znanym testem opartym na dystrybuancie empirycznej weryfikacji hipotezy o wykładniczości rozkładu są testy Kołmogorowa-Smirnowa i Cramera-von Misesa, których statystyki wyrażają się odpowiednio wzorami (por. [3]):

$$KS_n = \sup_{x \geq 0} |F_n(x) - (1 - \exp(-x))| = \max \left\{ \max_{1 \leq j \leq n} \left[\frac{j}{n} - z_{(j)} \right], \max_{1 \leq j \leq n} \left[z_{(j)} - \frac{j-1}{n} \right] \right\} \quad (12)$$

i

$$\omega_n^2 = \int_0^\infty (F_n(x) - (1 - \exp(-x)))^2 \exp(-x) dx = \frac{1}{12n} + \sum_{j=1}^n \left(z_{(j)} - \frac{j-1}{2n} \right)^2. \quad (13)$$

Inną ciekawą propozycją testu wykładniczości zasugerowali Baringhaus i Henze [1]. Zgodnie z założeniem $0 < \mu < \infty$, rozkład X jest wykładniczy wtedy i tylko wtedy, gdy $E(X - t | X > t) = \mu$ dla każdego $t > 0$. Ponieważ ten warunek jest równoznaczny z $E[\min(X, t)] = \mu \cdot F(t)$ dla każdego $t > 0$, to statystyka Kołmogorowa–Smirnova oraz Cramera–von Misesa przyjmuje odpowiednio postać:

$$\overline{KS}_n = \sqrt{n} \sup_{t \geq 0} \left| \frac{1}{n} \sum_{j=1}^n \min(Y_j, t) - \frac{1}{n} \sum_{j=1}^n \mathbf{1}\{Y_j \leq t\} \right| = \sqrt{n} \max(KS_n^+, KS_n^-), \quad (14)$$

gdzie

$$KS_n^+ = \max_{j=0,1,\dots,n-1} \left[\frac{1}{n} (y_{(1)} + \dots + y_{(j)} + y_{(j+1)}) \left(1 - \frac{j}{n} \right) - \frac{j}{n} \right],$$

$$KS_n^- = \max_{j=0,1,\dots,n-1} \left[\frac{j}{n} - \frac{1}{n} (y_{(1)} + \dots + y_{(j)} + y_{(j+1)}) \left(1 - \frac{j}{n} \right) \right],$$

oraz statystykę Cramera–von Misesa

$$\begin{aligned} \overline{CM}_n &= n \int_0^\infty \left(\frac{1}{n} \sum_{j=1}^n \min(Y_j, t) - \frac{1}{n} \sum_{j=1}^n \mathbf{1}\{Y_j \leq t\} \right)^2 \exp(-t) dt = \\ &= \frac{1}{n} \sum_{j,k=1}^n \left[2 - 3e^{-\min(Y_j, Y_k)} - 2mn(Y_j, Y_k)(e^{-Y_j} + e^{-Y_k}) + 2e^{-Y_k} + 2e^{-\max(Y_j, Y_k)} \right]. \end{aligned} \quad (15)$$

Odrzucamy hipotezę H_0 dla dużych wartości $\overline{KS}_n$ oraz $\overline{CM}_n$. Asymptotyczne rozkłady zerowe $\overline{KS}_n$ oraz $\overline{CM}_n$ są takie same jak graniczne rozkłady klasycznych statystyk Kołmogorowa–Smirnova i Cramera–von Misesa w przypadku testowania jednorodności na jednostkowym przedziale.

4. Badanie mocy testów

Wyniki badania Monte Carlo zostały przeprowadzone w celu oceny mocy omawianych testów. Empiryczne wartości krytyczne o wielkości α dla rozważanych testów otrzymano z 50 000 replikacji. Z tymi wartościami krytycznymi dokonano symulacji mocy tych testów na próbie o liczebności $n = 20$ i $n = 50$ z takich rozkładów, jak:

- rozkład Weibulla o gęstości $\theta x^{\theta-1} \exp(-x^\theta)$, oznaczony przez $W(\theta)$,
- rozkład Gamma o gęstości $\Gamma(\theta)^{-1} x^{\theta-1} \exp(-x)$, oznaczony przez $\Gamma(\theta)$,

- rozkład lognormalny $LN(\theta)$ o gęstości $(\theta x^{-1})(2\pi)^{-1/2} \exp(-\log x)^2/(2\theta^2)$,
- rozkład półnormalny HN o gęstości $(2/\pi)^{-1/2} \exp(-x^2/2)$,
- rozkład jednostajny U o gęstości $1, 0 \leq x \leq 1$,
- zmodyfikowana wartość ekstremalna $EV(\theta)$, o dystrybuancie $1 - \exp(\theta^{-1}(1 - e^2))$,
- Rozkład wskaźnika awaryjności $LF(\theta)$ o gęstości $(1 + \theta x) \exp(-x - \theta x^2/2)$,
- Rozkład Dhillona [4] $DL(\theta)$ o dystrybuancie $1 - \exp(-(\log(x+1))^{\theta+1})$,
- Rozkład Chena [2] $CH(\theta)$ o dystrybuancie $1 - \exp(2(1 - e^{x^\theta}))$.

5. Wnioski i uwagi końcowe

1. Dla $n = 20$ najmocniejszymi testami spośród rozważanych testów z tab. 2 są testy oparte na $\overline{KS}_n, \overline{CM}_n, \overline{S}_n, \overline{\omega}_n^2, \overline{KS}_n$. Wniosek ten jest prawdziwy dla $n = 20$, z wyjątkiem testu opartego na $\overline{KS}_n$, który jest mniej korzystny i zastępowany jest klasycznym testem Cramera–von Misesa.

Tablica 2. Empiryczna moc niektórych testów wykładniczości dla $n = 20; n = 50, \alpha = 0,05$ i rozkładów alternatywnych

Rozkłady alternatywne	$n = 20$				$n = 50$			
	$\overline{KS}$	$\overline{CM}$	$\overline{\omega}^2$	KS	$\overline{KS}$	$\overline{CM}$	$\overline{\omega}^2$	KS
$W(0,8)$	14	22	20	17	35	46	44	38
$W(1,4)$	35	35	34	28	71	77	75	64
$\Gamma(0,4)$	62	75	76	71	97	99	99	98
$\Gamma(1,0)$	5	5	5	5	5	5	5	5
$\Gamma(2,0)$	46	47	47	40	86	90	90	83
$LN(0,8)$	28	27	33	30	62	60	76	71
$LN(1,5)$	55	66	62	58	92	95	94	91
HN	24	22	21	18	50	53	48	39
U	72	70	66	52	99	99	98	93
$CH(0,5)$	47	61	61	56	90	94	95	92
$CH(1,0)$	18	16	14	13	36	37	32	26
$CH(1,5)$	79	83	79	67	*	*	*	98
$LF(2,0)$	32	30	28	24	65	69	64	53
$LF(4,0)$	44	43	41	34	82	87	83	72
$EV(0,5)$	18	16	14	13	36	37	32	26
$EV(1,5)$	48	47	43	35	88	90	86	75
$DL(1,0)$	22	21	23	20	43	44	52	46
$DL(1,5)$	62	63	65	56	96	97	98	95

Źródło: opracowanie własne na podstawie [7].

2. Dla $n = 50$ różnice mocy są mniej wyraźne i z wyjątkiem nieco mniej mocniejszego klasycznego testu Cramera–von Misesa, wszystkie te testy zachowują się podobnie.

Literatura przedmiotu dotycząca nowych i klasycznych testów wykładniczości ciągle się rozwija. Wynika to z jednej strony z potrzeby weryfikacji założenia wykładniczości oraz szerszego wykorzystania w wielu modelowych sytuacjach testowania życia i niezawodności, a z drugiej strony – z precyzji wnioskowania statystycznego. Autor zamierza kontynuować badania w tym zakresie, które pozwolą, jak się wydaje, z grupy kilkudziesięciu testów wybrać kilka, które zapewnią „najlepsze” wnioskowanie statystyczne.

Literatura

- [1] Baringhaus L., Henze N., *Tests of Fit for Exponentiality Based on a Characterization via the Mean Residual Life Function*, „Statistical Papers” 2000, 41, 225-236.
- [2] Chen Z., *A New Two-parameter Lifetime Distribution with Bathtub Shape or Increasing Failure Rate Function*, „Statistics and Probability Letters” 2000, 49, 155-161.
- [3] D’Agostino R.B., *Tests for the Normal Distribution, Goodness-of-Fit Techniques*, New York 1986, s. 367-419.
- [4] Dhillon B.S., *Lifetime Distributions*, „IEEE Transactions on Reliability” 1981, 30, 457-459.
- [5] Domański C., *Testy statystyczne*, PWE, Warszawa 1990.
- [6] Durbin J., Kolmogorov-Smirnov, *Tests when Parameters are Estimated with Application to Tests of Exponentiality and Test on Spacings*, „Biometrika” 1975, 62, s. 5-22.
- [7] Henze N., Klar B., Meintanis S.G., *Invariant Tests for Symmetry About an Unspecified Point Based on the Empirical Characteristics Function*, „Journal of Multivariate Analysis” 2003, 87, 275-297.
- [8] Henze N., Meintanis S.G., *Recent and Classical Tests for Exponentiality: A Partial Review with Comparisons*, „Metrika” 2005, 61, 29-45.
- [9] Pearson E.S., Hartley H.O., *Biometrika Tables for Statisticians*, Cambridge University Press, Cambridge 1972.
- [10] Stephens M.A., *Use of the Kolmogorov-Smirnov, Cramer-von Mises and Related Statistics without Extensive Tables*, „Journal of the Royal Statistical Society” 1970, 32, 115-122.

THE GOODNESS OF FIT TESTS FOR KOLMOGOROV-SMIRNOV EXPONENTIAL DISTRIBUTIONS

Summary

The paper presents tests verifying hypothesis stating that the population distribution is exponential. We will limit ourselves to the presentation of a couple of tests based on the Kolmogorov–Smirnov and Cramer–von Mises statistics. The tests' power is presented for a few alternative distributions for $n = 20$ and $n = 50$.

In the light of the results presented the Cramer–von Mises test generally turns out to be more powerful.