Тестови експоненцијалности засновани на емпиријским Лапласовим трансформацијама

Бојана Милошевић
Марко Обрадовић

Универзитет у Београду,
Математички факултет

Шести Симпозијум - Математика и примене, Београд

Тестови сагласности

Тестови сагласности с расподелом (класом расподела) су статистички тестови где се тестира $H_0(F \in \mathcal{F})$ против $H_1(F \notin \mathcal{F})$.

- Колмогоров-Смирнов тест

$$T_n = \sup_{t \in R} |F_n(t) - F_0(t)|,$$

- Крамер-фон Мизесов тест

$$T_n = \int_{-\infty}^{\infty} (F_n(t) - F_0(t))^2 f_0(t) dt,$$

- засновани су на растојању претпостављене функције расподеле F_0 и њене оцене, емпиријске функције расподеле F_n
- Гливенко-Кантелијева теорема о униформној конвергенцији емпиријских функција расподеле ка одговарајућој функцији расподеле

Тестови сагласности засновани на карактеризацијама

- идеја је потекла од Линика (1952)
- први тестови: Вашичек (1976) (максимална ентропија), Кул (1977,1978) (одсуство меморије)
- често су слободни од неког параметра расподеле па се могу тестирати сложене хипотезе
- иако се многи од њих формирају на сличан начин сваки је оригиналан јер користи сопствено “скривено” својство

Тестови сагласности засновани на карактеризацијама на основу једнаке расподељености статистика

$X_1, X_2, ...X_n$ п.с.у.

$$H_0 : X \in \mathcal{E}(\lambda).$$

Тестови су засновани на карактеризацији облика

$$\omega_1(X_1, ..., X_m) \stackrel{d}{=} \omega_2(X_1, ..., X_m),$$

где су ω_1 и ω_2 линеарне функције чланова узорка и његових статистика поретка

- карактеризација: Ахсанулах (1978)

  $$(n-j)(X_{(j+1),n} - X_{(j),n}) \stackrel{d}{=} (n-k)(X_{(k+1),n} - X_{(k),n}),$$

  за неко j, k, $1 \leq j \leq k < n$

- тест: Никитин, Волкова (2015)

Тестови сагласности засновани на карактеризацијама на основу једнаке расподељености

- карактеризација:
  - Чакраборти, Јанев (2013)

$$X_1 + \frac{1}{2}X_2 + \frac{1}{3}X_3 \stackrel{d}{=} \max\{X_1, X_2, X_3\}$$

  - Обрадовић(2014)

$$X_0 + X_{(2),3} \stackrel{d}{=} X_{(3),3}$$

  уопштења: Милошевић, Обрадовић (2014)
- тестови: Никитин, Волкова (2014), Милошевић (2015)

Тестови сагласности засновани на карактеризацијама на основу једнаке расподељености

$$
\begin{aligned}
I_n &= \int_0^\infty (H_n^1(t) - H_n^2(t)) dF_n(t) \\
K_n &= \sup_{t\geq 0} |H_n^1(t) - H_n^2(t)|
\end{aligned}
$$

$$
\begin{aligned}
H_n^k(t) &= \frac{1}{\binom{n}{m}} \sum_{1\leq i_1 < \cdots i_m \leq n} I\{\omega_k(X_{i_1}, ..., X_{i_m}) < t\} \\
H_n^k(t) &= \frac{1}{n^m} \sum_{1\leq i_1, i_m \leq n} I\{\omega_k(X_{i_1}, ..., X_{i_m}) < t\}
\end{aligned}
$$

U-статистике и V-статистике

$$U_n = \frac{1}{\binom{n}{m}} \sum_{1 \le i_1 < \cdots i_m \le n} \Phi(X_{i_1}, \ldots, X_{i_m}),$$

$$V_n = \frac{1}{n^m} \sum_{i_1, \ldots, i_m = 1}^{n} \Phi(X_{i_1}, \ldots, X_{i_m})$$

$$\phi(s) = E(\Phi(X_{i_1}, \ldots, X_{i_m}) | X_1 = s) \quad \sigma^2 = D(\phi(X_1))$$

недегенерисане $(\sigma^2 > 0)$

$\sqrt{n}(U_n - E\Phi) \sim \sqrt{n}(V_n - E\Phi) \sim \mathcal{N}(0, m^2\sigma^2)$ (Хефдинг 1948.)

U-емпиријске V-емпиријске функције расподеле

$$H_n(t) = \frac{1}{\binom{n}{m}} \sum_{1 \le i_1 \le \cdots \le i_m \le n} \mathrm{I}\{h(X_{i_1}, \ldots, X_{i_m}) \le t\}$$

$$H_n(t) = \frac{1}{n^m} \sum_{1 \le i_1, \cdots, i_m \le n} \mathrm{I}\{h(X_{i_1}, \ldots, X_{i_m}) \le t\}, \ \ t \in \mathbb{R}$$

- за $m = 1$ функција $H_n(t) = F_n(t)$
- за недегенерисане U (V)-емпиријске функције расподеле важи теорема Гливенко-Кантелијевог типа о униформној конвергенцији ка одговарајућој функцији расподеле (Силверман, 1983.)

Друге емпиријске функције

Лапласова трансформација случајне величине X

$$L(t) = E(e^{-tX}) = \int_0^\infty e^{-tx} dF(x)$$

Емпиријска Лапласова трансформација

$$L_n(t) = \frac{1}{n} \sum_{i=1}^{n} e^{-tX_i} = \int_0^\infty e^{-tx} dF_n(x)$$

Карактеристична функција

$$\varphi(t) = E(e^{itX}) = \int_{-\infty}^\infty e^{itx} dF(x)$$

Емпиријска карактеристична функција

$$\varphi_n(t) = \frac{1}{n} \sum_{i=1}^{n} e^{itX_i} = \int_0^\infty e^{itx} dF_n(x)$$

Тестови сагласности засновани на другим емпиријским функцијама

- Хензе (1993), Хензе и Меинтанис (2002)

$$Y_i = \frac{X_i}{\bar{X}_n},\ i = 1, 2...n, \hat{\lambda} = \frac{1}{\bar{X}_n}$$

$$H_0 : Y \sim \mathcal{E}(1)$$

$$T_n = \int_0^\infty (\tilde{L}_n(t) - \tilde{L}(t))^2 e^{-at} dt,\ \ \tilde{L}_n(t) = \frac{1}{n}\sum_{i=1}^{n} e^{-tY_i}$$

- Хензе и Клар (2002) за инверзну Гаусову расподелу

Тестови сагласности засновани на другим емпиријским функцијама

- Барингхаус, Хензе (1991)

$$(\lambda + t)L'(t) + L(t) = 0$$
$$(\hat{\lambda} + t)L_n'(t) + L_n(t) = 0$$

$$T_n = \int_0^\infty ((\hat{\lambda} + t)L_n'(t) + L_n(t))^2 \bar{X}_n e^{-a\bar{X}_n t} dt$$

- Меинтанис, Илиопоулос (2003) за Рејлијеву расподелу
- Хензе, Меинтанис, Ебнер (2012) за гама расподелу

Тестови сагласности засновани на другим емпиријским функцијама

- нађене су граничне расподеле тест статистика
- за узорке малог обима критичне вредности се добијају Монте Карло методом
- лако се имплементирају
- правилним избором параметра a се може повећати моћ теста

Нови тестови сагласности

Тест статистике које предлажемо су облика

$$J_n = \int\limits_0^\infty (L_n^{(1)}(t) - L_n^{(2)}(t))\bar{X}_n e^{-a\bar{X}t} dt,$$

$$L^{(k)} = \frac{1}{m!\binom{n}{m}} \sum_{i_1<\ldots<i_m} \sum_{\pi(m)} e^{-t\omega_k(X_{\pi(i_1)},\ldots,X_{\pi(i_m)})}, \quad k = 1, 2.$$

Нови тестови сагласности

$$J_n = \frac{\bar{X}_n}{n^{[m]}} \sum_{i_1<\dots<i_m} \sum_{\pi(m)} \frac{1}{a\bar{X}_n + \omega_1(X_1,..,X_m)} - \frac{1}{a\bar{X}_n + \omega_2(X_1,..,X_m)}.$$

Асимптотска својства тест статистика

Расподела статистике J_n не зависи од параметра скалирања λ. Користимо помоћне функције

$$J^*(\mu) = \frac{\mu}{n^{[m]}} \sum_{i_1<\ldots<i_m} \sum_{\pi(m)} \frac{1}{a\mu + \omega_1(X_1,..,X_m)} - \frac{1}{a\mu + \omega_2(X_1,..,X_m)}.$$

$J^*(1)$ je U-статистика са језгром

$$\Phi(X_1,...,X_m) = \frac{1}{m!} \sum_{\pi(m)} \frac{1}{a + \omega_1(X_1,..,X_m)} - \frac{1}{a + \omega_2(X_1,..,X_m)}.$$

Асимптотска својства тест статистика

Ако је језгро недегенерисано

$$\begin{aligned}\sqrt{n}J_n^*(1) &\longrightarrow \mathcal{N}(0, m^2\sigma_\Phi^2),\\ \sigma_\Phi^2 &= E(\phi^2(X_1))\\ \phi(s) &= E(\Phi(X_1,...,X_m)|X_1 = s).\end{aligned}$$

Како $\bar{X}_n \longrightarrow \mu$ с. с, користећи резултате Рандлеса из 1982. показали смо да J_n и $J^*(\mu)$ имају исту граничну расподелу .

Приближна Бахадурова ефикасност

$$e_{T,V}(\alpha,\beta,\theta) = \frac{N_T(\alpha,\beta,\theta)}{N_V(\alpha,\beta,\theta)}.$$

$$\lim_{n\to\infty} \frac{\ln(1 - G_{T_n}(t))}{n} = -f(t), \quad t \in I,$$

Претпоставимо да

- $\lim_{n\to\infty} P\{T_n \le t\} = F(t)$ постоји, где је F недегенерисана функција расподеле,
- $\lim_{t\to\infty} \log(1 - F(t)) = -\frac{a_T t^2}{2}$
- $\lim_{n\to\infty} T_n \stackrel{P_\theta}{=} b_T(\theta) > 0$, постоји за $\theta \in \Theta_1$.

Приближна релативна Бахадурова ефикасност статистике T_n у односу на другу статистику V_n је

$$e^*_{T,V} = \frac{c^*_T}{c^*_V}, \quad c^*_T = a_T b^2_T(\theta).$$

Асимптотска својства тест статистика

Посматрајмо класу блиских алтернатива

$$\mathcal{G} = \{G(x,\theta), 0 < \theta < c\},$$

$$G(x,0) = 1 - e^{-\lambda x},\ x > 0,\ \lambda > 0.$$

Функција густине $g(x,\theta)$ је диференцијаблина по θ у околини нуле. Означимо $h(x) = \frac{\partial}{\partial\theta} g(x,\theta)|_{\theta=0}$.

Асимптотска својства тест статистика

$$\bar{X}_n \longrightarrow \mu(\theta) = E_\theta(X_1) \text{ с. с.}$$

Користећи закон великих бројева за U-статистике са оцењеним параметром (Ајверсон, Рандлес 1989) добијамо да се лимеси у вероватноћи P_θ статистике J_n и помоћне функције J_n^* поклапају.

Асимптотска својства тест статистика

Лема

Лимес у вероватноћи P_θ *статистике* J_n *ако важи блиска алтернатива из фамилије* $\mathcal{G}$ *је*

$$b(\theta) = m \int_0^\infty \phi(x)h(x)dx \cdot \theta + o(\theta),\ \theta \to 0.$$

Асимптотска својства тест статистика

Лема

Коефицијент a_J у изразу за приближни Бахадуров нагиб статистике J_n

$$a_J = \frac{1}{m^2\sigma_\Phi^2}.$$

Карактеризације експоненцијалне расподеле

- Десу (1971)

$$2\min(X, Y) \stackrel{d}{=} X$$

- Пури-Рубин (1970)

$$|X - Y| \stackrel{d}{=} X$$

Тест статистике

$$J_n^{\mathcal{D}} = \frac{\bar{X}_n}{n(n-1)} \sum_{i_1<i_2} \sum_{\pi(2)} \frac{1}{a\bar{X}_n + 2\min(X_{\pi(i_1)}, X_{\pi(i_2)})} - \frac{1}{a\bar{X}_n + X_{\pi(i_1)}},$$

$$J_n^{\mathcal{P}} = \frac{\bar{X}_n}{n(n-1)} \sum_{i_1<i_2} \sum_{\pi(2)} \frac{1}{a\bar{X}_n + |X_{\pi(i_1)} - X_{\pi(i_2)}|} - \frac{1}{a\bar{X}_n + X_{\pi(i_1)}},$$

Тест статистике

Пројекције језгра одговарајућих U-статистика $J^{\mathcal{D}*}$ и $J^{\mathcal{P}*}$ на X_1 ако важи H_0 су

$$\begin{aligned}\varphi^{\mathcal{D}}(s) =& E(\Phi^{\mathcal{D}}(X_1, X_2)|X_1 = s) = -\frac{1}{2(a+s)} + \frac{1}{2}e^a Ei(-a)\\ &+ \frac{1}{2(a+2s)}e^{-s}\Big(2 + ae^{a/2+s}\Gamma(0, \frac{a}{2}) + 2e^{\frac{a}{2}+s}s\Gamma(0, \frac{a}{2})\\ &- ae^{a/2+s}\Gamma(0, \frac{a}{2}+s) - 2e^{\frac{a}{2}+s}s\Gamma(0, \frac{a}{2}+s)\Big),\end{aligned}$$

$$\begin{aligned}\varphi^{\mathcal{P}}(s) =& E(\Phi^{\mathcal{P}}(X_1, X_2)|X_1 = s) = \frac{1}{2(a+s)} - \frac{1}{2}e^a Ei(-a)\\ &+ e^{-a-s}(e^{2a}Ei(-a) + Ei(a) - Ei(a+s)),\end{aligned}$$

Конкурентске тест статистике

$$I_n^{\mathcal{D}} = \int_0^{\infty} (G_n(t) - F_n(t)) dF_n(t),$$

$$I_n^{\mathcal{P}} = \int_0^{\infty} (G_n(t) - F_n(t)) dF_n(t),$$

где су

$$F_n(t) = \frac{1}{n} \sum_i I\{X_i < t\}; \qquad G_n(t) = \frac{1}{\binom{n}{2}} \sum_{i<j} I\{2\min(X_i, X_j) < t\};$$

$$H_n(t) = \frac{1}{\binom{n}{2}} \sum_{i<j} I\{|X_i - X_j| < t\};$$

Алтернативне расподеле

- Вејбулова расподела са густином

$$g(x,\theta) = e^{-x^{1+\theta}}(1+\theta)x^{\theta},\ \theta > 0,\ x \geq 0;$$

- Макехамова расподела са густином

$$g(x,\theta) = (1+\theta(1-e^{-x}))\exp(-x-\theta(e^{-x}-1+x)),\ \theta > 0,\ x \geq 0;$$

- гама расподела са густином

$$g(x,\theta) = \frac{x^{\theta}}{\Gamma(\theta+1)}e^{-x}, \theta > 0, x \geq 0;$$

- расподела линеарне стопе отказа (ЛСО) са густином

$$g(x,\theta) = (1+\theta x)e^{-x-\theta\frac{x^2}{2}}, \theta > 0, x \geq 0;$$

Приближна асимптотска релативна Бахадурова ефикасност

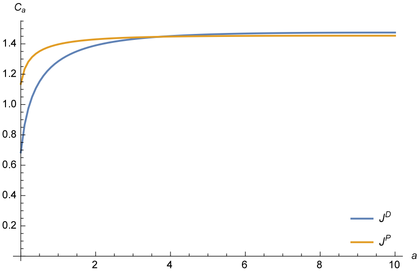
Табела : ПАРБЕ $J_{n,a}^{\mathcal{D}}$ против $I_n^{\mathcal{D}}$

	$a=1$	$a=2$	$a=5$
Вејбул	1.11	1.21	1.27
гама	1.09	1.08	1.03
Макехам	1.11	1.44	1.76
ЛСО	1.33	2.05	3.12

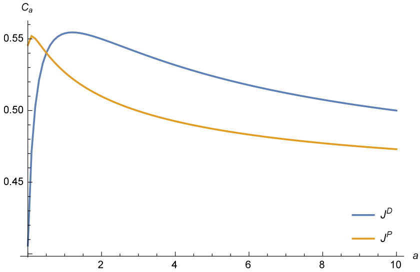
Табела : ПАРБЕ $J_{n,a}^{\mathcal{P}}$ против $I_n^{\mathcal{P}}$

	$a=1$	$a=2$	$a=5$
Вејбул	1.03	1.06	1.07
гама	1.04	1	0.96
Макехам	1	1.11	1.2
ЛСО	1.32	1.68	2.06

Приближна асимптотска релативна Бахадурова ефикасност

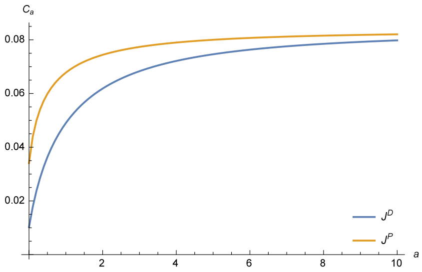


Слика : Приближан нагиб за Вејбул алтернативе

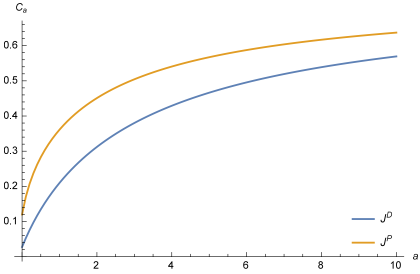


Слика : Приближан нагиб за гама алтернативе

Приближна асимптотска релативна Бахадурова ефикасност



Слика : Приближан нагиб за Макехам алтернативе



Слика : Приближан нагиб за ЛСО алтернативе

Емпиријске моћи тестова

Табела : $n = 20,\ \alpha = 0.05$

Алт.	$W(1.4)$	$\Gamma(2)$	HN	U	$CH(0.5)$	$CH(1)$	$CH(1.5)$	$LF(2)$	$LF(4)$	$EW(0.5)$	$EW(1.5)$
ЕП	36	48	21	66	63	15	84	28	42	15	45
$\overline{KS}$	35	46	24	72	47	18	79	32	44	18	48
$\overline{CM}$	35	47	22	70	61	16	83	30	43	16	47
ω^2	34	47	21	66	61	14	79	28	41	14	43
КС	28	40	18	52	56	13	67	24	34	13	35
КЛ	29	44	16	61	77	11	76	23	34	12	37
С	35	46	21	70	63	15	84	29	42	15	46
ЦО	37	54	19	50	80	13	81	25	37	13	37
$J^{\mathcal{D},1}$	47	70	24	51	17	17	17	31	44	18	40
$J^{\mathcal{D},2}$	48	67	26	61	20	19	19	35	49	20	46
$J^{\mathcal{D},5}$	50	67	30	71	22	23	22	40	50	22	55
$J^{\mathcal{P},1}$	49	65	29	73	21	21	21	38	52	21	53
$J^{\mathcal{P},2}$	50	64	31	77	23	23	23	40	56	23	57
$J^{\mathcal{P},5}$	49	62	31	79	22	22	23	41	56	23	58

Емпиријске моћи тестова

Табела : $n = 50,\ \alpha = 0.05$

Алт.	$W(1.4)$	$\Gamma(2)$	HN	U	$CH(0.5)$	$CH(1)$	$CH(1.5)$	$LF(2)$	$LF(4)$	$EW(0.5)$	$EW(1.5)$
ЕП	80	91	54	98	94	38	100	69	87	38	90
$\overline{KS}$	71	86	50	99	90	36	100	65	82	36	88
$\overline{CM}$	77	90	53	99	94	37	100	69	87	37	90
ω^2	75	90	48	98	95	32	100	64	83	32	86
КС	64	83	39	93	92	26	98	53	72	26	75
КЛ	72	93	37	97	99	23	100	54	75	23	79
С	79	90	54	99	94	38	100	69	87	38	90
ЦО	82	96	45	91	99	30	100	60	80	30	78
$J^{\mathcal{D},1}$	82	97	41	80	28	27	28	56	75	28	70
$J^{\mathcal{D},2}$	87	98	53	92	37	38	38	70	86	38	82
$J^{\mathcal{D},5}$	88	97	60	98	45	44	45	76	91	44	90
$J^{\mathcal{P},1}$	86	97	53	98	40	41	40	72	89	41	88
$J^{\mathcal{P},2}$	89	97	62	99	46	46	47	77	92	47	92
$J^{\mathcal{P},5}$	88	96	66	99	49	48	50	80	93	50	94

Тестови сагласности

- L. Baringhaus, N. Henze, A class of consistent tests for exponentiality based on the empirical Laplace transform, Ann. Inst. Statist. Math. vol.43(3) (1991) 551–564.
- N. Henze, A new flexible class of omnibus tests for exponentiality, Comm. Statist. Theory Methods vol.22(1) (1993) 115–133.
- N. Henze, B. Klar, Goodness-of-Fit Tests for the inverse Gaussian distribution based on the empirical Laplace transform, Ann. Inst. Statist. Math. vol.54(2) (2002) 425–444.
- N. Henze, S.G. Meintanis, Goodness-of-fit tests based on a new characterization of the exponential distribution, Comm. Statist. Theory Methods vol.31(9) (2002) 1479–1497.

Тестови сагласности

- N. Henze, S.G. Meintanis, Recent and classical tests for exponentiality: a partial review with comparisons, Metrika vol.61(1) (2005) 29–45.
- N. Henze, S.G. Meintanis, B. Ebner, Goodness-of-fit tests for the gamma distribution based on the empirical Laplace transform, Comm. Statist. Theory Methods vol.41(9) (2012) 1543-1556.
- S.G. Meintanis, G. Iliopoulos, Tests of fit for the Rayleigh distribution based on the empirical Laplace transform, Ann. Inst. Statist. Math. vol.55(1) (2003) 137–151.
- S.G. Meintanis, Ya.Yu. Nikitin, A.V. Tchirina, Testing exponentiality against a class of alternatives which includes the RNBUE distributions based on the empirical Laplace transform, J. Math. Sci. (N.Y.) vol.145(2) (2007) 4871–4879.

Карактеризације

- M. Ahsanullah, On a characterization of the exponential distribution by spacings, Ann. Inst. Statist. Math. vol.30(1) (1978) 163–166.
- M. Ahsanullah, M. Rahman, A Characterization of the Exponential Distribution, J. Appl. Probab. vol.9(2) (1972) 457–461.
- B.C. Arnold, J.A. Villasenor, Exponential characterizations motivated by the structure of order statistics in samples of size two, Statist. Probab. Lett. vol.83(2) (2013) 596–601.
- S. Chakraborty, G.P. Yanev, Characterization of exponential distribution through equidistribution conditions for consecutive maxima, J. Stat. Appl. Probab. vol.2(3) (2013) 237–242.
- M.M. Desu, A characterization of the exponential distribution by order statistics, Ann. Math. Statist. vol.42(2) (1971) 837–838.

Карактеризације

- B. Milošević, M. Obradović, Some characterizations of exponential distributions based on order statistics, arXiv:1412.5019 (2014)
- M. Obradović, Three Characterizations of Exponential Distribution Involving the Median of Sample of Size Three, arXiv preprint (2014) arXiv:1412.2563.
- P.S. Puri, H. Rubin, A Characterization Based on Absolute Difference of Two I.I.D. Random Variables, Ann. Math. Statist. vol.41(6) (1970) 2113-2122.
- H.-J. Rossberg, Characterization of the exponential and the Pareto distributions by means of some properties of the distributions which the differences and quotients of order statistics are subject to, Statistics vol.3(3) (1972) 207–216.

Остало

- W. Hoeffding, A Class of Statistics with Asymptotically Normal Distribution, Ann. Math. Statist. vol.19(3) (1948) 293–395.
- H.K. Iverson, R.H. Randles, The effects on convergence of substituting parameter estimates into U-statistics and other families of statistics. Probability Theory and Related Fields, vol. 81(3) (1989) 453–471.
- Ya.Yu. Nikitin, Asymptotic efficiency of nonparametric tests. Cambridge University Press, New York, 1995.
- R.H. Randles, On the asymptotic normality of statistics with estimated parameters, Ann. Statist. vol.10(2) (1982) 462–474.