Обзор критериев показательности

А. П. Рогожников, Б. Ю. Лемешко
Новосибирский государственный технический университет

Аннотация **– Рассматривается широкий набор критериев для проверки гипотезы о показательности распределения вероятностей. Методами статистического моделирования исследуются распределения статистик критериев при истинности проверяемой гипотезы, вычисляются оценки и проводится сравнительный анализ мощности критериев по отношению к конкурирующим альтернативам с различными формами функции интенсивности отказов, делаются выводы о предпочтительности использования тех или иных критериев при наличии определенных конкурирующих альтернатив.**

Ключевые слова **– критерий показательности, показательное распределение, мощность критерия.**

I. Введение

Проверке гипотез о принадлежности выборки показательному закону распределения посвящено множество работ, в которых авторами предлагаются различные статистические критерии. Обилие критериев обусловлено частым использованием модели показательного закона в приложениях. А частота использования не в последнюю очередь определяется тем, что применение такой простой модели во многих ситуациях позволяет найти решение задачи с опорой только на аналитические методы.

Наличие множества критериев ставит перед практиками непростую задачу выбора, так как имеющаяся в публикациях информация не позволяет однозначно отдать предпочтение некоторому определенному критерию, особенно при наличии определенных конкурирующих законов (конкурирующих гипотез). Конечно, для проверки показательности может использоваться совокупность критериев согласия, однако опыт показывает [1, 2], что именно среди критериев, целенаправленно построенных для проверки гипотезы о принадлежности выборки только одному конкретному закону, оказываются критерии наиболее мощные.

В некоторых работах, например, в [3] и [4], рассматривался достаточно широкий ряд критериев показательности, и методами статистического моделирования исследовалась их мощность по отношению к некоторым важным конкурирующим гипотезам. Полученные результаты позволили выделить критерии, перспективные как для применения в случае конкурирующих законов с определенной формой функции интенсивности отказов, так и предпочтительные для применения в случае наличия широкого класса конкурирующих гипотез.

Некоторые критерии в данной работе исключены из рассмотрения, так как они показали неудовлетворительные свойства в важных случаях. В данной работе проведен сравнительный анализ критериев, которые относятся к перспективной группе. Кроме этого рассмотрен критерий показательности Большева, в своё время предложенный для проверки гипотезы о показательности нескольких малых выборок.

II. Постановка задачи

Пусть $Exp(\theta)$ — показательное распределение с плотностью $f(x) = \exp(-x/\theta)/\theta$, $x \ge 0$, $\theta \equiv \lambda^{-1} > 0$ а $X_1, ..., X_n$ – заданные независимые наблюдения неотрицательной случайной величины X с плотностью $f(x)$ и функцией распределения $F(x)$. Проверяемая сложная гипотеза H_0: X подчиняется закону распределения $Exp(\theta)$ при некотором $\theta > 0$.

В статистиках критериев будем использовать масштабированные наблюдения $Y_j = X_j / \hat{\theta}_n$ или их преобразованные значения $Z_j = 1 - \exp(-Y_j)$, $1 \le j \le n$, где $\hat{\theta}_n = \bar{X}_n = n^{-1} \sum_{j=1}^{n} X_j$ – оценка максимального правдоподобия параметра θ.

Порядковые статистики величин X_j, Y_j и Z_j будем обозначать $X_{(j)}$, $Y_{(j)}$ и $Z_{(j)}$ соответственно. Положим $D_j = (n - j + 1)\left(X_{(j)} - X_{(j-1)}\right)$, $1 \le j \le n$, $X_{(0)} \equiv 0$.

III. Теория

2.1. F-критерий Гнеденко

F-критерий Гнеденко [3, 5] предназначен для проверки гипотезы показательности против конкурирующей гипотезы H_1: распределение имеет монотонную интенсивность отказов. Статистика критерия имеет вид

$$Q_R = \sum_{j=1}^{R} D_j / R \Bigg/ \sum_{j=R+1}^{n} D_j / (n - R).$$

Если проверяемая гипотеза верна, статистика Q_R подчиняется *F*-распределению Фишера с $2R$ и

$2(n-R)$ степенями свободы. Гипотеза H_0 отклоняется при малых и больших значениях Q_R, при этом в первом случае интенсивность отказов принимает убывающий, а во втором – возрастающий характер. Проведенное моделирование с получением оценок мощности критерия Гнеденко показало, что при конкурирующих законах с монотонной интенсивностью отказов предпочтительным является выбор $R=[0.3n]$ из рассмотренных целых значений $R=[0.1n]$, $[0.2n]$, …, $[0.9n]$.

2.2. Критерий Харриса

Критерий Харриса представляет собой модификацию F-критерия Гнеденко. Данный критерий был предложен Харрисом [6] и обсуждался в [3] и [5]. Статистика критерия

$$Q_R' = \left(\sum_{j=1}^{R} D_j + \sum_{j=n-R+1}^{n} D_j\right)\left(\sum_{j=R+1}^{n-R} D_j \Big/ (n-2R)\right)^{-1} \Big/ 2R .$$

подчиняется F-распределению с $4R$ и $2(n-2R)$ степенями свободы при истинности проверяемой гипотезы. Гипотеза отклоняется при малых и больших значениях Q_R'.

Нами получено, что данный критерий является достаточно мощным против конкурирующих законов с выпуклой интенсивностью отказов, в том числе против логнормального распределения, и имеет низкую мощность по отношению к конкурирующим законам с монотонной интенсивностью отказов. Проведенное нами моделирование показало, что в данном случае наибольшей мощности критерий достигает при $R=[0.1n]$.

2.3. Критерий Холландера-Прошана

Критерий Холландера-Прошана [7, 3] применяется при наличии альтернатив со свойством «новое лучше старого» («старое лучше нового»). Это свойство можно интерпретировать как утверждение, что вероятность $\bar{F}(x)$ безотказной работы нового устройства к возрасту x больше (меньше) вероятности $\bar{F}(x+y)/\bar{F}(y)$ безотказной работы устройства в дополнительное время x после периода безотказной работы y. Статистика критерия:

$$T = \sum_{i>j>k} \psi\left(X_{(i)}, X_{(j)} + X_{(k)}\right),\ \psi(a,b) = \begin{cases} 1, a > b, \\ 0, a \le b. \end{cases}$$

Критерий двусторонний, авторы критерия приводят таблицы приближенных нижних и верхних критических значений и следующую нормальную аппроксимацию:

$$T^* = \left(T - \mathrm{E}[T \mid H_0]\right)\left(\mathrm{D}[T \mid H_0]\right)^{-1/2},$$

где $\mathrm{E}(T \mid H_0) = n(n-1)(n-2)/8$ и $\mathrm{D}[T \mid H_0] = 1.5n(n-1)(n-2) \times$ $\times\left[2(n-3)(n-4)/2592 + 7(n-3)/432 + 1/48\right]$.

2.4. Критерий Гини

Этот двусторонний критерий со статистикой

$$G_n = \sum_{j,k=1}^{n} \left|Y_j - Y_k\right| \Big/ 2n(n-1)$$

рассмотрен в [8, 3, 4].

Предельным распределением величины

$$G_n^* = \left[12(n-1)\right]^{1/2}\left\{G_n - 1/2\right\}$$

является стандартный нормальный закон [8], который, как показала наша проверка, хорошо описывает G_n^* при $n \ge 10$. Критерий Гини эквивалентен критерию вкладов [4] со статистикой $S_n = 2n - 2n^{-1}\sum_{j=1}^{n} jY_{(j)}$, которая связана со статистикой Гини соотношением $(n-1)^{-1} S_n = 1 - G_n$. Мы рассматривали статистику G_n, выражая ее через S_n, так как последняя имеет меньшую вычислительную сложность.

2.5. Критерии, основанные на эмпирической функции распределения

2.5.1. Критерий Колмогорова

В критерии согласия Колмогорова в качестве меры отличия эмпирического распределения от показательного закона используется величина

$$D_n = \max\left\{\max_{i \le j \le n}\left[\frac{j}{n} - Z_{(j)}\right], \max_{i \le j \le n}\left[Z_{(j)} - \frac{j-1}{n}\right]\right\}.$$

Для уменьшения зависимости распределения статистики критерия Колмогорова от объема выборки следует использовать статистику с поправкой Большева [9]:

$$K_n = \left(6n \cdot D_n + 1\right)\Big/ 6\sqrt{n}\ .$$

2.5.2. Критерий Крамера-Мизеса-Смирнова

В критерии согласия Крамера-Мизеса-Смирнова при проверке гипотезы о принадлежности выборки показательному закону используется статистика

$$CMS_n = \frac{1}{12n} + \sum_{j=1}^{n}\left(Z_{(j)} - \frac{2j-1}{2n}\right)^2 .$$

2.5.3. Критерий Андерсона-Дарлинга

Статистика критерия согласия Андерсона-Дарлинга при проверке гипотезы о принадлежности выборки показательному закону имеет вид

$$AD_n = -n - 2\sum_{j=1}^{n}\left[\frac{2j-1}{2n}\ln Z_j + \left(1 + \frac{2j-1}{2n}\right)\ln\left(1 - Z_j\right)\right].$$

Гипотеза о показательности по критериям Колмогорова, Крамера-Мизеса-Смирнова и Андерсона-

Дарлинга отвергается при больших значениях статистик.

Хорошей [10] моделью распределения статистики K_n при справедливости сложной проверяемой гипотезы H_0 и объемах выборок $n > 25$ является гамма-распределение $\gamma(5.1092; 0.0861; 0.2950)$ (рис. 1) с функцией плотности

$$f(x) = \frac{1}{\theta_1^{\theta_0}\Gamma(\theta_0)}(x-\theta_2)^{\theta_0 - 1} e^{-(x-\theta_2)/\theta_1},\ x > \theta_2,$$

статистики CMS_n – распределение Sb Джонсона $Sb(3.3738; 0.2145; 1.0792; 0.011)$ (рис. 1) с функцией плотности

$$f(x) = \frac{\theta_1\theta_2}{\sqrt{2\pi}(x-\theta_3)(\theta_2+\theta_3-x)}\times$$
$$\times\exp\left\{-\frac{1}{2}\left(\theta_0+\theta_1\ln\frac{x-\theta_3}{\theta_2+\theta_3-x}\right)^2\right\},\ x\in[\theta_3, \theta_3+\theta_2],$$

статистики AD_n – распределение Sb Джонсона $Sb(3.8386; 1.3429; 7.500; 0.090)$ (рис. 1).

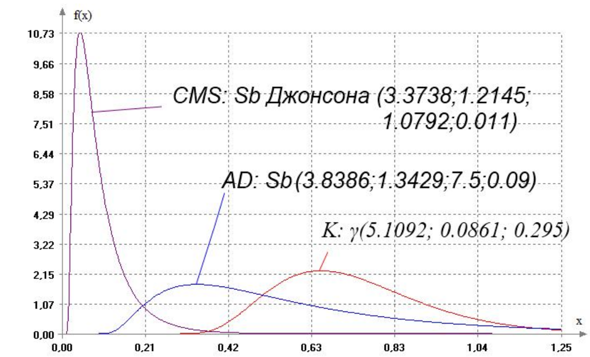


Рис. 1. Плотности распределений статистик *K*, *CMS* и *AD* при истинности проверяемой гипотезы.

2.6. Критерии, основанные на характеризации через функцию среднего остаточного времени безотказной работы

При $0 < \mu < \infty$ величина X распределена по показательному закону тогда и только тогда, когда $E(X - t \mid X > t) = \mu$ для всех $t > 0$. Это эквивалентно условию $E[\min(X,t)] = \mu F(t)\quad \forall t > 0$, и, основываясь на этом, Барингхаус и Хензе [11, 4] предложили статистики типа Колмогорова и Крамера-Мизеса-Смирнова.

Статистика Барингхауса-Хензе критерия типа Колмогорова имеет вид

$$\bar{K}_n = \sqrt{n}\sup_{t\ge 0}\left|\frac{1}{n}\sum_{j=1}^{n}\min(Y_j, t) - \frac{1}{n}\sum_{j=1}^{n}\mathbf{1}\{Y_j \le t\}\right| =$$
$$= \sqrt{n}\max(K_n^+, K_n^-),$$

где

$$K_n^+ = \max_{j=0,1,\dots,n-1}\left[n^{-1}\left(Y_{(1)}+\dots+Y_{(j)}\right) + Y_{(j+1)}(1 - j/n) - j/n\right],$$
$$K_n^- = \max_{j=0,1,\dots,n-1}\left[j/n - n^{-1}\left(Y_{(1)}+\dots+Y_{(j)}\right) - Y_{(j)}(1-j/n)\right].$$

В данном случае также целесообразно использовать статистику с поправкой Большева

$$K_n^* = \left(6n\cdot\bar{K}_n/\sqrt{n}+1\right)\Big/6\sqrt{n}.$$

Статистика Барингхауса-Хензе критерия типа Крамера-Мизеса-Смирнова определяется соотношением

$$CMS_n^* = n^{-1}\sum_{j,k=1}^{n}\left[2 - 3e^{-\min(Y_j,Y_k)} - 2\min(Y_j,Y_k)\times\right.$$
$$\left.\times\left(e^{-Y_j}+e^{-Y_k}\right) + 2e^{-\max(Y_j,Y_k)}\right].$$

Результаты моделирования показали, что распределения статистик не совпадают с указанными в [11], что не должно удивлять [12], так как в критериях проверяется сложная гипотеза с вычислением ОМП параметра масштаба (см. также раздел IV).

2.7. Критерий Дешпанде

Критерий предложен в [13] и рассмотрен в [3] для проверки показательности против конкурирующих распределений с возрастающей интенсивностью отказов. Статистика критерия вычисляется в соответствии с соотношением

$$J_b = n(n-1)^{-1}\sum h_b(X_j, X_k),$$

$$\text{где } h_b = \begin{cases}1, & X_j > bX_k,\\ 0, & \textit{иначе},\end{cases}$$

а суммирование ведется по всем $1 \le j,k \le n$, $j \ne k$. При отсутствии априорного знания о конкурирующем распределении следует использовать двусторонние критические значения.

Автором критерия показано, что величина $n^{1/2}(J_b - M(F))$ имеет нормальное распределение с $\mu = 0$ и $\sigma^2 = 4\zeta_1$, где $M(F) = (b+1)^{-1}$ и

$$\zeta_1 = \frac{1}{4}\left\{1 + \frac{b}{b+2} + \frac{1}{2b+1} + \frac{2(1-b)}{b+1} - \frac{2b}{b^2+b+1} - \frac{4}{(b+1)^2}\right\}.$$

2.8. Критерий Кокса-Оукса

Проверяемая гипотеза отклоняется при малых и больших значениях статистики критерия

$$CO_n = n + \sum_{j=1}^{n}(1-Y_j)\log Y_j.$$

Нормализованная статистика $CO_n\sqrt{6/n}\cdot\pi^{-1}$ в пределе подчиняется стандартному нормальному закону.

Критерий со статистикой CO_n состоятелен против конкурирующих законов с конечным математическим ожиданием и $\mathrm{E}[X\log X - \log X] \ne 1$ при условии, что последнее математическое ожидание существует.

2.9. Критерий Большева

Данный критерий предназначен для проверки гипотезы о принадлежности совокупности малых выборок показательным законам распределения [14].

Пусть $X_{i1},...,X_{in_i}\left(n_i \ge 2;\ i=\overline{1,N}\right)$ – взаимно независимые случайные величины. Требуется проверить гипотезу H_0: величины X_{ij} подчиняются показательным распределениям с плотностями распределения вероятностей $a_i\exp\left(-a_i x\right)$ $\left(x>0,\ j=\overline{1,n_i};\ i=\overline{1,N}\right)$; значения параметров a_i неизвестны и, возможно, различны. Если гипотеза H_0 верна, то статистики $\zeta_{ir}=\sum_{j=1}^{r}X_{ij}\Big/\sum_{j=1}^{r+1}X_{ij}$ $\left(r=\overline{1,n_i-1}\right)$ взаимно независимы и подчиняются бета-распределениям с параметрами r и 1 [14]. Статистики ζ_{ir}^{r}, $\left(r=\overline{1,n_i-1};\ i=\overline{1,N}\right)$ взаимно независимы и одинаково равномерно распределены на отрезке [0,1], к ним следует применять непараметрические критерии согласия для проверки равномерности. В данной работе использовался критерий Андерсона-Дарлинга [15]. На мощность критерия Большева определяющее влияние оказывает суммарный объем малых выборок, поэтому без ограничения общности рассматриваются одиночные выборки.

2.10. Критерий Клара

Критерий Клара [16, 4] основан на интегрированной функции распределения и отвергает гипотезу о показательности при больших значениях статистики

$$KL_{n,a}=\frac{2(3a+2)n}{(2+a)(1+a)^2}-2a^3\sum_{j=1}^{n}\frac{\exp\left(-(1+a)Y_j\right)}{(1+a)^2}-$$
$$-\frac{2}{n}\sum_{j=1}^{n}\exp\left(-aY_j\right)+\frac{2}{n}\sum_{i<j}\left[a\left(Y_{(k)}-Y_{(j)}\right)-2\right]\exp\left(-aY_{(j)}\right).$$

Автор критерия предлагает [16] использовать комбинированный критерий, основанный на нескольких статистиках $KL_{n,a}$ с различными значениями a и отвергающий проверяемую гипотезу, если она отвергается хотя бы одним из критериев $KL_{n,a}$ На основе результатов моделирования автор [16] приходит к выводу, что наибольшую мощность по отношению к альтернативам различных типов имеет критерий $KL_n^{1,10}$, основанный на комбинации $KL_{n,1}$ и $KL_{n,10}$

2.11. Критерии, основанные на эмпирическом преобразовании Лапласа

В данных критериях преобразование Лапласа показательного распределения $\psi(t)=E\left[\exp(-tX)\right]=\lambda/(t+\lambda)$ оценивается эмпирическим аналогом $\psi_n(t)=\frac{1}{n}\sum_{j=1}^{n}\exp\left(-tY_j\right)$.

2.11.1. Критерий Барингхауса-Хензе

В критерии Барингхауса-Хензе [17, 4] используется факт, что ψ удовлетворяет дифференциальному уравнению $(\lambda+t)\psi'(t)+\psi(t)=0$, $t\in R$. Критерий отклоняет гипотезу показательности для больших значений статистики

$$BH_{n,a}=n^{-1}\sum_{j,k=1}^{n}\left[\frac{(1-Y_j)(1-Y_k)}{Y_j+Y_k+a}-\frac{Y_j+Y_k}{\left(Y_j+Y_k+a\right)^2}+\right.$$
$$\left.+\frac{2Y_jY_k}{\left(Y_j+Y_k+a\right)^2}+\frac{2Y_jY_k}{\left(Y_j+Y_k+a\right)^3}\right].$$

Описание распределения статистики BH и таблицы процентных точек для некоторых a приводятся в [17]. Выбор параметра a предлагается делать в зависимости от предполагаемой конкурирующей гипотезы.

2.11.2. Критерий Хензе

Критерий Хензе [18, 4] отвергает гипотезу о показательности при больших значениях статистики

$$HE_{n,a}=n^{-1}\sum_{j,k=1}^{n}\frac{1}{Y_j+Y_k+a}-\sum_{j=1}^{n}\exp\left(Y_j+a\right)E_1\left(Y_j+a\right)+$$
$$+n\left(1-a\exp(a)E_1(a)\right),$$

где $E_1(z)=\int_z^{\infty}t^{-1}\exp(-t)\,dt$ – показательный интеграл, и $a>0$ – константа. Описание предельного распределения статистики HE и таблицы процентных точек для некоторых a автор приводит в [18].

2.11.3. L-критерий Хензе-Мейнтаниса

В L-критерии Хензе-Мейнтаниса [19, 4] используется статистика

$$L_{n,a}=\frac{1}{n}\sum_{j,k=1}^{n}\frac{1+\left(Y_j+Y_k+a+1\right)^2}{\left(Y_j+Y_k+a\right)^3}-2\sum_{j=1}^{n}\frac{1+Y_j+a}{\left(Y_j+a\right)^2}+\frac{n}{a}.$$

Проверяемая гипотеза о принадлежности выборки показательному закону распределения отклоняется при больших значениях статистики. Описание распределения статистики L и таблицы процентных точек для некоторых a авторы приводят в [19].

2.12. Критерии, основанные на эмпирической характеристической функции

2.12.1. W-критерии Хензе-Мейнтаниса

В W-критериях Хензе-Мейнтаниса [20, 4] проверяемая гипотеза отклоняется при больших значениях статистик

$$W_{n,a}^{(1)} = \frac{a}{2n}\sum_{j,k=1}^{n}\left[\frac{1}{a^2+\left(Y_j-Y_k\right)^2}-\frac{1}{a^2+\left(Y_j+Y_k\right)^2}\right.$$
$$\left.-\frac{4\left(Y_j+Y_k\right)}{\left(a^2+\left(Y_j+Y_k\right)^2\right)^2}+\frac{2a^2-6\left(Y_j-Y_k\right)^2}{\left(a^2+\left(Y_j-Y_k\right)^2\right)^3}+\frac{2a^2-6\left(Y_j+Y_k\right)^2}{\left(a^2+\left(Y_j+Y_k\right)^2\right)^3}\right],$$

$$W_{n,a}^{(2)} = \frac{\sqrt{\pi}}{4n\sqrt{a}}\sum_{j,k=1}^{n}\left[\left(1+\frac{2a-\left(Y_j-Y_k\right)^2}{4a^2}\right)\exp\left(-\frac{\left(Y_j-Y_k\right)^2}{4a}\right)+\right.$$
$$\left.+\left(\frac{2a-\left(Y_j+Y_k\right)^2}{4a^2}-\frac{Y_j+Y_k}{a}-1\right)\exp\left(-\frac{\left(Y_j+Y_k\right)^2}{4a}\right)\right].$$

Описание распределений статистик W и таблицы процентных точек для некоторых a авторы приводят в [20].

2.12.2. Критерий показательности Эппса-Палли

Статистика критерия показательности Эппса-Палли [4]

$$EP_n = \left(48n\right)^{1/2}\left[\frac{1}{2}\sum_{j=1}^{n}\exp\left(-Y_j\right)-1/2\right]$$

описывается стандартным нормальным законом при $n\to\infty$, проверяемая гипотеза отклоняется при больших значениях $\left|EP_n\right|$. Критерий является состоятельным против конкурирующих распределений с монотонной интенсивностью отказов, абсолютно-непрерывной функцией распределения F, $0<\mu<\infty$ и $F(0)=0$.

IV. Результаты экспериментов

Часть авторов приводит различные нормализующие преобразования статистик критериев, что позволяет для нормализованной статистики при проверке гипотезы использовать стандартный нормальный закон для вычисления достигаемых уровней значимости. На практике такие асимптотические результаты могут оказаться неприемлемыми для выборок конечного объема вследствие существенного отличия распределения конкретной статистики от асимптотического. Это в свою очередь способно привести к некорректным статистическим выводам при имеющемся объеме выборки.

Для проверки того, насколько хорошо действительные распределения статистик согласуются с теоретическими распределениями, была использована методика статистического моделирования [21]. При построении эмпирических распределений статистик критериев, соответствующих справедливой проверяемой гипотезе, к статистикам T, G, J_b и CO применялись предложенные авторами нормализующие преобразования. Количество испытаний в каждом случае составило 16600, в качестве истинного выступало показательное распределение $F(x)=1-\exp(-x)$. Полученные выборки статистик критериев показательности проверялись на согласие с соответствующими предельными законами по критерию Колмогорова. Достигнутые уровни значимости при проверке простой гипотезы о согласии эмпирического распределения статистики с соответствующим теоретическим законом распределения приведены в Табл. I.

При использовании критериев Q_R, Q_R', G, K, CMS, AD, B применение предельных распределений статистик является корректным и позволяет при проверке гипотезы точно оценить достигнутый уровень значимости.

Критерии K^* и CMS^* не избавлены полностью от влияния объема выборки на распределение статистики. При $n\geq 20$ в качестве модели распределения для K^* можно применять распределение Sb-Джонсона Sb(2.1275;1.6849;2.5437;0.26888), а в качестве модели распределения для CMS^* – Sb(2.756;0.98223;1.8645;0.01602). При $n=10$ использование данных моделей приводит к занижению достигаемого уровня значимости по критерию K^* и к завышению – по критерию CMS^*.

Таблица I

Достигнутые уровни значимости при проверке согласия эмпирических распределений статистик с соответствующими теоретическими

n	$Q_{[0.3n]}$	$Q'_{[0.1n]}$	HP	G	K	K^*	CMS	CMS^*	AD	$J_{0.5}$	EP	CO	B
10	0.04	0.60	0.00	0.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13
20	0.27	0.23	0.00	0.58	0.06	0.00	0.13	0.00	0.03	0.00	0.00	0.00	0.31
50	0.74	0.91	0.00	0.95	0.92	0.05	0.25	0.00	0.14	0.00	0.00	0.00	0.52
100	0.40	0.11	0.00	0.41	0.94	0.04	0.39	0.00	0.33	0.00	0.00	0.00	0.04
200	1.00	0.05	0.00	0.95	0.43	0.16	0.80	0.00	0.16	0.00	0.00	0.00	0.17
300	0.19	0.16	0.00	0.82	0.81	0.06	0.93	0.00	0.80	0.00	0.00	0.00	0.97
400	0.94	0.29	0.00	0.40	0.86	0.09	0.17	0.00	0.13	0.00	0.00	0.00	0.78
500	0.54	0.84	0.00	0.80	0.21	0.30	0.94	0.00	0.18	0.01	0.00	0.00	0.53

Нормальную аппроксимацию для распределения статистики *HP* критерия Холландера-Прошана можно использовать лишь с оговорками. При $n \le 300$ лучшим выбором было бы построение таблиц процентных точек для статистики критерия. При $n \ge 400$ использование нормальной модели уже оправданно.

В критерии со статистикой $J_{0.5}$ применение нормальной аппроксимации не приводит к существенным ошибкам при $n \ge 50$, а в критериях со статистиками *EP* и *CO* – при $n \ge 100$.

V. Обсуждение результатов

Мощность критериев сравнивалась на небольших объемах выборок n=20 и n=50. Эмпирические распределения статистик критериев, по которым находились оценки мощности, соответствующие проверяемой и конкурирующим гипотезам, для получения приемлемой точности строились по 1 660 000 испытаниям. В качестве проверяемой гипотезы H_0 рассматривался показательный закон $F(x) = 1 - \exp(-x)$. Показательному закону соответствует постоянная интенсивность отказов, поэтому в качестве конкурирующих гипотез рассматривались распределения, принадлежащие к трем классам: с возрастающими, убывающими и немонотонными интенсивностями отказов:

– Вейбулла $W(\theta)$ с плотностью $f(x) = \theta x^{\theta-1} \exp\left(-x^{\theta}\right)$;
– гамма $\Gamma(\theta)$ – $f(x) = \Gamma(\theta)^{-1} x^{\theta-1} \exp(-x)$;
– бета $\mathrm{B}(\theta_0, \theta_1)$ – $f(x) = \mathrm{B}(\theta_0, \theta_1)^{-1} x^{\theta_0 - 1} (1-x)^{\theta_1 - 1}$;
– равномерное $U(0,1)$ на интеравале $[0,1]$;
– логнормальное $LN(\theta)$ – $f(x) = \left(\theta x \sqrt{2\pi}\right)^{-1} \exp\left(-(\ln x)^2 / 2\theta^2\right)$;
– полунормальное HN – $f(x) = (2/\pi)^{1/2} \exp\left(-x^2/2\right)$.

К распределениям с возрастающими интенсивностями отказов относятся $W(\theta)$ и $\Gamma(\theta)$ $(\theta > 1)$, $U(0,1)$, HN, $\mathrm{B}(1,2)$, $\mathrm{B}(2,1)$; с убывающими – $W(\theta)$ и $\Gamma(\theta)$ $(\theta < 1)$; с немонотонными – LN, $\mathrm{B}(0.5,1)$.

При вычислении критических значений статистик и оценок мощности мы исходили из предположения, что при проверке гипотезы не известен тип конкурирующей гипотезы. Поэтому в критериях, допускающих выбор между левосторонней и правосторонней критическими областями, мы использовали двусторонние критические области.

Полученные оценки мощности критериев по отношению к конкурирующим законам с возрастающими, убывающими и немонотонными интенсивностями отказов приведены в Табл. II и III.

Критерии со статистиками BH и HE ведут себя похожим образом, что отмечено и в [4], поэтому далее будет упоминаться только критерий со статистикой BH. Выбор параметра $a = 0.5$ (и соответственно статистики $BH_{0.5}$) обеспечивает более высокую мощность по сравнению с другими значениями a, а также с другими критериями. В L-критерии в общем случае имеет смысл выбирать статистику L_1, в W-критериях – статистику $W_1^{(1)}$, в KL-критерии, очевидно, – статистику $KL^{1,10}$.

В случае конкурирующих законов с возрастающей интенсивностью отказов (см. Табл. II) необходимо отметить следующие недостатки исследуемых критериев. При n=20 критерий Большева смещен относительно конкурирующих гипотез (конкурирующих законов) $W(1.2)$, $\Gamma(1.5)$, полунормального HN и $\mathrm{B}(1,2)$ (т.е. мощность критерия оказывается меньше заданного уровня значимости $\alpha = 0.05$); критерий $L_{0.1}$ – относительно тех же законов и $W(1.4)$; критерий $W_{2.5}^{(2)}$ оказывается смещенным относительно $W(1.2)$.

Относительно конкурирующих законов с убывающей интенсивностью отказов примечательно низкую мощность показывает критерий со статистикой $Q'_{0.1}$ (см. Табл. III).

В случае конкурирующих законов с немонотонными интенсивностями отказов критерии со статистиками $W_{2.5}^{(2)}$ и BH_5 смещены относительно альтернативы $\mathrm{B}(0.5,1)$, критерий $L_{0.1}$ – относительно $LN(1)$ и $LN(0.8)$ (см. Табл. III).

VI. Выводы и заключение

Очевидно, среди исследованных критериев нельзя однозначно выбрать критерий, обладающий наибольшей мощностью относительно всех рассмотренных альтернатив (конкурирующих законов). Расположить критерии в некотором безусловном порядке, например, по убыванию мощности также не представляется возможным.

В то же время, можно выделить группы критериев, одинаково перспективных для применения в случае выдвижения альтернатив определенного вида.

Так, по отношению к конкурирующим законам с возрастающими и убывающими интенсивностями отказов стабильно высокую мощность показывают критерии Кокса-Оукса (CO), Андерсона-Дарлинга (AD), Хензе-Мейнтаниса (L_1 и $W_1^{(1)}$), Барингхауса-Хензе ($BH_{0.5}$), Хензе ($HE_{0.5}$).

По отношению к альтернативам с немонотонными интенсивностями отказов высокой мощностью обладают критерии Харриса ($Q'_{0.1}$) и Андерсона-Дарлинга (AD).

При малых объемах выборок или без указания конкретной альтернативы нежелательно использование критериев Харриса ($Q'_{0.1}$), Большева (B), Хензе-Мейнтаниса ($L_{0.1}$ и $W_{2.5}^{(2)}$), Барингхауса-Хензе ($BH_{0.5}$), Хензе ($HE_{0.5}$) (вследствие возможной смещенности критериев).

Чтобы выбрать наиболее мощный критерий показательности при наличии заданной альтернативы, выходящей за рамки рассмотренных в данной работе конкурирующих гипотез, необходимо провести исследование мощности критериев по аналогичной методике. При этом учесть знания об интенсивности отказов, характеризующей данную альтернативу.

Критерий показательности Большева (B) имеет достаточно высокую мощность против конкурирующих законов с убывающими интенсивностями отказов, а в случае иных альтернатив он уступает другим рассмотренным критериям. При этом следует иметь в виду, что основным достоинством критерия Большева является подход, позволяющий проверять гипотезу о показательности по совокупности малых выборок.

Список литературы

[1] Лемешко Б.Ю., Лемешко С.Б. Сравнительный анализ критериев проверки отклонения распределения от нормального закона // Метрология. 2005. № 2. С. 3-24.

[2] Лемешко Б.Ю., Рогожников А.П. Исследование особенностей и мощности некоторых критериев нормальности // Метрология. 2009. № 4. С. 3-24.

[3] Ascher S. A survey of tests for exponentiality // Communications in Statistics - Theory and Methods. 1990. Vol. 19. No. 5. pp. 1811-1825.

[4] Henze N., Meintanis S.G. Recent and classical tests for exponentiality: a partial review with comparisons // Metrika. 2005. Vol. 61. pp. 29-45.

[5] Lin C.C., Mudholkar G.S. A test of exponentiality based on the bivariate F distribution // Technometrics. Feb 1980. Vol. 22. No. 1. pp. 79-82.

[6] Harris C.M. A note on testing for exponentiality // Naval Research Logistics Quarterly. Mar 1976. Vol. 23. No. 1. pp. 169-175.

[7] Hollander M., Proschan F. Testing whether new is better than used // The Annals of Mathematical Statistics. 1972. Vol. 43. No. 4. pp. 1136-1146.

[8] Gail M.H., Gastwirth J.L. A scale-free goodness-of-fit test for the exponential distribution based on the Gini statistic // Journal of the Royal Statistical Society. Series B (Methodological). 1978. Vol. 40. No. 3. pp. 350-357.

[9] Большев Л.Н. Асимптотические пирсоновские преобразования // Теория вероятностей и ее применения. 1963. Т. 8. № 2. С. 129-155.

[10] Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н., Чимитова Е.В. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход : монография. Новосибирск: Изд-во НГТУ, 2011. 888 с.

[11] Baringhaus L., Henze N. Tests of fit for exponentiality based on a characterization via the mean residual life function // Statistical Papers. 2000. No. 41. pp. 225-236.

[12] Lemeshko B.Y., Lemeshko S.B. Construction of statistic distribution models for nonparametric goodness-of-fit tests in testing composite hypotheses: the computer approach // Quality Technology & Quantitative Management. 2011. Vol. 8. No. 4. pp. 359-373.

[13] Deshpande V.J. A class of tests for exponentiality against increasing failure rate average alternatives // Biometrika. 1983. Vol. 70. No. 2. pp. 514-518.

[14] Большев Л.Н. К вопросу о проверке «показательности» // Теория вероятностей и ее применения. 1966. Т. 11. № 3. С. 542-544.

[15] Лемешко Б.Ю., Рогожников А.П. Исследование мощности критерия показательности Большева // Сборник научных трудов НГТУ. 2012. № 1(67).

[16] Klar B. Goodness-of-fit tests for the exponential and the normal distribution based on the integrated distribution function // Ann. Inst. Statist. Math. 2001. Vol. 53. No. 2. pp. 338-353.

[17] Baringhaus L., Henze N. A class of consistent tests for exponentiality based on the empirical Laplace transform // Ann. Inst. Statist. Math. 1991. Vol. 43. No. 3. pp. 551-564.

[18] Henze N. A new flexible class of omnibus tests for exponentiality // Commun. Statist. - Theory Meth. 1993. Vol. 22. No. 1. pp. 115-133.

[19] Henze N., Meintanis S.G. Tests of fit for exponentiality based on the empirifcal Laplace transform // Statistics. 2002. Vol. 36. No. 2. pp. 147-161.

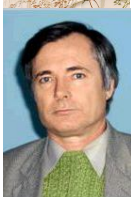
[20] Henze N., Meintanis S.G. Goodness-of-fit tests based on a new characterization of the exponential distribution // Comm. Statist. Theory Meth. 2002. Vol. 31. No. 9. pp. 1479-1497.

[21] Лемешко Б.Ю., Постовалов С.Н. Компьютерные технологии анализа данных и исследования статистических закономерностей: Учеб. пособие. Новосибирск: Изд-во НГТУ, 2004. 120 с.

[22] Fercho W.W., Ringer L.J. Small sample power of some tests of the constant failure rate // Technometrics. 1972. No. 14. pp. 713-724.

Андрей Павлович Рогожников

Магистр прикладной математики и информатики (2009), аспирант кафедры прикладной математики НГТУ



Борис Юрьевич Лемешко

Профессор кафедры прикладной математики НГТУ, д.т.н., декан Факультета прикладной математики и информатики НГТУ

Таблица II

Мощность критериев показательности относительно конкурирующих законов с возрастающей интенсивностью отказов $\times 1000$ (n=20, α=0.05).

	$W(1.2)$	Г(1.5)	HN	B(1,2)	$W(1.4)$	Г(2)	$W(1.5)$	$U(0.1)$	Г(4)	B(2,1)
CO	138	217	191	220	381	551	527	528	996	999
$J_{0.5}$	123	186	177	200	322	466	448	545	984	998
EP	133	194	216	270	366	490	511	672	989	1000
G	130	187	216	277	356	473	498	714	987	1000
$Q_{0.3}$	107	148	163	180	265	358	367	445	926	985
$Q'_{0.1}$	54	78	60	73	93	176	126	120	692	607
HP	124	184	191	234	333	465	466	669	985	1000
K	119	169	178	204	290	407	398	528	961	994
K^*	152	204	244	303	358	462	480	729	975	1000
CMS	135	197	210	252	350	483	482	673	988	1000
CMS^*	134	191	221	279	358	477	496	716	988	1000
AD	109	168	170	209	307	451	438	628	987	1000
B	39	43	45	49	58	77	77	126	457	823
$KL^{1,10}$	102	160	159	203	294	439	423	619	986	1000
$L_{0.1}$	10	12	20	21	27	61	50	48	661	673
L_1	140	220	192	214	380	554	523	526	996	999
$W_1^{(1)}$	132	194	189	221	328	465	450	661	982	1000
$W_{2.5}^{(1)}$	123	161	229	327	320	389	447	809	954	1000
$W_1^{(2)}$	118	150	228	342	303	356	424	829	923	1000
$W_{2.5}^{(2)}$	46	59	99	174	134	160	206	696	710	1000
$BH_{0.5}$	134	213	184	206	368	542	510	523	996	999
BH_1	140	213	205	240	381	537	526	597	995	1000
$BH_{1.5}$	138	206	211	253	378	520	523	631	993	1000
$BH_{2.5}$	131	192	210	263	363	489	507	662	989	1000
$HE_{0.5}$	139	219	192	214	379	552	522	529	996	999
HE_1	142	215	209	244	385	538	531	601	995	1000
$HE_{1.5}$	139	207	213	257	380	520	525	634	993	1000
$HE_{2.5}$	131	192	211	264	363	489	507	664	989	1000

Таблица III

Мощность критериев показательности относительно конкурирующих законов с убывающей (слева) и немонотонной (справа) интенсивностью отказов $\times 1000$ (n=20, α=0.05)

	Γ(0.7)	*W*(0.8)	Γ(0.5)	Γ(0.4)
CO	281	277	730	913
$J_{0.5}$	196	184	564	786
EP	200	236	543	759
G	203	239	547	759
$Q_{0.3}$	206	193	567	787
$Q'_{0.1}$	85	84	131	165
HP	226	216	601	811
K	156	173	470	706
K^*	112	134	380	617
CMS	178	199	525	756
CMS^*	185	218	523	748
AD	273	269	706	898
B	175	172	511	759
$KL^{1,10}$	272	279	686	879
$L_{0.1}$	363	309	785	933
L_1	240	252	645	851
$W_1^{(1)}$	155	162	464	693
$W_{2.5}^{(1)}$	155	191	429	638
$W_1^{(2)}$	148	184	404	606
$W_{2.5}^{(2)}$	182	235	428	611
$BH_{0.5}$	251	259	664	866
BH_1	225	248	614	827
$BH_{1.5}$	213	242	583	800
$BH_{2.5}$	202	238	548	765
$HE_{0.5}$	243	255	650	855
HE_1	220	245	602	816
$HE_{1.5}$	210	240	575	791
$HE_{2.5}$	200	237	543	759

	LN(1)	B(0.5,1)	*LN*(0.8)	*LN*(1.5)	*LN*(0.6)
CO	106	261	348	595	890
$J_{0.5}$	73	144	356	247	900
EP	132	66	259	663	801
G	117	60	246	659	801
$Q_{0.3}$	41	254	197	421	707
$Q'_{0.1}$	215	460	312	326	657
HP	61	137	307	314	863
K	138	154	304	572	851
K^*	122	117	287	549	841
CMS	152	188	341	616	891
CMS^*	151	112	279	652	841
AD	139	397	334	625	893
B	73	208	65	398	234
$KL^{1,10}$	149	290	347	652	914
$L_{0.1}$	5	530	42	389	430
L_1	109	171	375	619	923
$W_1^{(1)}$	105	226	350	513	888
$W_{2.5}^{(1)}$	133	100	173	631	640
$W_1^{(2)}$	128	94	147	618	549
$W_{2.5}^{(2)}$	193	41	91	680	291
$BH_{0.5}$	117	221	379	623	929
BH_1	125	137	334	646	892
$BH_{1.5}$	132	100	303	656	859
$BH_{2.5}$	142	69	266	666	809
$HE_{0.5}$	110	175	372	623	921
HE_1	121	113	326	646	883
$HE_{1.5}$	129	87	298	656	851
$HE_{2.5}$	140	64	263	665	805