УДК 519.248 EDN: KDBJJP

АНАЛИЗ МОЩНОСТИ ДВУХ КРИТЕРИЕВ ТИПА КОЛМОГОРОВА — СМИРНОВА ПРОВЕРКИ СТЕПЕННОЙ МОДЕЛИ КОКСА ДЛЯ ПРОГРЕССИВНО ЦЕНЗУРИРОВАННЫХ ВЫБОРОК

В.И. Тимонин timonin@bmstu.ru
Н.Д. Тянникова tiannikova@bmstu.ru

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Аннотация

При испытаниях технических систем часто возникает задача сравнения показателей надежности их элементов в различных режимах (эксплуатация в различных климатических поясах, сравнение результатов лабораторных испытаний и реальных данных эксплуатации и др.). При этом законы распределения наработок до отказа элементов неизвестны. Дополнительные трудности в решении этой задачи вызывает характер имеющихся данных — в последовательных системах отказ одного элемента приводит к тому, что наработки до отказа оставшихся годными элементов остаются неизвестными (цензурируются). В связи с этим актуальным является решение такой задачи непараметрическими (свободными от знания распределения) методами. Ранее авторами был разработан критерий типа Колмогорова — Смирнова, позволяющий проверять степенную зависимость функций надежности элементов последовательных систем (непараметрическая степенная модель Кокса). В статистике этого критерия использовались оценки Каплана — Мейера функций надежности элементов, построенные по наработкам составленных из них систем. Однако особенностью оценок Каплана — Мейера является медленная сходимость к теоретической функции надежности на правом хвосте распределения, что может приводить к уменьшению мощности (снижению чувствительности) этого критерия. Здесь для решения аналогичной задачи предложен другой критерий типа Колмогорова — Смирнова, в статистике которого не используются оценки Каплана — Мейера и который основан на сравнении оценок

Ключевые слова

Критерий Колмогорова — Смирнова, модель Кокса, оценка Каплана — Мейера, мощность критерия, последовательные системы, цензурирование

функций надежности систем. Получены точное и асимптотическое распределения его статистики. Проведено подробное исследование мощности двух критериев методами численного анализа и статистического моделирования. Методами Монте-Карло исследована точность оценок параметра Кокса, полученных минимизацией статистик сравниваемых критериев

Поступила 14.04.2023
Принята 07.06.2023

Введение. Задача проверки гипотезы о пропорциональности интенсивности отказов составляющих систем элементов для различных режимов [1, 2] часто возникает при испытаниях сложных технических систем. Эта задача эквивалентна проверке гипотезы

$$H_0 : P_1(t) = \left(P_2(t)\right)^k, \tag{1}$$

где $P_1(t)$, $P_2(t)$ — функции надежности элементов систем, работающих в режимах ε_1 и ε_2. Степенная модель $P_1(t) = \left(P_2(t)\right)^k$ [3], используемая в гипотезе (1), впервые была предложена Д. Коксом. Она эквивалентна пропорциональности интенсивностей отказов элементов систем, работающих в режимах ε_1 и ε_2: $\lambda_1(t) = k\lambda_2(t)$, где $k \geq 1$ — некоторый заданный параметр [4, 5].

Использовать критерий типа Колмогорова — Смирнова [6] предложено в [7] для проверки (1) по двум прогрессивно цензурированным выборкам, составленным из отказов элементов систем. Необходимым условием проверки гипотезы является применение оценок функций надежности элементов Каплана — Мейера, построенным по наработкам до отказа систем, составленных из элементов [8, 9]. В настоящей работе для проверки (1) предложено рассматривать системы как «черный ящик», что позволяет не использовать оценки функций надежности элементов. Разработан критерий проверки (1), основанный на использовании только эмпирических функций надежности систем. Получены точные и асимптотические распределения статистики этого критерия и проведено сравнение мощностей этих критериев.

Постановка задачи. Пусть имеются две цензурированные выборки [10–12]: первая выборка состоит из n_1 наработок систем, структурно состоящих из m_1 последовательных элементов, вторая — из n_2 наработок систем из m_2 аналогичных, последовательно соединенных элементов. Под наработкой системы понимается минимальная наработка до отказа одного из элементов, из которых состоит система.

Для оценки функций надежности $P_1(t)$, $P_2(t)$ по прогрессивно цензурированным выборкам предложено использовать оценки Каплана — Мейера:

$$\hat{P}_{\theta_1}(t)=\begin{cases} 1, & d_1(t)=0, \\ \prod\limits_{i=1}^{d_1(t)}\left(1-\dfrac{1}{m_1(n_1-i+1)}\right), & 1\le d_1(t)\le(n_1-1); \\ 0, & d_1(t)=n_1, \end{cases}$$

$$\hat{P}_{\theta_2}(t)=\begin{cases} 1, & d_2(t)=0, \\ \prod\limits_{j=1}^{d_2(t)}\left(1-\dfrac{1}{m_2(n_2-j+1)}\right), & 1\le d_2(t)\le(n_2-1). \\ 0, & d_2(t)=n_2, \end{cases}$$

Здесь $d_1(t)$, $d_2(t)$ — число элементов выборок Θ_1 и Θ_2, меньших t.

Статистика критерия, введенного в [6], позволяющего проверить гипотезу (1) в виде, более удобном для применения:

$$T=$$

$$=\frac{m_1m_2\sqrt{n_1n_2}}{\sqrt{k^2n_1m_1^2+n_2m_2^2}}\max_t\frac{\left(V(t)\right)^{m_2/k+m_1-1}}{k_2\left(V(t)\right)^{m_2/k}+k_1\left(V(t)\right)^{m_1}}\left|\hat{P}_{\theta_1}(t)-\left(\hat{P}_{\theta_2}(t)\right)^k\right|, \tag{2}$$

где

$$V(t)=k_2\left(1-\hat{F}_1\right)^{1/m_1}+k_1\left(1-\hat{F}_2\right)^{k/m_2},$$

$$k_2=\frac{m_2^2}{\rho m_1^2k^2+m_2^2},\quad k_1=\frac{\rho m_1^2k^2}{\rho m_1^2k^2+m_2^2},\quad \rho=\frac{n_1}{n_2}.$$

Если рассмотреть выборки из отказов систем Θ_1 и Θ_2 как полные независимые, то их функции распределения можно оценить обычными эмпирическими функциями распределения $\hat{F}_1$, $\hat{F}_2$, равными отношению числа отказов систем $d_1(t)$, $d_2(t)$ к числу систем n_1, n_2.

Рассмотрим системы как «черные ящики», т. е. по результатам испытаний наблюдаются только наработки до отказа систем $\left(\theta_1^1,\ldots,\theta_1^{n_1}\right)$, $\left(\theta_2^1,\ldots,\theta_2^{n_2}\right)$ в режимах ε_1 и ε_2. В таком случае интенсивности отказов систем $\lambda_1^s(t)$ связаны с интенсивностями отказов элементов $\lambda_1(t)$ соотношением $\lambda_1^s(t) = m_1\lambda_1(t)$. Следовательно, функция надежности систем $P_1^s(t)$ и функция надежности составляющих ее элементов $P_1(t)$ связаны соотношением $P_1^s(t) = (P_1(t))^{m_1}$, где m_1 — кратность систем. Аналогичные соотношения можно записать для режима ε_2: $P_2^s(t) = (P_2(t))^{m_2}$. Тогда задача проверки гипотезы (1) сводится к задаче проверки гипотезы

$$H_0 : P_1^s(t) = \left(P_2^s(t)\right)^q, \quad q = \frac{km_1}{m_2}. \tag{3}$$

Индекс s в (3) указывает на то, что эта функция есть функция надежности систем, а не элементов, как было в (1).

Для проверки степенной гипотезы (3) предлагается использовать критерий типа Колмогорова — Смирнова, в котором вместо оценок функций надежности наработок до отказа элементов Каплана — Мейера используют оценки функций надежности наработок до отказа систем $\widehat{P_1^s}(t)$, $\widehat{P_2^s}(t)$, построенные по их полным выборкам [1]. Статистика критерия:

$$T^s = \frac{\sqrt{n_1 n_2}}{\sqrt{n_1 q^2 + n_2}} \max_t \frac{\left(U(t)\right)^{1/q}}{a_2\left(U(t)\right)^{1/q} + a_1\left(U(t)\right)} \left| \widehat{P_1^s}(t) - \left(\widehat{P_2^s}(t)\right)^q \right|, \tag{4}$$

где $U(t) = a_2 \widehat{P_1^s}(t) + a_1 \left(\widehat{P_1^s}(t)\right)^q$, $a_2 = \dfrac{1}{\rho q^2 + 1}$, $a_1 = \dfrac{\rho q^2}{\rho q^2 + 1}$.

Отличие статистик (2) и (4) заключается в замене оценок Каплана — Мейера $\hat{P}_{\theta_1}(t)$, $\hat{P}_{\theta_2}(t)$ функций надежности элементов оценками функции надежности систем $\widehat{P_1^s}(t)$, $\widehat{P_2^s}(t)$, что влечет за собой замену степени k степенью $q = km_1 / m_2$.

Точные распределения. Точные распределения статистики (4) T^s могут быть вычислены с использованием модели случайного блуждания [6, 13]. Введем массив ячеек случайного блуждания частицы $A = \{a_{ij}\}$, $i = \overline{0, n_1}$, $j = \overline{0, n_2}$. Частица выходит из ячейки (0, 0) и заканчивает блуждание в ячейке (n_1, n_2), перемещаясь вниз или вправо. Введем вектор $\vec{Z}$, составленный из n_1 единиц и n_2 нулей. Для этого примем $z_i = 1$, если происходит отказ системы, работающей в режиме ε_1, и $z_i = 0$, если происходит отказ системы, работающей в режиме ε_2.

Все допустимые траектории частицы случайного блуждания обладают взаимно однозначным соответствием со всеми допустимыми векторами $\vec{Z}$.

Обозначим

$$Q(t)=\frac{\sqrt{n_1n_2}}{\sqrt{n_1q^2+n_2}}\frac{\left(U(t)\right)^{1/q}}{a_2\left(U(t)\right)^{1/q}+a_1\left(U(t)\right)}\left|\widehat{P_1^s}(t)-\left(\widehat{P_2^s}(t)\right)^q\right|,$$

т. е. $T^s=\max Q(t).$ При прохождении блуждания через ячейку a_{ij} функция $Q(t),$ соответствующая статистике (4), принимает значение $t_{ij}{}^s,$ равное

$$t_{ij}{}^s=\frac{\sqrt{n_1n_2}}{\sqrt{n_1q^2+n_2}}\times$$

$$\times\frac{\left(a_2\left(1-i/n_1\right)+a_1\left(1-j/n_2\right)^q\right)^{1/q}}{a_2\left(a_2\left(1-i/n_1\right)+a_1\left(1-j/n_2\right)^q\right)^{1/q}+a_1\left(a_2\left(1-i/n_1\right)+a_1\left(1-j/n_2\right)^q\right)}\times$$

$$\times\left|\left(1-\left(1-i/n_1\right)\right)-\left(1-j/n_2\right)^q\right|.$$

Общий метод вычисления точных распределений статистик типа Колмогорова — Смирнова, построенный на основе представленной модели случайного блуждания, предложен в [6]. На его основе можно получить утверждение следующей леммы.

Лемма 1. *Вероятность* $P\left\{T^s<h\right\}$ *равна величине* $\pi_{n_1,n_2}\left(h\right),$ *которая получается применением рекуррентного соотношения*

$$\pi_{ij}\left(h\right)=\begin{cases}1, \text{ если } i=0, j=0;\\ \left(\dfrac{n_2-j+1}{q\left(n_1-i\right)+n_2-j+1}\pi_{i,j-1}\left(h\right)\right)\chi_{ij}\left(h\right), \quad i=0,\ 1\le j\le n_2;\\ \left(\dfrac{q\left(n_1-i+1\right)}{q\left(n_1-i+1\right)+n_2-j}\pi_{i-1,j}\left(h\right)\right)\chi_{ij}\left(h\right), \quad 1\le i\le n_1,\ j=0;\\ \left(\begin{gathered}\dfrac{q\left(n_1-i+1\right)}{q\left(n_1-i+1\right)+n_2-j}\pi_{i-1,j}\left(h\right)+\\ +\dfrac{n_2-j+1}{q\left(n_1-i\right)+n_2-j+1}\pi_{i,j-1}\left(h\right)\end{gathered}\right)\chi_{ij}\left(h\right),\ 1\le i\le n_1,\ 1\le j\le n_2.\end{cases}\qquad(5)$$

Здесь $\chi_{ij}(h)=\begin{cases}1, & t_{ij}{}^{s}<h,\\ 0, & t_{ij}{}^{s}\geq h.\end{cases}$

Точные вероятности, которые позволяют сравнить скорость сходимости точных распределений T и T^s к асимптотическим [14, 15] для $m_1=2,\ m_2=3,\ k=3,$ приведены в табл. 1. По этой причине выбраны аргументы 1,22, 1,36, которые являются квантилями предельного распределения Колмогорова — Смирнова уровней 0,9 и 0,95 соответственно [15].

Таблица 1

Точные распределения статистик T и T^s

$n_1=n_2$	$P(T^s<1{,}22)$	$P(T<1{,}22)$	$P(T^s<1{,}36)$	$P(T<1{,}36)$
10	0,9468	0,9015	0,9803	0,9052
50	0,9196	0,9047	0,9620	0,9517
100	0,9141	0,9060	0,9579	0,9543
300	0,9074	0,9057	0,9558	0,9545
500	0,9057	0,9042	0,9546	0,9538
700	0,9039	0,9036	0,9536	0,9533
900	0,9034	0,9030	0,9535	0,9531
1100	0,9029	0,9025	0,9531	0,9529
1300	0,9025	0,9022	0,9529	0,9527
1500	0,9022	0,9020	0,9528	0,9526

Представляет интерес определение мощности предложенных критериев для альтернативных гипотез [16]. Очевидно, что мощности можно точно рассчитать в случае альтернатив вида $H_1: P_1(t)=\left(P_2(t)\right)^{k+\delta}$, $k\geq 1,\ \ \delta>0,$ где δ — параметр, определяющий разность проверяемого значения k и альтернативного значения $k_1=k+\delta.$ В приведенных ниже расчетах использовано $\delta=0{,}5.$

В этом случае точные вероятности для статистик (2) и (4) рассчитываются по формулам (5), но с заменой k на $k+\delta$ за исключением определения

$$\chi_{ij}(h)=\begin{cases}1, & t_{ij}{}^{s}<h,\\ 0, & t_{ij}{}^{s}\geq h.\end{cases}$$

Критерий будет тем мощнее, чем меньше будут получаемые вероятности. Результаты расчета вероятностей ошибок второго рода при $m_1=2,$ $m_2=3,\ \delta=0{,}5,\ h=1{,}22$ для равных объемов выборок приведены в табл. 2.

Сравнение мощностей рассматриваемых критериев типа Колмогорова — Смирнова при $m_1 = 2$, $m_2 = 3$, $k = 2$, $\delta = 0{,}5$, $h = 1{,}22$ для равных объемов выборок представлено на рис. 1.

Таблица 2

Вероятности ошибок второго рода для критериев типа Колмогорова — Смирнова

$n_1 = n_2$	$k = 2$		$k = 2{,}5$		$k = 3$	
	$P(T^s < h)$	$P(T < h)$	$P(T^s < h)$	$P(T < h)$	$P(T^s < h)$	$P(T < h)$
10	0,939369	0,860857	0,95907	0,792251	0,964026	0,851601
20	0,891687	0,859161	0,929755	0,819301	0,937275	0,830096
30	0,87261	0,803194	0,893098	0,798719	0,91967	0,793967
40	0,824589	0,794633	0,882784	0,795422	0,907695	0,81493
50	0,817511	0,766495	0,866698	0,789251	0,899371	0,798283
500	0,127392	0,115071	0,285659	0,246961	0,435448	0,373325
1000	0,008636	0,007412	0,057954	0,047911	0,157294	0,129244

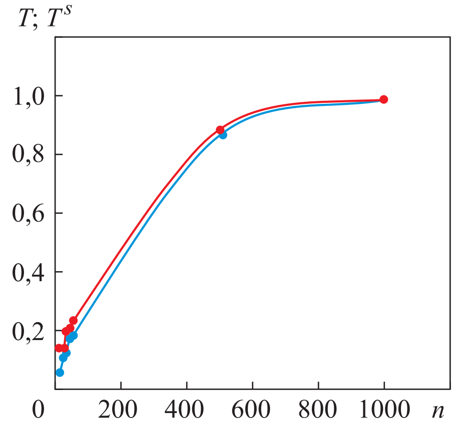

Рис. 1. Сравнение мощностей критериев типа Колмогорова — Смирнова T^s (—•—) и T (—•—) при $k = 2$

Асимптотическое распределение. Вид статистик T и T^s обусловлен тем, что их асимптотические распределения совпадают с классическим распределением Колмогорова — Смирнова.

Примем $P_1^s(t) = 1 - t$, $P_2^s(t) = (1-t)^{1/q}$, $0 \le t \le 1$. Введем процесс $X_1(t) = \sqrt{n_1}\left(\widehat{P_1^s}(t) - (1-t)\right)$. Доказано, что распределение процесса $X_1(t)$ слабо сходится к распределению броуновского моста $W(t)$ [17].

Аналогично введем вспомогательный случайный процесс $\tilde{X}_2(t) = \sqrt{n_2}\left(\widehat{P_2^s}(t) - (1-t)^{1/q}\right)$, $0 \le t \le 1$.

Лемма 2. *Пусть* $0 \le s \le t \le 1$, *тогда в асимптотике функция ковариации процесса* $\tilde{X}_2(t)$ *имеет вид*

$$\tilde{K}_2(s,t) = E\tilde{X}_2(s)\tilde{X}_2(t) - E\tilde{X}_2(s)E\tilde{X}_2(t) \underset{n_2 \to \infty}{\to} 1 - (1-s)^{1/q}(1-t)^{1/q}.$$

Для получения асимптотической функции ковариации процесса $X_2(t)=\sqrt{n_2}\left(\widehat{P_2^s}(t)^q-(1-t)\right),\ 0\le t\le 1,$ используем утверждение, ранее доказанное авторами [6].

Утверждение. *Пусть последовательность*

$$h_n(x)=\sqrt{n}\left(\left((1-t)^{1/k}+x/\sqrt{n}\right)^k-(1-t)\right),$$

$$h(x)=k(1-t)^{1-1/k}\,x,\ x\in D[0,1],\ D[0,1]$$

— пространство функций без разрывов второго рода на $[0,1]$. *Тогда* $h_n(x)$ *сходится в метрике Скорохода [18] к* $h(x)$ *равномерно на ограниченных множествах из* $D[0,1]$.

Очевидно, что асимптотически $X_2(t)=h\left(\tilde{X}_2(t)\right)$. Применив это утверждение к случайному процессу $\tilde{X}_2(t)$, получим, что асимптотическая функция ковариаций рассматриваемого процесса имеет вид:

$$K_2(s,t)=$$

$$=EX_2(s)X_2(t)-EX_2(s)EX_2(t)\underset{n_2\to\infty}{\to}(1-t)(1-s)\left(\frac{q^2\left(1-(1-s)^{1/q}\right)}{(1-s)^{1/q}}\right).$$

Распределение статистики (4) в асимптотике совпадает с распределением максимума модуля случайного процесса

$$Z_n(t)=\sqrt{n_1}\left(\left(\widehat{P_1^s}(t)-(1-t)\right)-\left(\left(\widehat{P_2^s}(t)\right)^q-(1-t)\right)\right)=$$

$$=X_1(t)-\sqrt{\rho}\,X_2(t),\ 0\le t\le 1.$$

Обозначим функцию ковариаций этого процесса как $K_n(s,t)$.

Лемма 3. *Пусть* $0\le s\le t\le 1,$ *тогда*

$$K_n(s,t)\underset{\substack{n_1\to\infty\\ n_2\to\infty}}{\to}K(s,t)=K_1(s,t)+\rho K_2(s,t)=$$

$$=(1-t)\left(s+\frac{\rho q^2\left(1-(1-s)^{1/q}\right)}{(1-s)^{1/q-1}}\right).$$

Обозначим

$$Y_n(t) = \frac{1}{\sqrt{q^2\rho+1}} Z_n(t).$$

При стандартных ограничениях $0 \le s \le t \le 1$ процесс $Y_n(t)$ сходится к гауссову процессу $Y(t)$ [17] с параметрами

$$E[Y(t)] = 0,\ E[Y(s)Y(t)] = \frac{1}{q^2\rho+1}(1-t)\left(s + \frac{\rho q^2\left(1-(1-s)^{1/q}\right)}{(1-s)^{1/q-1}}\right) =$$

$$= (1-t)\frac{a_2(1-s)^{1/q-1} + a_1 - (1-s)^{1/q}}{(1-s)^{1/q-1}}, \quad 0 \le s \le t \le 1.$$

Введем преобразование

$$\tau(t) = \frac{a_2(1-t)^{1/q-1} + a_1 - (1-t)^{1/q}}{a_2(1-t)^{1/q-1} + a_1}: \quad [0,1] \to [0,1].$$

Ввиду того что производная этого преобразования неотрицательна, существует обратное преобразование $t = \psi(\tau)$.

Лемма 4. *Процесс*

$$W(\tau) = Y(\psi(\tau))\frac{(1-\psi(\tau))^{1/q-1}}{a_2(1-\psi(\tau))^{1/q-1} + a_1}$$

является броуновским мостом.

◄ Имеем

$$W(\tau) = Y(\psi(\tau))\frac{(1-\psi(\tau))^{1/q-1}}{a_2(1-\psi(\tau))^{1/q-1} + a_1} = Y(\psi(\tau))\lambda(\psi(\tau)).$$

При $0 \le u \le v \le 1,\ \psi(u) = s,\ \ \psi(v) = t$:

$$E[W(\tau)] = 0,$$

$$E[W(u)W(v)] = \lambda(\psi(u))\lambda(\psi(v))E[Y(\psi(u))Y(\psi(v))] =$$

$$= \lambda(s)\lambda(t)E[Y(s)Y(t)] =$$

$$= \frac{(1-s)^{1/q-1}}{a_2(1-s)^{1/q-1} + a_1}\frac{(1-t)^{1/q-1}}{a_2(1-t)^{1/q-1} + a_1}(1-t)\frac{a_2(1-s)^{1/q-1} + a_1 - (1-s)^{1/q}}{(1-s)^{1/q-1}} =$$

$$= \frac{(1-t)^{1/q}}{a_2(1-t)^{1/q-1} + a_1}\frac{a_2(1-s)^{1/q-1} + a_1 - (1-s)^{1/q}}{a_2(1-s)^{1/q-1} + a_1} = u(1-\nu).$$ ►

Лемма 5.

$$P\left(\sup_t\left|\widehat{P_1^s}(t)-(1-t)\right|\underset{n_1\to\infty}{\to}0\right)=1,\ P\left(\sup_t\left|\left(\widehat{P_2^s}(t)\right)^q-(1-t)\right|\underset{n_2\to\infty}{\to}0\right)=1.$$

◄ В силу теоремы Гливенко и равномерной непрерывности на [0, 1] степенных функций с положительным показателем степени утверждение леммы выполняется.

В результате получаем, что асимптотическим распределением предложенной статистики (4) является распределение максимума модуля броуновского моста, которое является распределением Колмогорова [15]. ►

Оценка параметра модели Кокса — Лемана. На практике значение коэффициента k чаще всего неизвестно. Его оценка проводится максимизацией по k частной функции правдоподобия по полученным выборкам [19] или минимизацией по k «расстояния» между эмпирическими функциями надежности $\widehat{P_1^s}(t)$, $\left(\widehat{P_2^s}(t)\right)^q$. Последний метод оценки носит название минимаксного. В качестве «расстояния» между $\widehat{P_1^s}(t)$, $\left(\widehat{P_2^s}(t)\right)^q$ удобно брать статистику (4). В силу того, что цензурированные данные имеют сложную структуру анализ статистических свойств оценок (среднее значение, дисперсия) как первым, так и вторым методами проводят методом Монте-Карло [20]. Отсюда следует, что в качестве оценки $\widehat{k^s}$ необходимо использовать значение $\widehat{k^s}=m_2\hat{q}/m_1$, где $\hat{q}$ доставляет минимум статистике типа Колмогорова — Смирнова (4), т. е. $\hat{q}=\arg\min T^s(\tilde{q})$, $\tilde{q}=m_1\tilde{k}/m_2$, где $\tilde{k}$, $1\le\tilde{k}\le K$, K — некоторое граничное значение (для расчетов $K=5$).

В целях сравнения свойств оценок $\hat{k}$ (минимаксная оценка k на основе (2)) и $\widehat{k^s}$ (минимаксная оценка k на основе (4)) моделирование проводилось следующим образом.

1. Моделируется выборка из экспоненциального распределения с параметром 0,001, состоящая из одинаково распределенных случайных величин $\vec{\zeta}=\left(\xi_1^1,\ldots,\xi_{m_1n_1}^1\right)$.

2. Элементы выборки $\vec{\zeta}$ случайным образом разбиваются на n_1 групп по m_1 элементов в каждой группе. Определяются минимумы наработок до отказа в каждой группе $\theta_1^i=\min\left(\xi_1^{1,i},\ldots,\xi_{m_1}^{1,i}\right)$, $i=\overline{1,n_1}$.

3. Задается некоторое значение параметра Кокса — Лемана k. Для второго режима работы аналогично моделируются $m_2 n_2$ одинаково распределенных случайных величин с функцией распределения $1-\left(1-\left(F_0\left(t\right)\right)^k\right)^{1/k}$, где $F_0\left(t\right)$ — функция распределения экспоненциального распределения из п. 1. Как и в п. 2, для второго режима работы находятся минимумы наработок до отказа в каждой группе $\theta_2^j = \min\left(\xi_1^{2,j},\ldots,\xi_{m_2}^{2,j}\right)$, $j=\overline{1, n_2}$.

4. По двум выборкам из минимумов наработок до отказа в полученных группах в двух режимах работы $\Theta_1=\left(\theta_1^1,\ldots,\theta_1^{n_1}\right)$, $\Theta_2=\left(\theta_2^1,\ldots,\theta_2^{n_2}\right)$ соответственно для переменного значения $\tilde{k}$ вычисляется значение статистики типа Колмогорова — Смирнова $T\left(\tilde{k}\right)$ (2), где в качестве параметра Кокса — Лемана использовано моделируемое значение $\tilde{k}$.

5. Аналогично работе [6] методом перебора $\tilde{k}, 1\le\tilde{k}\le K$, определяется оценка $\hat{k}$, минимизирующая значение статистики $T\left(\tilde{k}\right)$, $\hat{k} = \arg\min T\left(\tilde{k}\right)$.

6. Предполагается, что не наблюдаются отказы элементов, а наблюдаются только отказы систем. Для некоторого значения $\tilde{k}$, $1\le\tilde{k}\le K$, вычисляется $\tilde{q}=m_1\tilde{k}/m_2$. По двум полным выборкам $\Theta_1=\left(\theta_1^1,\ldots,\theta_1^{n_1}\right)$, $\Theta_2=\left(\theta_2^1,\ldots,\theta_2^{n_2}\right)$ определяется значение статистики типа Колмогорова — Смирнова $T^s\left(\tilde{q}\right)$.

7. Методом перебора $\tilde{q}$ определяется оценка $\hat{q}$, минимизирующая значение статистики $T^s\left(\tilde{q}\right)$, $\hat{q}=\arg\min T^s\left(\tilde{q}\right)$. Вычисляется соответствующее значение оценки параметра Кокса — Лемана $\widehat{k^s}=m_2\hat{q}/m_1$.

Алгоритм моделирования повторяется 500 раз и по полученным данным строятся гистограммы оценок параметра Кокса — Лемана $\hat{k}$, $\widehat{k^s}$.

В качестве примера на рис. 2 показаны результаты моделирования оценок при значениях $m_1=2$, $m_2=3$, $k=3$, $n_1=100$, $n_2=200$. Приведены гистограммы оценок и значения средних и среднеквадратических отклонений выборок $M\hat{k}=3{,}0642, \sigma=0{,}4834$, $M\widehat{k^s}=3{,}0704, \sigma^s=0{,}4527$. На основе этих результатов можно сделать выводы, что распределения обеих оценок можно полагать приближенно нормальным; средние значения практически

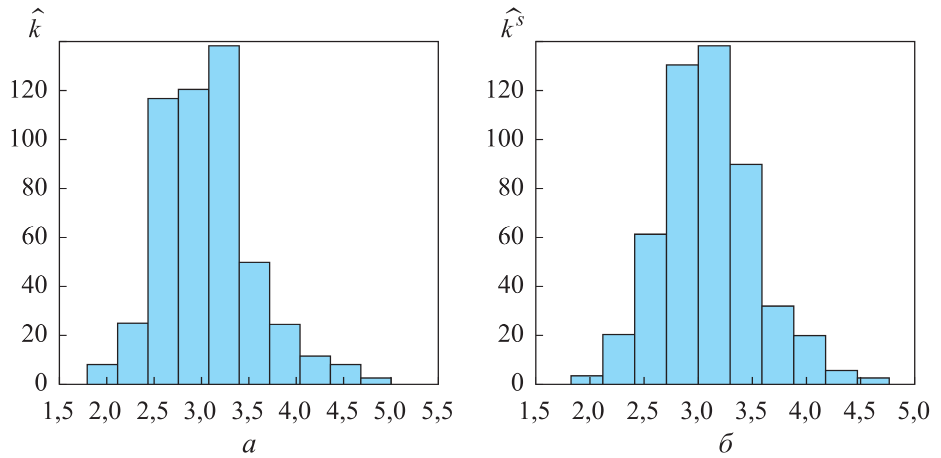

Рис. 2. Гистограммы оценок $\hat{k}$, $\widehat{k^s}$ при $m_1 = 2$, $m_2 = 3$, $k = 3$, $n_1 = 100$, $n_2 = 200$ для $\hat{k} = \arg\min T\left(\tilde{k}\right)$ (*а*) и $\widehat{k^s} = m_2\hat{q} / m_1$, $\hat{q} = \arg\min T^s\left(\tilde{q}\right)$ (*б*)

одинаковыми, а среднеквадратическое отклонение оценки $\widehat{k^s}$ несколько меньшим, хотя и статистически незначимым.

Сравнение результатов моделирования при различных наборах параметров m_1, m_2, k, n_1, n_2 показало, что статистические свойства оценок параметра модели Кокса — Лемана зависят от их значений. В общем случае нельзя отдать предпочтение одной из оценок $\hat{k}$ или $\widehat{k^s}$.

Заключение. Проведено сравнение двух разработанных авторами критериев типа Колмогорова — Смирнова для непараметрической проверки моделей Кокса — Лемана. Ранее в [10] непараметрическая проверка этих моделей не рассматривалась, поэтому какие-либо результаты в этой области отсутствуют.

Задача решена в терминах теории испытаний сложных технических систем. Показано, что применение обычной оценки функции надежности систем вместо оценок функции надежности элементов Каплана — Мейера не приводит к существенному ухудшению скорости сходимости распределения к предельному. При этом преимуществом новой статистики является увеличение скорости вычисления точных вероятностей.

Исследование мощности обоих критериев при проверке гипотезы (1) против альтернатив (6) показало, что, как и для обычных критериев однородности Колмогорова — Смирнова, значимое уменьшение вероятности ошибок второго рода при одинаковых объемах выборок наблюдается начиная с $n_1 = n_2 = 40$. При этом с увеличением k, т. е. при возрастании нелинейности гипотез, вероятность ошибки второго рода увеличивается.

При заданных значениях параметров $m_1 = 2,\ m_2 = 3$ для статистики с использованием оценок Каплана — Мейера вероятность ошибки второго рода меньше, т. е. чувствительность критерия, использующего данные о структуре систем, оказывается выше.

Наряду со сравнением мощности двух критериев методом Монте-Карло исследование тесно связано с задачей проблемы точности минимаксных оценок параметра модели Кокса — Лемана. Показано, что их точность зависит от параметров испытаний $m_1,\ m_2,\ n_1,\ n_2$.

ЛИТЕРАТУРА

[1] Гнеденко Б.В., Беляев Ю.К., Соловьев А.Д. Математические методы в теории надежности. М., Либроком, 2012.

[2] Abdushukurov A.A. Estimation of survival function in cox regression model under random censoring from both sides. *Commun. Stat. Theory Methods*, 2015, vol. 44, iss. 3, pp. 533–553. DOI: https://doi.org/10.1080/03610926.2012.746981

[3] Nikulin M., Wu H.I. The cox model and its applications. Berlin, Springer, 2016.

[4] Escobar L.A., Meeker W.Q. A review of accelerated test models. *Statist. Sci.*, 2006, vol. 21, iss. 4, pp. 552–577. DOI: https://doi.org/10.1214/088342306000000321

[5] Balakrishnan N., Tripathi R.C., Kannan N., et al. Some nonparametric precedence type tests based on progressively censored samples and evaluation of power. *J. Stat. Plan. Inference*, 2010, vol. 140, iss. 2, pp. 559–573.
DOI: https://doi.org/10.1016/j.jspi.2009.08.003

[6] Dimitrova D.S., Kaishev V.K., Tan S. Computing the Kolmogorov — Smirnov distribution when the underlying CDF is purely discrete, mixed or continuous. *J. Stat. Softw.*, 2020, vol. 95, iss. 10, pp. 1–42. DOI: https://doi.org/10.18637/jss.v095.i10

[7] Тимонин В.И., Тянникова Н.Д. Применение оценок Каплана — Мейера для проверки степенной гипотезы Кокса по двум прогрессивно цензурированным выборкам. *Вестник МГТУ им. Н.Э. Баумана. Сер. Естественные науки*, 2015, № 6 (63), с. 68–84. DOI: http://dx.doi.org/10.18698/1812-3368-2015-6-68-84

[8] Ng N., Balakrishnan N. Precedence-type test based on Kaplan — Meier estimator of cumulative distribution function. *J. Stat. Plan. Inference*, 2010, vol. 140, iss. 8, pp. 2295–2311. DOI: https://doi.org/10.1016/j.jspi.2010.01.025

[9] Su P., Li C., Shyr Y. Sample size determination for paired right-censored data based on the difference of Kaplan — Meier estimates. *Comput. Stat. Data Anal.*, 2014, vol. 74, pp. 39–51. DOI: https://doi.org/10.1016/j.csda.2013.12.006

[10] Bagdonavichus V., Kruopis J. Nikulin M.S. Nonparametric tests for censored data. London, ISTE Ltd, 2011.

[11] Balakrishnan N., Cramer E. The art of progressive censoring. New York, Springer, 2014.

[12] Maturi T.A., Coolen-Schrijner P., Coolen F.P. Nonparametric predictive comparison of lifetime data under progressive censoring. *J. Stat. Plan. Inference*, 2010, vol. 140, iss. 2, pp. 515–525. DOI: https://doi.org/10.1016/j.jspi.2009.07.027

[13] Тимонин В.И., Тянникова Н.Д. Метод вычисления точных распределений статистик типа Колмогорова — Смирнова в случае нарушения однородности и независимости анализируемых выборок. *Наука и образование: научное издание,* 2014, № 11. DOI: 10.7463/1114.0740251

[14] Simard R., L'Ecuyer P. Computing the two-sided Kolmogorov — Smirnov distribution. *J. Stat. Softw*., 2011, vol. 39, iss. 11, pp. 1–18.
DOI: https://doi.org/10.18637/jss.v039.i11

[15] Лебедев А.В., Фадеева Л.Н. Теория вероятностей и математическая статистика. М., ЭКСМО, 2010.

[16] Lemeshko B., Veretel'nikova I. Power of *k*-sample tests aimed at checking the homogeneity of laws. *Meas. Tech*., 2018, vol. 61, no. 2, pp. 647–654.
DOI: https://doi.org/10.1007/s11018-018-1479-1

[17] Hajek J., Sidak Z. Theory of rank tests. London, Academic Press, 1999.

[18] Борзых Д.А. Об одном классе функционалов, непрерывных в топологии Скорохода. *Вестник Воронежского государственного университета. Сер. Физика. Математика*, 2016, № 4, с. 83–88. EDN: WYMNWL

[19] Han D., Bai T. On the maximum likelihood estimation for progressively censored lifetimes from constant-stress and step-stress accelerated tests. *Electron. J. Appl. Stat. Anal*., 2019, vol. 12, no. 2, pp. 392–404.
DOI: https://doi.org/10.1285/i20705948v12n2p392

[20] Kroese D.P., Brereton T., Taimre T., et al. Why the Monte Carlo method is so important today. *WIREs Comput Stat.*, 2014, vol. 6, iss. 6, pp. 386–392.
DOI: https://doi.org/10.1002/wics.1314

Тимонин Владимир Иванович — д-р физ.-мат. наук, доцент, профессор кафедры «Высшая математика» МГТУ им. Н.Э. Баумана (Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5, стр. 1).

Тянникова Нина Дмитриевна — канд. физ.-мат. наук, доцент кафедры «Высшая математика» МГТУ им. Н.Э. Баумана (Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5, стр. 1).

Просьба ссылаться на эту статью следующим образом:

Тимонин В.И., Тянникова Н.Д. Анализ мощности двух критериев типа Колмогорова — Смирнова проверки степенной модели Кокса для прогрессивно цензурированных выборок. *Вестник МГТУ им. Н.Э. Баумана. Сер. Естественные науки*, 2024, № 2 (113), с. 57–73. EDN: KDBJJP

POWER ANALYSIS OF TWO KOLMOGOROV — SMIRNOV TYPE CRITERIA IN TESTING THE COX POWER MODEL FOR THE PROGRESSIVELY CENSORED SAMPLES

V.I. Timonin timonin@bmstu.ru
N.D. Tiannikova tiannikova@bmstu.ru

Bauman Moscow State Technical University, Moscow, Russian Federation

Abstract

Testing a technical system often leads to a problem of comparing reliability indicators of their elements in different modes (operation in different climatic zones, comparing results of laboratory tests and real operation data, etc.). At the same time, laws of the element time to failure distribution are unknown. Nature of the available data causes additional difficulties in solving this problem, i.e. a failure of one element in a sequential system leads to the fact that time to failure of the remaining usable elements remains unknown (censored). In this regard, it becomes relevant to solve such a problem using the nonparametric (free from knowing distribution) methods. Previously, the authors developed a Kolmogorov — Smirnov type criterion making it possible to check power dependence of the reliability functions with elements of the sequential systems (nonparametric Cox power model). Statistics of this criterion uses the Kaplan — Meier estimates of the elements reliability functions constructed from developments in systems composed of them. However, a feature of the Kaplan — Meier estimates is their slow convergence to the theoretical reliability function on the distribution right tail, which could lead to a decrease in the criterion power (decrease in sensitivity). Here, another criterion of the Kolmogorov — Smirnov type is proposed to solve a similar problem, its statistics is not using the Kaplan — Meier estimates and is based on comparing estimates of the system reliability functions. Exact and asymptotic distributions of its statistics are obtained. Detailed study of the two criteria power is carried out using numerical analysis and statistical simulation. The Cox parameter estimates accuracy is obtained by minimizing the compared criteria statistics and is analyzed using the Monte Carlo methods

Keywords

Kolmogorov — Smirnov criterion, Cox model, Kaplan — Meier estimate, criterion power, sequential systems, censoring

Received 14.04.2023
Accepted 07.06.2023

REFERENCES

[1] Gnedenko B.V., Belyaev Yu.K., Solovyev A.D. Matematicheskie metody v teorii nadezhnosti [Mathematical methods in reliability theory]. Moscow, Librokom Publ., 2012.

[2] Abdushukurov A.A. Estimation of survival function in cox regression model under random censoring from both sides. *Commun. Stat. Theory Methods*, 2015, vol. 44, iss. 3, pp. 533–553. DOI: https://doi.org/10.1080/03610926.2012.746981

[3] Nikulin M., Wu H.I. The cox model and its applications. Berlin, Springer, 2016.

[4] Escobar L.A., Meeker W.Q. A review of accelerated test models. *Statist. Sci.*, 2006, vol. 21, iss. 4, pp. 552–577. DOI: https://doi.org/10.1214/088342306000000321

[5] Balakrishnan N., Tripathi R.C., Kannan N., et al. Some nonparametric precedence type tests based on progressively censored samples and evaluation of power. *J. Stat. Plan. Inference*, 2010, vol. 140, iss. 2, pp. 559–573.
DOI: https://doi.org/10.1016/j.jspi.2009.08.003

[6] Dimitrova D.S., Kaishev V.K., Tan S. Computing the Kolmogorov — Smirnov distribution when the underlying CDF is purely discrete, mixed or continuous. *J. Stat. Softw.*, 2020, vol. 95, iss. 10, pp. 1–42. DOI: https://doi.org/10.18637/jss.v095.i10

[7] Timonin V.I., Tyannikova N.D. Application of Kaplan — Meier estimates to testing Cox power hypothesis for two progressive lycensored samples. *Herald of the Bauman Moscow State Technical University, Series Natural Sciences*, 2015, no. 6 (63), pp. 68–84 (in Russ.). DOI: http://dx.doi.org/10.18698/1812-3368-2015-6-68-84

[8] Ng N., Balakrishnan N. Precedence-type test based on Kaplan — Meier estimator of cumulative distribution function. *J. Stat. Plan. Inference*, 2010, vol. 140, iss. 8, pp. 2295–2311. DOI: https://doi.org/10.1016/j.jspi.2010.01.025

[9] Su P., Li C., Shyr Y. Sample size determination for paired right-censored data based on the difference of Kaplan — Meier estimates. *Comput. Stat. Data Anal.*, 2014, vol. 74, pp. 39–51. DOI: https://doi.org/10.1016/j.csda.2013.12.006

[10] Bagdonavichus V., Kruopis J. Nikulin M.S. Nonparametric tests for censored data. London, ISTE Ltd, 2011.

[11] Balakrishnan N., Cramer E. The art of progressive censoring. New York, Springer, 2014.

[12] Maturi T.A., Coolen-Schrijner P., Coolen F.P. Nonparametric predictive comparison of lifetime data under progressive censoring. *J. Stat. Plan. Inference*, 2010, vol. 140, iss. 2, pp. 515–525. DOI: https://doi.org/10.1016/j.jspi.2009.07.027

[13] Timonin V.I., Tyannikova N.D. The method of calculating the exact distributions of the Kolmogorov — Smirnov statistics in case of violation of homogeneity and independence of the analyzed samples. *Nauka i obrazovanie: nauchnoe izdanie* [Science and Education: Scientific Publication], 2014, no. 11 (in Russ.).
DOI: 10.7463/1114.0740251

[14] Simard R., L'Ecuyer P. Computing the two-sided Kolmogorov — Smirnov distribution. *J. Stat. Softw.*, 2011, vol. 39, iss. 11, pp. 1–18.
DOI: https://doi.org/10.18637/jss.v039.i11

[15] Lebedev A.V., Fadeeva L.N. Teoriya veroyatnostey i matematicheskaya statistika [Theory of probability and mathematical statistics]. Moscow, EKSMO Publ., 2010.

[16] Lemeshko B., Veretel'nikova I. Power of *k*-sample tests aimed at checking the homogeneity of laws. *Meas. Tech.*, 2018, vol. 61, no. 2, pp. 647–654.
DOI: https://doi.org/10.1007/s11018-018-1479-1

[17] Hajek J., Sidak Z. Theory of rank tests. London, Academic Press, 1999.

[18] Borzykh D.A. On a class of functionals continuous in the Skorokhod topology. *Vestnik Voronezhskogo gosudarstvennogo universiteta. Ser. Fizika. Matematika* [Proceedings of Voronezh State University. Ser. Physics. Mathematics], 2016, no. 4, pp. 83–88 (in Russ.). EDN: WYMNWL

[19] Han D., Bai T. On the maximum likelihood estimation for progressively censored lifetimes from constant-stress and step-stress accelerated tests. *Electron. J. Appl. Stat. Anal.*, 2019, vol. 12, no. 2, pp. 392–404.
DOI: https://doi.org/10.1285/i20705948v12n2p392

[20] Kroese D.P., Brereton T., Taimre T., et al. Why the Monte Carlo method is so important today. *WIREs Comput Stat.*, 2014, vol. 6, iss. 6, pp. 386–392.
DOI: https://doi.org/10.1002/wics.1314

Timonin V.I. — Dr. Sc. (Phys.-Math.), Assoc. Professor, Professor, Department of Higher Mathematics, Bauman Moscow State Technical University (2-ya Baumanskaya ul. 5, str. 1, Moscow, 105005 Russian Federation).

Tiannikova N.D. — Cand. Sc. (Phys.-Math.), Assoc. Professor, Department of Higher Mathematics, Bauman Moscow State Technical University (2-ya Baumanskaya ul. 5, str. 1, Moscow, 105005 Russian Federation).

Please cite this article in English as:

Timonin V.I., Tiannikova N.D. Power analysis of two Kolmogorov — Smirnov type criteria in testing the Cox power model for the progressively censored samples. *Herald of the Bauman Moscow State Technical University, Series Natural Sciences*, 2024, no. 2 (113), pp. 57–73 (in Russ.). EDN: KDBJJP