

НОВЫЕ МАТЕМАТИЧЕСКИЕ И ФИЗИЧЕСКИЕ МЕТОДЫ

УДК 519.248

ПРОВЕРКА ГИПОТЕЗЫ ЛЕМАНА ПО НЕПОЛНЫМ ДАННЫМ ДЛЯ ПАРАЛЛЕЛЬНЫХ СИСТЕМ

© Авторы, 2024

doi: 10.25210/jfop-2403-QGSOUB | edn: QGSOUB

Тимонин В. И. — д.ф.-м.н., доцент, МГТУ им. Н.Э. Баумана, Москва. E-mail: timonin@bmstu.ru

Тяникова Н. Д. — к.ф.-м.н., доцент, МГТУ им. Н.Э. Баумана, Москва. E-mail: tiannikova@bmstu.ru

Аннотация

В работе рассматривается проблема проверки гипотезы о степенной зависимости между функциями распределения наработок до отказа изделий (элементов), функционирующих в двух различных режимах (гипотеза Лемана). В каждом режиме исходными данными являются наработки до отказа параллельных систем, составленных из этих элементов. Нарботки систем совпадают с наибольшей наработкой составляющих ее элементов. Системы в различных режимах могут иметь различную кратность. Время испытаний ограничено, то есть не все системы отказывают в процессе испытаний. Авторами предложен критерий типа Реньи проверки гипотезы, статистика которого основана на сравнении предложенных авторами ранее оценок типа Каплана-Мейера функций распределения элементов. Получен метод вычисления точных распределений статистики критерия. Доказана сходимость распределения данной статистики к стандартному распределению Реньи при условии справедливости проверяемой гипотезы. Методом Монте-Карло исследованы статистические свойства минимаксной оценки параметра степенной модели Лемана.

Ключевые слова: Непараметрическая статистика, гипотеза Лемана, критерий типа Реньи, оценка Каплана-Мейера, параллельные системы

Abstract

The paper considers the problem of testing the hypothesis about the power-law dependence between the distribution functions of the operating times before failure of products (elements) operating in two different modes (Lehmann hypothesis). In each mode, the initial data are the operating times before failure of parallel systems composed of these elements. The operating times of the systems coincide with the greatest operating time of their constituent elements. Systems in different modes can have different multiplicity. The testing time is limited, i. e. not all systems fail during testing. The authors proposed a Renyi-type criterion for testing the hypothesis, the statistics of which are based on a comparison of the Kaplan-Meier-type estimates of the element distribution functions proposed earlier by the authors. A method for calculating the exact distributions of the criterion statistics is obtained. The convergence of the distribution of this statistics to the standard Renyi distribution is proved, provided that the hypothesis being tested is valid. The statistical properties of the minimax estimate of the Lehmann power-law model parameter are investigated using the Monte Carlo method.

Keywords: Renyi-type criteria, Leman hypothesis, nonparametric statistics, parallel systems

Введение

На практике часто возникает задача сравнения законов распределения наработок до отказа элементов в различных режимах в случае, когда имеются данные по отказам состоящих из этих элементов параллельных систем. В этом случае наблюдаются наработки только наиболее надежных элементов, входящих в эти системы. В работах авторов [1,2] для случая полных результатов испытаний (все системы доводятся до отказа) решалась задача проверки степенной зависимости функций распределения наработок элементов в разных режимах по наработкам систем, состоящих из этих элементов. Авторами было предложено использовать критерий типа Колмогорова-Смирнова, построенный на основе сравнения оценок функций распределения наработок до отказа элементов типа Каплана-Мейера. В настоящей работе авторы предлагают и исследуют критерий типа Реньи, который позволяет проверять гипотезу Лемана в случае, когда не все системы испытаны до отказа.

1. Постановка задачи

В двух различных режимах работы ε_1 и ε_2 испытывают n_1 сложных систем, состоящих из m_1 параллельно соединенных элементов, и n_2 систем, состоящих из m_2 параллельно соединенных элементов соответственно. Испытания проходят таким образом, что $(m_j - 1), j = 1, 2$ отказов не наблюдаются, а наблюдаются только отказы последних элементов систем. Предполагается, что на распределение наработок до отказа элементов в системах не влияют отказы составляющих систему элементов. По испытаниям такого рода наблюдают две выборки $\Theta_1 = (\theta_1^1, \dots, \theta_1^{n_1}), \Theta_2 = (\theta_2^1, \dots, \theta_2^{n_2})$ из максимумов наработок до отказа элементов систем, работающих в режимах ε_1 и ε_2 .



По испытаниям такого рода требуется проверить гипотезу Лемана [3]

$$H_0 : F_1(t) = [F_2(t)]^k, \quad (1)$$

где $F_1(t)$ — функция распределения наработок до отказа элементов в режиме ε_1 , $F_2(t)$ — функция распределения наработок до отказа в режиме ε_2 , k — некоторое фиксированное число.

В работе [1] для проверки гипотезы (1) в случае, когда наблюдаются отказы всех систем, был предложен критерий типа Колмогорова–Смирнова. Он основан на сравнении впервые предложенных в [2] оценок $\hat{F}_{\theta_1}(t), \hat{F}_{\theta_2}(t)$ функций распределения $F_1(t), F_2(t)$ — аналогов оценок Каплана–Мейера [4,5] функций надежности по цензурированным выборкам, вид и свойства которых были рассмотрены в работе [3].

Эти оценки имеют вид

$$\hat{F}_{\theta_j}(t) = \begin{cases} 0, d_j(t) = 0; \\ \prod_{i=1}^{n_j - d_j(t)} \left(1 - \frac{1}{m_j(n_j - i + 1)} \right), 1 \leq d_j(t) \leq (n_j - 1); j = 1, 2, \\ 1, d_j(t) = n_j, \end{cases} \quad (2)$$

где $d_j(t)$ — количество отказов систем к моменту времени t выборки Θ_j .

Свойства (2) более подробно рассмотрены в работе [3].

В силу того, что далеко не всегда удастся испытать все системы до отказа, важное значение имеет задача проверки (1) для случая, когда не все системы испытаны до отказа [6].

Введем некоторые обозначения. Обозначим $r_1 \leq n_1$ — число отказов систем за время испытаний в режиме ε_1 , аналогично $r_2 \leq n_2$ — число отказов систем в режиме ε_2 . Пусть $G_j(t) = F_j^{m_j}(t), j = 1, 2$ — функции распределения наработок до отказа систем. Объединённую эмпирическую функцию распределения изделий по выборкам $(\theta_1^1, \dots, \theta_1^{r_1}), (\theta_2^1, \dots, \theta_2^{r_2})$ обозначим $\hat{F}_0(t)$. Как и в работе [2]

$$\hat{F}_0(t) = \frac{n_1}{n_1 + n_2} \hat{G}_1^{\frac{1}{m_1}} + \frac{n_2}{n_1 + n_2} \hat{G}_2^{\frac{k}{m_2}}.$$

Для проверки справедливости (1) в настоящей работе предлагается критерий типа Реньи [7], статистика которого имеет вид

$$R_\lambda = \frac{m_1 m_2 \sqrt{\rho n_2}}{\sqrt{k^2 \rho m_1^2 + m_2^2}} \sqrt{\frac{\lambda}{1 - \lambda}} \max_{t: \psi(\hat{F}_0(t)) < 1 - \lambda} \frac{\hat{F}_0(t)^{m_1 + \frac{m_2}{k} - 1} \cdot \left| \hat{F}_{\theta_1}(t) - (\hat{F}_{\theta_2}(t))^k \right|}{k_2 \hat{F}_0(t)^{\frac{m_2}{k}} + k_1 \hat{F}_0(t)^{m_1} - \hat{F}_0(t)^{m_1 + \frac{m_2}{k}}}. \quad (3)$$

Здесь $\psi(x) = x^{\frac{m_2}{k}} / (k_2 x^{\frac{m_2}{k} - m_1} + k_1)$, $x \in (0, 1]$, $\psi(0) = 0$, строго возрастающая функция,

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

В (3) в случае $\hat{F}_0(t) = 0$ функция, максимум которой определяется, равна 0.

Тогда продолжительность испытаний можно определить следующим образом. Пусть $0 < \lambda < 1$ — некоторое фиксированное число. Испытания прекращаются в момент времени, когда нарушается неравенство $\psi(\hat{F}_0(t)) < 1 - \lambda$. Параметр λ назовём параметром Реньи.

2. Точные распределения статистики R_λ

Для вычисления точных распределений статистики R_λ , введем модель случайного блуждания частицы по целочисленной решетке [8,9]. Для этого запишем статистику (3) в виде $R_\lambda = \max_t H(t)$, где $H(t)$ определяется видом статистики.

Для вычисления точных распределений статистики (3) составим $Q = \{\theta_1^1, \dots, \theta_1^{n_1}, \theta_2^1, \dots, \theta_2^{n_2}\}$ — объединённую выборку из отказов систем в обоих режимах. Расположим элементы выборки Q в порядке убывания и составим из них ряд $\Gamma = \{\gamma_1 \geq \gamma_2 \geq \gamma_3 \geq \dots \geq \gamma_{n_1 + n_2}\}$.

Замечание. То, что рассматриваются полные выборки $Q = \{\theta_1^1, \dots, \theta_1^{n_1}, \theta_2^1, \dots, \theta_2^{n_2}\}$ и $\Gamma = \{\gamma_1 \geq \gamma_2 \geq \gamma_3 \geq \dots \geq \gamma_{n_1 + n_2}\}$ из отказов систем, введено лишь для удобства изложения. В дальнейшем не требуется, чтобы при расчете распределений использовались полные данные.

Введем вектор $\mathbf{Z} = (z_1, z_2, \dots, z_{n_1+n_2})$, состоящий из n_1 единиц и n_2 нулей, где

$$z_l = \begin{cases} 1, & \text{если } \gamma_l = \theta_1^i, i = \overline{1, n_1} \\ 0, & \text{если } \gamma_l = \theta_2^i, i = \overline{1, n_2} \end{cases}, l = 1, \dots, n_1 + n_2.$$

В работе [2] была доказана следующая лемма.

Лемма 1. Распределение вероятностей векторов $\mathbf{Z} = (z_1, z_2, \dots, z_{n_1+n_2})$ не зависит от распределения наработок до отказа и определяется следующим выражением:

$$p(\mathbf{Z}) = \prod_{l=1}^{n_1+n_2} \left(\frac{((n_1 - V_{l-1})m_1 k)^{z_l} ((n_2 - U_{l-1})m_2)^{1-z_l}}{(n_1 - V_{l-1})m_1 k + (n_2 - U_{l-1})m_2} \right),$$

где $V_l = \sum_{i=1}^l z_i$, $V_0 = 0$ — количество единиц в векторе $\mathbf{Z}$ до l -ого места включительно, $U_l = l - \sum_{i=1}^l z_i$, $U_0 = 0$ — количество нулей в векторе $\mathbf{Z}$ до l -ого места включительно.

Пусть (i, j) , $0 \leq i \leq n_1$, $0 \leq j \leq n_2$ — целочисленная решетка на плоскости. Частица на первом шаге выходит из точки $(0, 0)$ и на $(n_1 + n_2)$ -ом шаге она заканчивает блуждание в точке (n_1, n_2) , совершая n_1 скачков «вправо» и n_2 скачков «вверх». Траектории частицы будут находиться во взаимно однозначном соответствии с векторами $\mathbf{Z}$. Равенство $z_l = 1$, $l = 1, \dots, (n_1 + n_2)$ в векторе $\mathbf{Z}$ соответствует скачку вправо на l шаге, если же $z_l = 0$, $l = 1, \dots, (n_1 + n_2)$ — скачку вверх.

Точные вероятности $P(R_\lambda < h) = \pi_{n_1, n_2}(h)$ вычисляются при помощи алгоритма, основанного на модели случайного блуждания по множеству точек, которое показано на рис. 1.

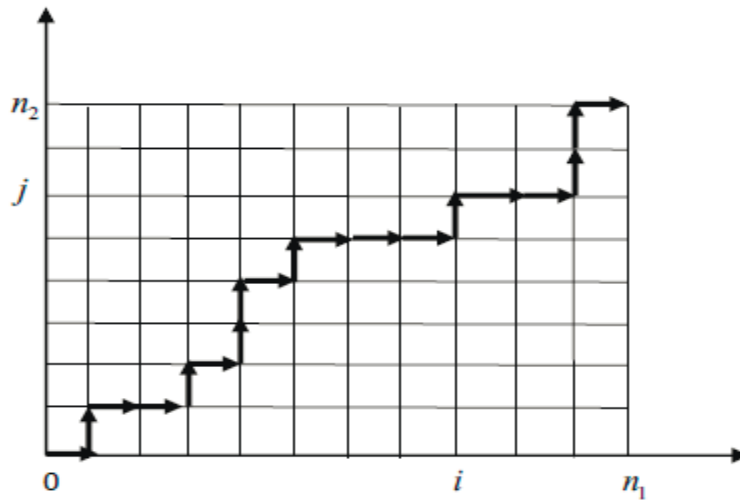


Рис. 1. Случайное блуждание частицы по целочисленной решетке

При прохождении блуждания через точку (i, j) (что соответствует $V_{i+j} = i$, $U_{i+j} = j$) функция $H(t)$ принимает значение, равное

$$H_{ij} = \frac{m_1 m_2 \sqrt{\rho n_2}}{\sqrt{k^2 \rho m_1^2 + m_2^2}} \sqrt{\frac{\lambda}{1-\lambda}} \frac{F_{ij}^{m_1 + \frac{m_2}{k} - 1} \cdot \Delta_{ij}}{k_2 F_{ij}^{\frac{m_2}{k}} + k_1 F_{ij}^{m_1} - F_{ij}^{m_1 + \frac{m_2}{k}}}.$$

Здесь

$$\Delta_{ij} = \left| \prod_{s_1=1}^i \left(1 - \frac{1}{m_1(n_1 - s_1 + 1)} \right) - \left(\prod_{s_2=1}^j \left(1 - \frac{1}{m_2(n_2 - s_2 + 1)} \right) \right)^k \right|$$

— значения модуля разности оценок типа

Каплана-Мейера функций распределения элементов двух выборок,

$$F_{ij} = \frac{n_1}{n_1 + n_2} \left(\frac{n_1 - i}{n_1} \right)^{\frac{1}{m_1}} + \frac{n_2}{n_1 + n_2} \left(\frac{n_2 - j}{n_2} \right)^{\frac{k}{m_2}}$$

— значение объединенной оценки $\hat{F}_0(t)$ функции распределения.

Теорема 1. Вероятность $P(R_\lambda < h) = \pi_{n_1, n_2}(h)$ может быть получена путем применения следующего рекуррентного соотношения:

$$\pi_{ij}(h) = \begin{cases} 1, & \text{если } i = 0, j = 0; \\ \left(\frac{m_2(n_2 - j + 1)}{m_1 k(n_1 - i) + m_2(n_2 - j + 1)} \pi_{i, j-1}(h) \right) \chi_{ij}(h), & i = 0, 1 \leq j \leq n_2; \\ \left(\frac{m_1 k(n_1 - i + 1)}{m_1 k(n_1 - i + 1) + m_2(n_2 - j)} \pi_{i-1, j}(h) \right) \chi_{ij}(h), & 1 \leq i \leq n_1, j = 0; \\ \left(\frac{m_1 k(n_1 - i + 1)}{m_1 k(n_1 - i + 1) + m_2(n_2 - j)} \pi_{i-1, j}(h) + \right. \\ \left. + \frac{m_2(n_2 - j + 1)}{m_1 k(n_1 - i) + m_2(n_2 - j + 1)} \pi_{i, j-1}(h) \right) \chi_{ij}(h), & 1 \leq i \leq n_1, 1 \leq j \end{cases} \quad (4)$$

$\chi_{ij}(h) = \begin{cases} 1, & a_{ij} \in A_0(h) \\ 0, & a_{ij} \notin A_0(h) \end{cases}$ — индикатор множества $A_0(h)$, где $A_0(h) = \{a_{ij}\}$, а индексы i, j удовлетворяют следующим условиям:

1. $i = n_1, j = n_2$.
2. $\frac{(F_{ij})^{\frac{m_2}{k}}}{k_2(F_{ij})^{\frac{m_2}{k}-m_1} + k_1} \geq 1 - \lambda$.
3. $\left(\frac{(F_{ij})^{\frac{m_2}{k}}}{k_2(F_{ij})^{\frac{m_2}{k}-m_1} + k_1} < 1 - \lambda \right) \wedge (H_{ij} < h)$.

Доказательство. Вероятность каждой траектории ω можно записать в следующем виде:

$$p(\mathbf{Z}) = p(\omega) = \prod_{l=1}^{n_1+n_2} \left(\frac{((n_1 - V_{l-1})m_1 k)^{z_l} ((n_2 - U_{l-1})m_2)^{1-z_l}}{(n_1 - V_{l-1})m_1 k + (n_2 - U_{l-1})m_2} \right) = \prod_{l=1}^{n_1+n_2} \lambda_l(\omega).$$

Пусть ω_{ij} — множество «частичных» траекторий, начинающихся в точке $(0,0)$ и оканчивающихся в точке (i,j) (соответствующие $\mathbf{Z}$ имеют i единиц и j нулей на l -ом месте, где $l = i + j$). Обозначим $p_{ij}(\omega) = \prod_{s=1}^l \lambda_s(\omega)$. Вероятность любой траектории, совершающей скачок $(i-1, j) \rightarrow (i, j)$ (что соответствует

$z_l = 1$), имеет множитель $\lambda_l(\omega) = \frac{m_1 k(n_1 - i + 1)}{m_1 k(n_1 - i + 1) + m_2(n_2 - j)}$. Если же происходит скачок $(i, j-1) \rightarrow (i, j)$

(что соответствует $z_l = 0$), то $\lambda_l(\omega) = \frac{m_2(n_2 - j + 1)}{m_1 k(n_1 - i) + m_2(n_2 - j + 1)}$. Пусть $\pi_{ij} = \sum_{\omega_{ij}} p_{ij}$. Тогда (4) следует из того, что в (i, j) за один скачок можно попасть только из точки $(i-1, j)$ или из $(i, j-1)$.

Множество $A_0(h)$ имеет такой вид вследствие того, что неравенство $R_\lambda < h$ не проверяется вне области $\psi(\hat{F}_0(t)) < 1 - \lambda$. При этом учитывается, что эмпирические функции распределения наработок до отказа систем равны $\hat{F}^1 = \frac{n_1 - i}{n_1}, \hat{F}^2 = \frac{n_1 - j}{n_2}, i = \overline{1, n_1}, j = \overline{1, n_2}$. Соотношения (1) и (2) задают начальные и граничные условия. ▯

В табл. 1 представлены вычисленные вероятности точного распределения статистики R_λ для квантилей $h=1.96, h=2.24$, которые являются соответственно квантилями уровней 0.9, 0.9498 асимптотического распределения статистики Реньи [10].

В целях сравнения сходимости распределения статистики (3) при различных m_1, m_2, k со сходимостью распределения классической статистики Реньи проверки однородности были вычислены точные вероятности $P(R_\lambda < h)$ при $m_1 = m_2 = k = 1$ для равных объёмов выборок при $\lambda = 0.2$.

Таблица 1. Точные вероятности $P(R_\lambda < h)$ в случае равных объёмов выборок при $m_1 = 2, m_2 = 3, \lambda = 0.2$

$n_1 = n_2$	$P(R_\lambda < h)$			
	$h=1.96$		$h=2.24$	
	$k=1$	$k=2$	$k=1$	$k=2$
100	0.9211	0.9180	0.9619	0.9592
400	0.9079	0.9080	0.9564	0.9551
700	0.9069	0.9082	0.9535	0.9532
1000	0.9057	0.9056	0.9530	0.9530
1300	0.9043	0.9050	0.9523	0.9526
1600	0.9043	0.9041	0.9523	0.9523
1900	0.9043	0.9040	0.9522	0.9520
2200	0.9037	0.9036	0.9519	0.9519
2500	0.9035	0.9033	0.9516	0.9516
∞	0.9000	0.9000	0.9498	0.9498

Таблица 2. Точные вероятности $P(R_\lambda < h)$ в случае равных объёмов выборок при $m_1 = m_2 = k = 1$

$n_1 = n_2$		100	1000	1600	1900	2200	2500	∞
$P(R_\lambda < h)$	$h=1.96$	0.9240	0.9063	0.9041	0.9038	0.9041	0.9034	0.9000
	$h=2.24$	0.9550	0.9530	0.9520	0.9522	0.9519	0.9518	0.9498

Результаты расчёта показали, что скорость сходимости к предельному распределению приблизительно одинакова в обоих случаях.

3. Асимптотическое распределение статистики R_λ

Без ограничения общности будем считать, что $F_1(t) = t$, $F_2(t) = t^{1/k}$, $0 \leq t \leq 1$. В работе [1] для асимптотической ковариации процесса $Z_n(t) = \sqrt{n_1} \left(\hat{F}_{t_1}(t) - \left(\hat{F}_{t_2}(t) \right)^k \right)$, $0 < \Delta \leq t \leq 1$, который определяет статистику (3), была доказана следующая теорема. Обозначим $K_n(t_1, t_2)$ функцию ковариации $Z_n(t)$.

Теорема 2. Пусть $0 < \Delta \leq t_1 \leq t_2 \leq 1$, тогда $K_n(t_1, t_2) \xrightarrow[n_2 \rightarrow \infty]{n_1 \rightarrow \infty} K(t_1, t_2) = K_1(t_1, t_2) + \rho K_2(t_1, t_2) =$

$$= t_1 t_2 \left(\frac{1 - t_2^{m_1}}{m_1^2 t_2^{m_1}} + \frac{\rho k^2 \left(1 - t_2^{\frac{m_2}{k}} \right)}{m_2^2 t_2^{\frac{m_2}{k}}} \right) = t_1 \left(\frac{1 - t_2^{m_1}}{m_1^2 t_2^{m_1-1}} + \frac{\rho k^2 \left(1 - t_2^{\frac{m_2}{k}} \right)}{m_2^2 t_2^{\frac{m_2}{k}-1}} \right).$$

Введём процесс $Y_n(t) = \frac{m_1 m_2}{\sqrt{k^2 \rho m_1^2 + m_2^2}} Z_n(t)$. При стандартных ограничениях $0 < \Delta \leq t_1 \leq t_2 \leq 1$ процесс $Y_n(t)$ сходится к гауссовскому процессу $Y(t)$ с нулевым математическим ожиданием и ковариацией

$$E[Y(t_1) \cdot Y(t_2)] = \frac{m_1^2 m_2^2}{k^2 \rho m_1^2 + m_2^2} t_1 \cdot \left(\frac{1 - t_2^{m_1}}{m_1^2 t_2^{m_1-1}} + \frac{\rho k^2 \left(1 - t_2^{\frac{m_2}{k}} \right)}{m_2^2 t_2^{\frac{m_2}{k}-1}} \right) = t_1 \cdot \frac{k_2 t_2^{\frac{m_2}{k} - m_1} + k_1 - t_2^{\frac{m_2}{k}}}{t_2^{\frac{m_2}{k}-1}} =$$

$$= t_1 \cdot \frac{k_2 t_2^{\frac{m_2}{k}+1} + k_1 t_2^{m_1+1} - t_2^{\frac{m_2}{k}+m_1+1}}{t_2^{\frac{m_2}{k}+m_1}}, \quad 0 < \Delta \leq t_1 \leq t_2 \leq 1.$$

Теорема 3. При $n_1 \rightarrow \infty, n_2 \rightarrow \infty, \frac{n_1}{n_2} \rightarrow \rho$ и справедливости гипотезы (1) распределение статистики (3) сходится к стандартному распределению Реньи

$$L(h) = \frac{4}{\pi} \sum_{i=0}^{\infty} \frac{(-1)^i}{2i+1} \exp \left\{ -\frac{(2i+1)^2 \pi^2}{8h^2} \right\}.$$

Доказательство. Рассмотрим преобразование времени

$$\tau(t) = \frac{t^{\frac{m_2}{k} + m_1}}{k_2 t^{\frac{m_2}{k}} + k_1 t^{m_1} - t^{\frac{m_2}{k} + m_1}}, t \neq 0, t \neq 1.$$

Тогда $\tau'(t) > 0$. Так как $\lim_{t \rightarrow 0} \tau(t) = 0, \lim_{t \rightarrow \infty} \tau(t) = 1$, то доопределяя $\tau(0) = 0$ получим строго возрастающее отображение $[0; 1) \rightarrow [0; \infty)$. Тогда существует обратное преобразование $t = t(\tau)$.

Пусть $\tau_1 \leq \tau_2 \Leftrightarrow t(\tau_1) = t_1 \leq t_2 = t(\tau_2)$.

Введем в рассмотрение процесс

$$\begin{aligned} W(\tau) &= Y(t(\tau)) \cdot \frac{(t(\tau))^{\frac{m_2}{k} + m_1 - 1}}{k_2 (t(\tau))^{\frac{m_2}{k}} + k_1 (t(\tau))^{m_1} - (t(\tau))^{\frac{m_2}{k} + m_1}}. \\ E[W(\tau)] &= 0, \\ E[W(\tau_1) \cdot W(\tau_2)] &= \frac{t_1^{\frac{m_2}{k} + m_1 - 1}}{k_2 t_1^{\frac{m_2}{k}} + k_1 t_1^{m_1} - t_1^{\frac{m_2}{k} + m_1}} \cdot \frac{t_2^{\frac{m_2}{k} + m_1 - 1}}{k_2 t_2^{\frac{m_2}{k}} + k_1 t_2^{m_1} - t_2^{\frac{m_2}{k} + m_1}} \cdot E[Y(t_1)Y(t_2)] = \\ &= \frac{t_1^{\frac{m_2}{k} + m_1 - 1}}{k_2 t_1^{\frac{m_2}{k}} + k_1 t_1^{m_1} - t_1^{\frac{m_2}{k} + m_1}} \cdot \frac{t_2^{\frac{m_2}{k} + m_1 - 1}}{k_2 t_2^{\frac{m_2}{k}} + k_1 t_2^{m_1} - t_2^{\frac{m_2}{k} + m_1}} \cdot t_1 \cdot \frac{k_2 t_2^{\frac{m_2}{k} + 1} + k_1 t_2^{m_1 + 1} - t_2^{\frac{m_2}{k} + m_1 + 1}}{t_2^{\frac{m_2}{k} + m_1}} = \\ &= \frac{t_1^{\frac{m_2}{k} + m_1}}{k_2 t_1^{\frac{m_2}{k}} + k_1 t_1^{m_1} - t_1^{\frac{m_2}{k} + m_1}} = \tau_1. \end{aligned}$$

Таким образом, процесс $W(\tau)$ является винеровским процессом.

Рассмотрим неравенство $\psi(t) < 1 - \lambda$, задающее область поиска максимума в статистике (3)

$$\psi(t(\tau)) = \frac{t^{\frac{m_2}{k}}}{k_2 t^{\frac{m_2}{k} - m_1} + k_1} = \frac{t^{\frac{m_2}{k} + m_1}}{k_2 t^{\frac{m_2}{k}} + k_1 t^{m_1}} = \frac{t^{\frac{m_2}{k} + m_1}}{k_2 t^{\frac{m_2}{k}} + k_1 t^{m_1} - t^{\frac{m_2}{k} + m_1}} \cdot \frac{k_2 t^{\frac{m_2}{k}} + k_1 t^{m_1} - t^{\frac{m_2}{k} + m_1}}{k_2 t^{\frac{m_2}{k}} + k_1 t^{m_1}} = \frac{\tau}{1 + \tau} < 1 - \lambda.$$

Решив неравенство относительно τ , получим $\tau < \frac{1 - \lambda}{\lambda}$.

В асимптотике $\hat{F}_0(t)$ — объединенную эмпирическую оценку функции распределения элементов по двум полным выборкам Θ_1, Θ_2 можно заменить [11] на t . Тогда, учитывая, что для винеровского процесса $W(t)$ справедливо $\sqrt{C} W(t) = W(Ct)$ [12] в пределе при $n_1 \rightarrow \infty, n_2 \rightarrow \infty$ получим

$$\begin{aligned} P(R_\lambda < h) &= P\left(\sqrt{\frac{\lambda}{1 - \lambda}} \max_{\tau < \frac{1 - \lambda}{\lambda}} |W(\tau)| < h\right) = P\left(\max_{\tau < \frac{1 - \lambda}{\lambda}} \left|W\left(\frac{\lambda}{1 - \lambda} \tau\right)\right| < h\right) = \\ &= P\left(\max_{u < 1} |W(u)| < h\right) = L(h), \text{ где } u = \frac{\lambda}{1 - \lambda} \tau. \end{aligned}$$

Доказанная теорема позволяет проверять гипотезы при достаточно больших объемах выборок n_1, n_2 . Но учитывая, что скорость сходимости распределений статистик типа Реньи медленная [13], на практике лучшим является использование точных распределений статистики R_λ .

4. Оценка степенного параметра

Аналогично работе [1] рассмотрим возможность нахождения оценки параметра k в модели Лемана минимизацией статистики (3). Пусть испытания проводились до момента времени τ , когда отказала лишь часть систем, находящихся на эксплуатации. В качестве оценки степенного параметра k будем использовать значение, которое даёт минимум статистики (3), т.е. $\hat{k} = \arg \min R_\lambda$. Рассмотрим алгоритм моделирования оценки параметра модели Лемана.

1. Моделируются $n_1 m_1$ одинаково распределенных случайных величин $(\xi_1^1, \dots, \xi_{m_1 n_1}^1)$ с функцией распределения $G_0(t)$.

2. Нарботки случайным образом разбиваются на n_i групп по m_i величин в каждой. Элементы i -й группы обозначим $(\xi_1^{1,i}, \dots, \xi_{m_i}^{1,i}), i = \overline{1, n_1}$. Определяются наработки до отказа систем $\theta_1^i = \max(\xi_1^{1,i}, \dots, \xi_{m_i}^{1,i})$.

3. Аналогичным образом моделируются $n_2 m_2$ одинаково распределенных случайных величин $(\xi_1^2, \dots, \xi_{m_2 n_2}^2)$ с функцией распределения $(G_0(t))^{\frac{1}{k}}$, где k — некоторое заданное значение степенного параметра. Нарботки $(\xi_1^2, \dots, \xi_{m_2 n_2}^2)$ случайным образом разбиваются на n_2 групп по m_2 величин в каждой. Элементы i -й группы обозначим $(\xi_1^{2,i}, \dots, \xi_{m_2}^{2,i}), i = \overline{1, n_2}$. Определяются наработки до отказа систем во втором режиме работы $\theta_2^i = \max(\xi_1^{2,i}, \dots, \xi_{m_2}^{2,i})$.

4. Задается глубина цензурирования испытаний Λ , $0 < \Lambda < 1$, равная доле числа отказавших систем в обоих режимах от общего количества систем. Определяется число r такое, что $\frac{r}{n_1 + n_2} \leq \Lambda < \frac{r+1}{n_1 + n_2}$.

Наработки $(\theta_1^1, \dots, \theta_1^{n_1}), (\theta_2^1, \dots, \theta_2^{n_2})$ располагаются в порядке возрастания. Пусть $\gamma_1 < \gamma_2 < \dots < \gamma_{n_1+n_2}$ — объединенный вариационный ряд из этих наработок. По вариационному ряду и числу r определяется наработка $\gamma_r = \tau$, а также r_1 и r_2 — число отказов систем в режимах ε_1 и ε_2 до момента γ_r соответственно. Определяются две выборки $\Theta_1 = (\theta_1^1, \dots, \theta_1^{r_1}), \Theta_2 = (\theta_2^1, \dots, \theta_2^{r_2})$ из наблюдаемых до момента γ_r наработок.

5. Для заданного значения $\tilde{k}$, $1 \leq \tilde{k} \leq K$ вычисляется текущая глубина цензурирования

$$\tilde{\Lambda}_k = \left(\frac{r_1}{n_1} \right)^{\frac{1}{m_1}} \cdot \frac{n_1}{n_1 + n_2} + \left(\frac{r_2}{n_2} \right)^{\frac{1}{m_2}} \cdot \frac{n_2}{n_1 + n_2} = \hat{F}_0(\gamma_r).$$

Параметр Реньи в статистике (3) полагается равным глубине цензурирования $\lambda(\tilde{k}) = 1 - \hat{F}_0(\gamma_r)$. Для этого $\lambda(\tilde{k})$ вычисляется значение статистики (3).

6. Определяется оценка $\hat{k}$, которая даёт минимальное значение статистики $R_\lambda(\tilde{k})$, т.е.

$$\hat{k} = \arg \min_{1 \leq \tilde{k} \leq K} R_\lambda(\tilde{k}).$$

В качестве примера для расчёта в качестве $G_0(t)$ использовались экспоненциальное распределение с параметром $\beta = 0.001$ и распределение Вейбулла с параметрами $\beta = 0.001, p = 1.5$. На рис. 2 изображены гистограммы оценок $\hat{k}$ для $m_1 = 2, m_2 = 3, n_1 = n_2 = 100$. Количество испытаний $N = 500$, действительное значение параметра Лемана $k = 3$.

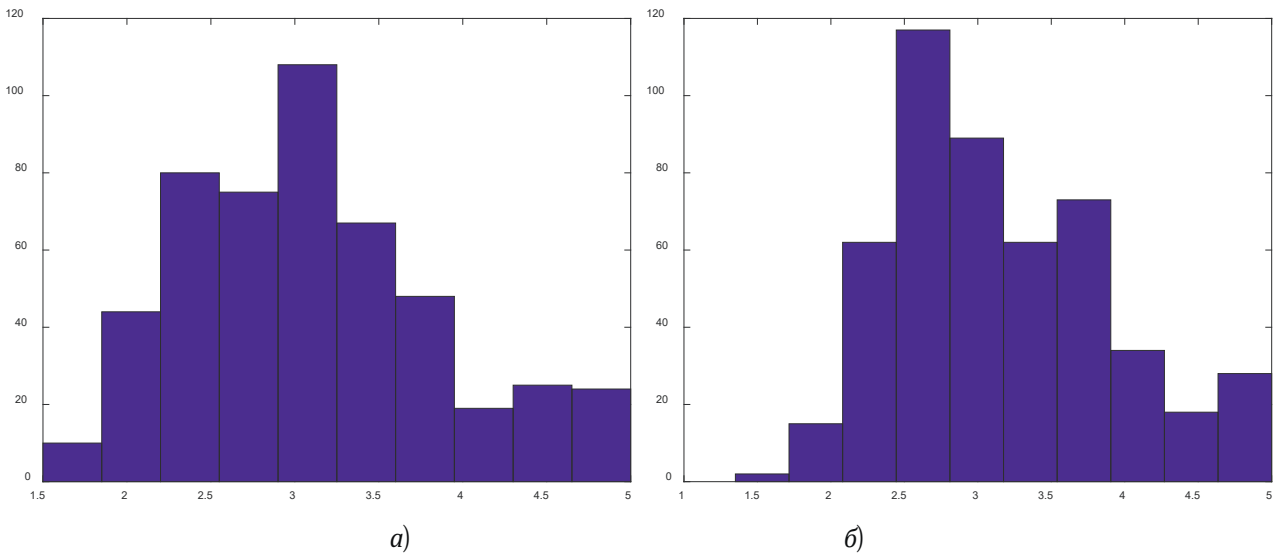


Рис. 2. Гистограммы оценок $\hat{k}$ при $\Lambda = 0,8$: Экспоненциальное распределение $M\hat{k} = 3.1198, \sigma = 0.7745$ (а), $M\hat{k} = 3.1522, \sigma = 0.7670$ (б).

Заключение

В работе представлена новая статистика типа Реньи для проверки степенной гипотезы Лемана для функций распределения наработок до отказа элементов при испытаниях, составленных из них систем, ограниченных во времени. Получены асимптотическое и точные распределения статистики типа Реньи для случая проверки основной гипотезы. Показано, что асимптотическое распределение статистики предложенного авторами критерия совпадает с предельным распределением Реньи. Проведено статистическое моделирование с целью анализа точности минимаксной оценки степенного параметра гипотезы Лемана.

Список литературы

1. Тимонин В.И., Тянникова Н.Д. Методы решения задачи непараметрической проверки гипотез Лемана при испытаниях параллельных систем // Математическое моделирование и численные методы. 2018. № 1. С. 98–112.
2. Тимонин В.И., Тянникова Н.Д. Проверка однородности распределений наработок до отказа элементов в двух различных режимах по результатам испытаний резервированных систем, составленных из этих элементов // Физические основы приборостроения. 2018. Т. 7. № 3(29). С. 16–25.
3. Гнеденко Б.В., Беляев Ю.К., Соловьев А.Д. Математические методы в теории надежности. Основные характеристики надежности и их статистический анализ. М.: Либроком, 2013. 584 с.
4. Kaplan, E.L., Meier, P. Nonparametric estimation from incomplete observations // JAm Stat Assoc. 1958. Vol. 53. Iss. 282. P. 457–481. DOI: 10.1080/01621459.1958.10501452
5. 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. P. 2295–2311. DOI: 10.1016/j.jspi.2010.01.025
6. Bagdonavichus, V., Kruopis, J., and Nikulin, M.S. Nonparametric tests for censored data. London: ISTE Ltd, 2011. 233 p.
7. Balakrishnan, N., Cramer, E. The Art of Progressive Censoring. Applications to Reliability and Quality. New York: Springer, 2014. 645 p.
8. Тимонин В.И., Тянникова Н.Д. Метод вычисления точных распределений статистик типа Колмогорова-Смирнова в случае нарушения однородности и независимости анализируемых выборок // Электронное научно-техническое издание «Наука и образование». 2014. № 11. С. 217–227.
9. 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. С. 68–84.
10. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. М.: Наука, 1983. 416 с.
11. Hajek, J., Sidak, Z. Theory of rank tests. London: Academic Press, 2004. 438 p.
12. Дуб Дж.Л. Вероятностные процессы. М.: Издательство иностранной литературы, 1956. 607 с.
13. Лемешко Б.Ю., Чимитова Е.В. О распределениях статистик критерия Реньи // Мат. VI международной конференции «Актуальные проблемы электронного приборостроения». АПЭП-2002. Новосибирск, 2002. Т. 6. С. 39–42.

Поступила 16 июля 2024 г.

TESTING LEHMAN'S HYPOTHESIS WITH INCOMPLETE DATA FOR PARALLEL SYSTEMS

TIMONIN V.I., TYANNIKOVA N.D.

Bauman Moscow State Technical University

doi: 10.25210/jfop-2403-QGSOUB | edn: QGSOUB

Let there be a system consisting of parallel connected elements of the same type (hot reserve). Based on tests of two types of systems of different multiplicities in different modes, two samples from the time to failure of the systems were obtained. The trials are time limited, so both samples are censored. The problem of checking the Lehmann power law dependence of the functions of distributions of time to failure of elements is solved using data of this kind. Estimates of the distribution function of time to failure of elements are determined on the basis of analogues of Kaplan-Meier estimates for reliability functions introduced by the authors in [1,2]. A Renyi-type criterion based on a comparison of distribution function estimates is considered. In a particular case, with unit multiplicity of systems and the power parameter in Lehman's hypothesis $k = 1$, the criterion coincides with the classical Renyi criterion. It is shown that the limiting distribution of the test statistics does not depend on the value of the Lehman model parameter. The method for calculating exact distributions of criterion statistics is a modification of the general method for calculating exact distributions of Kolmogorov-Smirnov type statistics for heterogeneous samples, introduced by the authors in [8]. The values of the exact quantiles of the statistics distribution function for finite sample sizes are calculated. For the asymptotic distribution of the criterion statistics, it is shown that the distribution of statistics reduces to the distribution of the maximum modulus of the Wiener process. The consequence of this is the convergence of exact distributions of statistics to the classical Renyi distribution. The question of the minimum sample sizes at which it is possible to use the limiting probabilities of the Renyi distribution is investigated. Based on the introduced criterion, a minimax estimate of the Lehmann parameter is proposed in the case where the parameter is unknown. A study of the statistical properties of this estimate using Monte Carlo methods was carried out. It is shown that with average sample sizes, the estimate gives a very slight bias with moderate scatter. Simulations were carried out for the exponential distribution and for the Weibull distribution.