ISSN 2309-3684

Математическое моделирование и численные методы

УДК 519.248 DOI 10.18698/2309-3684-2017-1-102117

Проверка справедливости модели Кокса по нескольким прогрессивно цензурированным выборкам

© В.И. Тимонин, Н.Д. Тянникова

МГТУ им. Н.Э. Баумана, Москва, 105005, Россия

Предложен непараметрический критерий типа Кифера — Гихмана для проверки справедливости модели Кокса по нескольким прогрессивно цензурированным выборкам. По каждой из выборок в качестве оценок функции надежности использованы оценки Каплана — Мейера. Доказано, что при справедливости гипотезы в качестве приближения асимптотического распределения статистики критерия может быть применено распределение Кифера — Гихмана. Разработан метод вычисления точных распределений статистики на основе модели случайного блуждания частицы по многомерному массиву ячеек. Приведены таблицы полученных значений вероятностей точных распределений предложенной статистики для широкого набора возможных значений объемов выборок. Методами статистического моделирования показана состоятельность метода оценки параметров Кокса, основанного на минимизации статистики. Представлены гистограммы полученных оценок для экспоненциального распределения наработок до отказа. Результаты исследования находят применение при анализе результатов испытаний резервированных технических систем различной кратности, функционирующих в различных условиях эксплуатации. Анализируемые системы используются во всех отраслях — от машиностроительных до радиоэлектронных.

Ключевые слова: *непараметрическая статистика, критерий типа Кифера — Гихмана, оценка Каплана — Мейера, прогрессивное цензурирование, модель Кокса*

Введение. Во многих задачах теории надежности [1–3] требуется проверить наличие степенных зависимостей между функциями надежности наработок до отказа изделий для нескольких режимов. Обозначим функцию надежности i-й выборки $P_i(t) = 1 - F_i(t)$, $i = \overline{1, q}$, где $F_i(t)$ — функция распределения i-й выборки. Тогда проверяемая гипотеза имеет вид

$$H_0 : P_1^{k_1}(t) = P_2^{k_2}(t) = ... = P_q^{k_q}(t) = P_0(t), \qquad (1)$$

где t — время; q — число режимов работы.

Без ограничения общности можно считать, что $k_1 = 1$, $k_1 \le k_2 \le ... \le k_q$. В случае $k_i = 1$, $\forall i$ гипотеза является классической гипотезой однородности, для проверки которой используют критерий Кифера — Гихмана [4, 5]. Его статистика имеет вид

$$T = \max_i \sum_{i=1}^{q} n_i \left(\widehat{F}_i - \overline{F} \right)^2,$$

где n_i — объем i-й выборки, $n = \sum_{i=1}^{q} n_i$; $\widehat{F}_i$ — эмпирическая функция распределения i-й выборки; $i = \overline{1, q}$; $\overline{F} = \frac{1}{n} \sum_{i=1}^{q} n_i \widehat{F}_i(t)$ — объединенная эмпирическая функция распределения.

В работе [6] рассмотрено обобщение критерия Кифера — Гихмана на случай проверки справедливости модели Кокса (1). Однако в силу конструктивных особенностей изделий зачастую отказ одного элемента системы может приводить к снятию с эксплуатации всей системы. Это вызывает появление прогрессивно цензурированных данных, для анализа которых необходимы специальные методы обработки. В работе [7] предложен многовыборочный критерий однородности типа Кифера — Гихмана для нескольких прогрессивно цензурированных выборок. В данной статье проведено обобщение этого критерия на случай проверки справедливости модели Кокса.

Постановка задачи. Имеется q режимов работы ε_i, $i = \overline{1, q}$. Во всех режимах ε_i испытывается n_i систем, каждая из которых состоит из m_i последовательно соединенных элементов. Ввиду конструктивных особенностей изделий в результате испытаний можно наблюдать из наработок до отказа элементов систем только минимумы θ_i^j, $i = \overline{1, q}$, $j = \overline{1, n_i}$, а оставшиеся $(m_i - 1)$ цензурируются. В итоге имеем q прогрессивно цензурированных выборок $\vec{\Theta}_1 = \left(\theta_1^1, ..., \theta_1^{n_1} \right)$, $\vec{\Theta}_2 = \left(\theta_2^1, ..., \theta_1^{n_2} \right)$, ..., $\vec{\Theta}_q = \left(\theta_q^1, ..., \theta_q^{n_q} \right)$, составленных из отказов систем [8, 9]. По результатам таких испытаний необходимо проверить гипотезу о том, что интенсивности отказов $\lambda_1(t), \lambda_2(t), ..., \lambda_q(t)$ элементов для q режимов пропорциональны, т. е. $k_1 \lambda_1(t) = k_2 \lambda_2(t) = ... = k_q \lambda_q(t)$, где $k_i \geq 1, i = \overline{1, q}$ — известные фиксированные числа. Тогда проверяемая гипотеза также имеет вид (1). Далее без ограничения общности будем считать, что $k_1 = 1, k_1 \leq k_2 \leq ... \leq k_q$.

Функции надежности элементов по каждой из выборок $\vec{\Theta}_i$ можно оценить с помощью оценок Каплана — Мейера [10, 11]:

$$\widehat{P}_i(t)=\begin{cases} 1,\ d_i(t)=0, \\ \prod\limits_{j=1}^{d_i(t)}\left(1-\dfrac{1}{m_i(n_i-j+1)}\right),\ 1\le d_i(t)\le(n_i-1), \\ 0,\ d_i(t)=n_i, \end{cases}$$

где $d_i(t)$ — количество элементов выборок $\vec{\Theta}_i$, меньших t.

Если справедлива гипотеза (1), то $\widehat{P}_i^{k_i}(t)$ оценивают одну и ту же функцию надежности $P_0(t)$. Заметим, что выборки $\left(\theta_1^1,...,\theta_1^{n_1}\right)$,..., $\left(\theta_q^1,...,\theta_q^{n_q}\right)$ можно рассматривать как полные (нецензурированные) независимые выборки из отказов систем с функциями распределения $G^1,G^2,...,G^q$, где $G^i=1-(1-F_i)^{m_i}$. Их можно оценить эмпирическими функциями распределения $\widehat{G}^1,\widehat{G}^2,...,\widehat{G}^q$, $\widehat{G}^i=\left(d_i(t)\right)/n_i, i=\overline{1,q}$.

Для проверки гипотезы (1) предлагается использовать критерий типа Кифера — Гихмана, статистика которого имеет вид

$$\tilde{T}^2=\max_t\frac{\sum\limits_{i=1}^{q}n_i\left(\widehat{P}_i^{k_i}-\overline{P}\right)^2+\tilde{P}\left(\sqrt{\Phi(q-1)}-\Phi_1\right)}{\Gamma^2}. \tag{2}$$

Здесь $\overline{P}=\sum\limits_{i=1}^{q}\rho_i\widehat{P}_i(t)^{k_i}$ (где $\rho_i=n_i/n$, $n=\sum\limits_{i=1}^{q}n_i$); $\tilde{P}=\sum\limits_{i=1}^{q}\rho_i\left(1-\widehat{G}^i\right)^{k_i/m_i}$; $\Phi=\left(\sum\limits_{i=1}^{q}\rho_iS_i\right)^2+\sum\limits_{i=1}^{q}(1-2\rho_i)S_i^2$ (где $S_i=\dfrac{1-\tilde{P}^{m_i/k_i}}{(m_i/k_i)^2\,\tilde{P}^{m_i/k_i-1}}$); $\Phi_1=$ $=\sum\limits_{i=1}^{q}(1-\rho_i)S_i$; $\Gamma=\tilde{P}+\sqrt{\dfrac{\Phi}{q-1}}$.

Если $d_i(t)=n_i$, $i=\overline{1,\,q}$, то положим

$$\frac{\sum\limits_{i=1}^{q}n_i\left(\widehat{P}_i^{k_i}-\overline{P}\right)^2+\tilde{P}\left(\sqrt{\Phi(q-1)}-\Phi_1\right)}{\Gamma^2}=0.$$

Приближение асимптотического распределения $\tilde{T}^2$. Без ограничения общности можно считать, что функции надежности каждой из выборок имеют вид $P_i(t)=(1-t)^{1/k_i}$, $i=\overline{1,\,q}$, $0\le t\le 1$. Для вывода приближения асимптотического распределения статистики рассмотрим эмпирический процесс $y_i(t)=\sqrt{n_i}\left(\hat{P}_i^{k_i}(t)-(1-t)\right)$. В работе [11] показано, что асимптотическая ковариация процесса $y_i(t)$ имеет вид

$$K_i(s,t)\xrightarrow[n_i\to\infty]{}(1-t)(1-s)\left(\frac{k_i^2\left(1-(1-s)^{m_i/k_i}\right)}{m_i^2(1-s)^{m_i/k_i}}\right),\ 0\le s\le t\le\Delta<1. \quad (3)$$

Рассмотрим эмпирический случайный процесс $Y(t)= =\sum_{i=1}^{q}n_i\left(\hat{P}_i^{k_i}-\bar{P}\right)^2=\sum_{i=1}^{q}\left(y_i(t)-\sqrt{\rho_i}\sum_{j=1}^{q}\sqrt{\rho_j}y_j(t)\right)^2$, определяющий статистику (2). Учитывая (3), легко показать, что эмпирический случайный процесс $Y(t)$ слабо сходится при $n_i\to\infty,\ n_i/n\to\rho_i>0$ к распределению процесса

$$Y^1(t)=\sum_{i=1}^{q}\left(y_i(t)-\sqrt{\rho_i}\sum_{j=1}^{q}\sqrt{\rho_j}y_j(t)\right)^2,$$

где $y_i(t)$ — нормальный случайный процесс с характеристиками

$$E(y_i(t))=0,\quad K_i(s,t)=Ey_i(s)y_i(t)=(1-t)(1-s)\left(\frac{k_i^2\left(1-(1-s)^{m_i/k_i}\right)}{m_i^2(1-s)^{m_i/k_i}}\right),$$

$0\le s\le t\le\Delta<1$.

Лемма 1. При $n_i\to\infty,\ n_i/n\to\rho_i>0,\ 0\le t\le\Delta<1\ \bar{P}\xrightarrow{\text{п.н.}}1-t$; при $0\le t\le1\ \tilde{P}\xrightarrow{\text{п.н.}}1-t$; $\Phi_1\xrightarrow{\text{п.н.}}\Phi_{1,\infty}(t)=\sum_{i=1}^{q}(1-\rho_i)\times$

$$\times\frac{k_i^2\left(1-(1-t)^{m_i/k_i}\right)}{m_i^2(1-t)^{m_i/k_i-1}};\ \Phi\xrightarrow{\text{п.н.}}\Phi_\infty(t)=\left(\sum_{i=1}^{q}\rho_i\frac{k_i^2\left(1-(1-t)^{m_i/k_i}\right)}{m_i^2(1-t)^{m_i/k_i-1}}\right)^2+$$

$$+\sum_{i=1}^{q}(1-2\rho_i)\left(\frac{k_i^2\left(1-(1-t)^{m_i/k_i}\right)}{m_i^2(1-t)^{m_i/k_i-1}}\right)^2;\ \Gamma\xrightarrow{\text{п.н.}}\Gamma_\infty(t)=(1-t)+\frac{\Phi_\infty(t)}{\sqrt{q-1}}.$$

Здесь символ «п. н.» означает сходимость с вероятностью 1, или «почти наверное».

Лемма 1 следует из теоремы Гливенко [1] и работы [12].◄

Перейдем к выводу математического ожидания и ковариации случайного процесса $Y^1(t)$.

Теорема 1. Математическое ожидание процесса $Y^1(t)$ имеет вид

$$EY^1(t)=(1-t)\sum_{i=1}^{q}(1-\rho_i)\frac{k_i^2\left(1-(1-t)^{m_i/k_i}\right)}{m_i^2(1-t)^{m_i/k_i-1}}=(1-t)\Phi_{1,\infty}(t),\ \ 0\le t\le\Delta<1.$$

Доказательство. Обозначим $\mu(t)=\sum_{j=1}^{q}\sqrt{\rho_j}\,y_j(t).$ Так как процессы $y_i(t)$ независимые, имеем

$$EY^1(t)=E\left(\sum_{i=1}^{q}\left(y_i(t)-\sqrt{\rho_i}\sum_{j=1}^{q}\sqrt{\rho_j}\,y_j(t)\right)^2\right)=$$

$$=E\left(\sum_{i=1}^{q}\left(y_i(t)-\sqrt{\rho_i}\,\mu(t)\right)^2\right)=$$

$$=E\sum_{i=1}^{q}\left(y_i^2(t)-2\sqrt{\rho_i}\,y_i(t)\mu(t)+\rho_i\mu^2(t)\right)=E\left(\sum_{i=1}^{q}y_i^2(t)-\mu^2(t)\right)=$$

$$=\sum_{i=1}^{q}(1-\rho_i)\frac{k_i^2\left(1-(1-t)^{m_i/k_i}\right)}{m_i^2(1-t)^{m_i/k_i-2}}=(1-t)\Phi_{1,\infty}(t).\ ◄$$

Аналогично работе [7], воспользуемся свойствами двумерного нормального распределения. Если (ξ,η) — двумерный нормальный вектор с нулевым математическим ожиданием и матрицей ковариаций $D=\begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{pmatrix}$, то $E\xi^2\eta^2=\sigma_{11}\sigma_{22}+2\sigma_{12}^2.$

Теорема 2. При $0\le s\le t\le\Delta<1$ ковариация эмпирического случайного процесса $Y^1(t)$ имеет вид

$$K(s,t)=2(1-t)^2\left[\left(\sum_{i=1}^{q}\rho_i\frac{k_i^2\left(1-(1-s)^{m_i/k_i}\right)}{m_i^2(1-s)^{m_i/k_i-1}}\right)^2+\right.$$

$$+\sum_{i=1}^{q}\left(1-2\rho_i\right)\left(\frac{k_i^2\left(1-\left(1-s\right)^{m_i/k_i}\right)}{m_i^2\left(1-s\right)^{m_i/k_i-1}}\right)^2\Bigg] = 2(1-t)^2\Phi_\infty\left(s\right).$$

Доказательство. Получим выражение для ковариации эмпирического случайного процесса $Y^1\left(t\right)$. В соответствии с результатами, приведенными в работе [7], имеем

$$E\left(Y^1\left(s\right)Y^1\left(t\right)\right)=$$

$$=E\left(\sum_{i=1}^{q}\sum_{j=1}^{q}y_i^2\left(s\right)y_j^2\left(t\right)-\sum_{i=1}^{q}y_i^2\left(s\right)\mu^2\left(t\right)-\mu^2\left(s\right)\sum_{i=1}^{q}y_i^2\left(t\right)+\mu^2\left(s\right)\mu^2\left(t\right)\right).$$

Для первого слагаемого получим

$$E\sum_{i=1}^{q}\sum_{j=1}^{q}y_i^2\left(s\right)y_j^2\left(t\right)=\sum_{i=1}^{q}E\left(y_i^2\left(s\right)y_i^2\left(t\right)\right)+\sum_{i\neq j}^{q}\sum_{j=1}^{q}E\left(y_i^2\left(s\right)y_j^2\left(t\right)\right)=$$

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

$$\left.+\,2\left(1-t\right)^2\left(\frac{k_i^2\left(1-\left(1-s\right)^{m_i/k_i}\right)}{m_i^2\left(1-s\right)^{m_i/k_i-1}}\right)^2\right)+$$

$$+\sum_{i\neq j}^{q}\sum_{j=1}^{q}\frac{k_i^2\left(1-\left(1-s\right)^{m_i/k_i}\right)}{m_i^2\left(1-s\right)^{m_i/k_i-2}}\frac{k_j^2\left(1-\left(1-t\right)^{m_i/k_i}\right)}{m_j^2\left(1-t\right)^{m_i/k_i-2}}=$$

$$=2\sum_{i=1}^{q}\frac{\left(1-t\right)^2k_i^4\left(1-\left(1-s\right)^{m_i/k_i}\right)^2}{m_i^4\left(1-s\right)^{2\left(m_i/k_i\right)-2}}+$$

$$+\sum_{i=1}^{q}\sum_{j=1}^{q}\left(\frac{k_i^2\left(1-\left(1-s\right)^{m_i/k_i}\right)}{m_i^2\left(1-s\right)^{m_i/k_i-2}}\frac{k_j^2\left(1-\left(1-t\right)^{m_i/k_i}\right)}{m_j^2\left(1-t\right)^{m_i/k_i-2}}\right).$$

Применив аналогичные рассуждения для оставшихся слагаемых и суммировав полученные результаты, получим утверждение теоремы.◄

Преобразуем процесс $Y^1(t)$ так, чтобы при некотором преобразовании времени $u(t):[0,1]\to[0,1]$ его математическое ожидание и ковариация совпадали с математическим ожиданием и функцией ковариации процесса $Z(t)=\sum_{i=1}^{q} n_i\left(\widehat{F}_i(t)-\overline{F}(t)\right)^2$, $t\in[0,1]$. Этот процесс используется при проверке гипотезы однородности q полных выборок, и его математическое ожидание и функция ковариаций равны

$$EZ(t)=(q-1)t(1-t),$$

$$EZ(t_1)Z(t_2)=2(q-1)t_1^2(1-t_2)^2,\ 0\le t_1\le t_2\le 1.$$

Для этого введем процесс

$$Y^*(t)=\frac{Y^1(t)}{h(t)},$$

где $h(t)>0,\ h(0)=1$ — некоторая функция.

Очевидно, что $EY^*(t)=\left(1/h(t)\right)EY^1(t)$, $EY^*(s)Y^*(t)=$ $=\left(1/\left(h(s)h(t)\right)\right)EY^1(s)Y^1(t)$. Учитывая полученную функцию ковариации процесса $Y^1(t)$, имеем $\operatorname{cov}\left(Y^*(s),\,Y^*(t)\right)=\dfrac{2(1-t)^2}{h(t)}\dfrac{\Phi_\infty(s)}{h(s)}$.

Положим $h(t)=\left(1-t+\sqrt{\dfrac{\Phi_\infty(t)}{q-1}}\right)^2$. Рассмотрим функцию $u(t):[0,1]\to[0,1]$, $u=1-\dfrac{1-t}{\sqrt{h(t)}}$, $t\ne 1$, $u(1)=1$. Легко показать, что $u(t)$ монотонно возрастает. Имеем $(1-u)^2=\dfrac{(1-t)^2}{h(t)}$. Заметим, что $\Gamma_\infty(t)=\sqrt{h(t)}$.

Тогда функцию ковариации процесса $W(u)=Y^*(t(u))=\dfrac{Y^1(t(u))}{\Gamma_\infty^2(t(u))}$ запишем так:

$$EW(u_1)W(u_2) = E\frac{Y^1(t(u_1))Y^1(t(u_2))}{\Gamma_\infty^2(t(u_1))\Gamma_\infty^2(t(u_2))} - \frac{EY^1(t(u_1))EY^1(t(u_2))}{\Gamma_\infty^2(t(u_1))\Gamma_\infty^2(t(u_2))} =$$

$$= 2(q-1)u_1^2(1-u_2)^2, \text{ при } u_1 < u_2.$$

Здесь $t(u)$ — обратная функция к $u(t)$.

Исходя из того, что при изменении математического ожидания случайного процесса ковариация не меняется, преобразуем случайный процесс $W(u)$ таким образом, чтобы получить математическое ожидание, которое тоже будет совпадать с математическим ожиданием процесса $Z(t)$.

Введем новый процесс

$$W^*(u) = \frac{Y^1(t(u)) + (1-t(u))\left[\sqrt{(q-1)\Phi_\infty(t(u))} - \Phi_{1,\infty}(t(u))\right]}{\Gamma_\infty^2(t(u))}.$$

Учитывая, что

$$EY^*(t) = \frac{(1-t)\Phi_{1,\infty}(t)}{\Gamma_\infty^2(t)} = (1-u)\frac{\sqrt{\Phi_\infty(t)}}{\Gamma_\infty(t)}\sqrt{q-1}\frac{\Phi_{1,\infty}(t)}{\sqrt{\Phi_\infty(t)}}\frac{1}{\sqrt{q-1}} =$$

$$= u(1-u)(q-1)\left[1 + \left(\frac{\Phi_{1,\infty}(t)}{\sqrt{\Phi_\infty(t)}}\frac{1}{\sqrt{q-1}} - 1\right)\right] =$$

$$= u(1-u)(q-1) + \frac{u(1-u)(q-1)\left(\Phi_{1,\infty}(t) - \sqrt{(q-1)\Phi_\infty(t)}\right)}{\sqrt{(q-1)\Phi_\infty(t)}} =$$

$$= u(1-u)(q-1) - \frac{1-t}{\Gamma^2(t)}\left(\sqrt{\Phi_\infty(t)(q-1)} - \Phi_{1,\infty}(t)\right),$$

получим требуемое математическое ожидание $EW^*(u) = (q-1)u(1-u)$.

В результате математическое ожидание и ковариация процесса $W^*(u) = \left(Y^1(s(u)) + (1-t(u))\left[\sqrt{(q-1)\Phi_\infty(t(u))} - \Phi_{1,\infty}(t(u))\right]\right) / \left(\Gamma_\infty^2(t(u))\right)$ совпадают с математическим ожиданием и ковариацией суммы квадратов броуновских мостов. Тогда, аналогично работе [7], можем считать, что для больших объемов выборок в качестве приближения асимптотического распределения статистики $\tilde{T}^2$ может быть использовано распределение Кифера — Гихмана [7, 13].

Точные распределения $\tilde{T}^2$. Для вычисления точных распределений $\tilde{T}^2$ воспользуемся q-мерной моделью случайного блуждания. Частица двигается по q-мерному массиву ячеек $\Omega = \left\{ a_{j_1, \ldots, j_q} \middle| 0 \le j_i \le n_i, \, i = \overline{1, q} \right\}$, выходя на первом шаге из ячейки $a_{0, \ldots, 0}$ и оканчивая блуждание в ячейке $a_{n_1, n_2, \ldots, n_q}$. Более подробно рассматриваемый далее метод описан в общем виде в работе [12].

Зададим объединенный вариационный ряд выборок $\vec{\Theta}_1, \vec{\Theta}_2, \ldots, \vec{\Theta}_q$, составленных из наработок до отказа систем, $\vec{\gamma} = (\gamma_1, \ldots, \gamma_n)$, $\gamma_j < \gamma_{j+1}$, $j = \overline{1, n-1}$. Введем вектор $\vec{z} = (z_1, \ldots, z_n)$, положим $z_j = i$, если γ_j — элемент i-й выборки. Для удобства вычислений обозначим $\tilde{k}_i = 1/k_i, \, i = \overline{1, q}$.

Пусть ν_j^i — количество элементов выборки $\vec{\Theta}_i$, на первых j местах в векторе $\vec{z} = (z_1, \ldots, z_n)$, $\sum_{i=1}^{q} \nu_j^i = j$. Тогда $\sum_{i=1}^{q} \left(n_i - \nu_{j-1}^i\right) m_i \tilde{k}_i$ — объем множества риска перед j-м отказом в вариационном ряду γ_j. Ниже в несколько измененном виде приведена лемма, доказанная в работе [7].

Лемма 2. Распределение векторов $\vec{z}$ при справедливости (2) имеет вид

$$p(\vec{z}) = \prod_{j=1}^{n} \frac{\prod_{i=1}^{q} \left(\left(n_i - \nu_{j-1}^i\right) m_i \tilde{k}_i \right)^{\nu_j^i - \nu_{j-1}^i}}{\sum_{i=1}^{q} \left(n_i - \nu_{j-1}^i\right) m_i \tilde{k}_i}. \blacktriangleleft$$

При прохождении частицы через ячейку $a_{j_1, \ldots, j_q}$ функция $\left(\sum_{i=1}^{q} n_i \left(\hat{P}_i^{k_i} - \bar{P} \right)^2 + \tilde{P}\left(\sqrt{\Phi(q-1)} - \Phi_1 \right) \right) / \Gamma^2$ принимает значение, равное $\omega_{j_1, \ldots, j_q} = A_{j_1, \ldots, j_q} / B_{j_1, \ldots, j_q}$.

Здесь

$$A_{j_1,\ldots,j_q} = \Delta_{j_1,\ldots,j_q} + \left(\sum_{i=1}^{q} \rho_i \left(1 - \frac{j_i}{n_i} \right)^{k_i/m_i} \right) \left(\sqrt{\varphi_{j_1,\ldots,j_q}(q-1)} - {\varphi^1}_{j_1,\ldots,j_q} \right),$$

$$\Delta_{j_1,\ldots,j_q} =$$

$$= \sum_{i=1}^{q} n_i \left[\left(\prod_{s_i=1}^{j_i} \left(1 - \frac{1}{m_i \left(n_i - s_i + 1 \right)} \right) \right)^{k_i} - \sum_{i=1}^{q} \left(\rho_i \left(\prod_{s_i=1}^{j_i} \left(1 - \frac{1}{m_i \left(n_i - s_i + 1 \right)} \right)^{k_i} \right) \right) \right],$$

$$B_{j_1,\ldots,j_q} = \sqrt{\frac{\varphi_{j_1,\ldots,j_q}}{q-1}} + \sum_{i=1}^{q} \rho_i \left(1 - \frac{j_i}{n_i} \right)^{k_i/m_i},$$

$$W_i = \frac{k_i^2 \left(1 - \left(\sum_{i=1}^{q} \rho_i \left(1 - \frac{j_i}{n_i} \right)^{k_i/m_i} \right)^{m_i/k_i} \right)}{m_i^2 \left(\sum_{i=1}^{q} \rho_i \left(1 - \frac{j_i}{n_i} \right)^{k_i/m_i} \right)^{m_i/k_i - 1}},$$

$$\varphi^1_{j_1,\ldots,j_q} = \sum_{i=1}^{q} (1-\rho_i) W_i, \;\; \varphi_{j_1,\ldots,j_q} = \left(\sum_{i=1}^{q} \rho_i W_i \right)^2 + \sum_{i=1}^{q} (1 - 2\rho_i) W_i^2.$$

Теорема 3. Вероятность $P\left(T^2 < h\right)$ равна величине $\pi_{n_1,n_2,\ldots,n_q}(h)$. Ее можно получить, применив рекуррентное соотношение

$$\pi_{j_1,j_2,\ldots,j_q}(h) =$$

$$= \left(\begin{array}{l} \pi_{j_1-1,j_2,\ldots,j_q}(h) \dfrac{m_1 \tilde{k}_1 (n_1 - j_1 + 1)}{\left(\sum_{i=1}^{q} m_i \tilde{k}_i (n_i - j_i) \right) + m_1 \tilde{k}_1} + \ldots + \\ + \pi_{j_1,j_2,\ldots,j_q-1}(h) \dfrac{m_q \tilde{k}_q (n_q - j_q + 1)}{\left(\sum_{i=1}^{q} m_i \tilde{k}_i (n_i - j_i) \right) + m_q \tilde{k}_q} \end{array} \right) \chi_{j_1,j_2,\ldots,j_q}(h), \qquad (4)$$

с начальными и граничными условиями: $\pi_{0,\ldots,0}(h) = 1,\; \pi_{-1,0,\ldots,0} = \ldots = \pi_{0,\ldots,0,-1} = 0$.

Здесь $\chi_{j_1,j_2,\ldots,j_q}(h) = \begin{cases} 1, & \omega_{j_1,j_2,\ldots,j_q} < h, 0 \le j_i \le n_i, \\ 0, & \text{в противном случае.} \end{cases}$

Доказательство. Условие $\omega_{i_1,i_2,\ldots,i_q} < h$ определяет множество ячеек, при прохождении через которые траектории случайного блуждания, значение случайного процесса функции $\left(\sum_{i=1}^{q} n_i \left(\widehat{P}_i^{k_i} - \overline{P}\right)^2 + \right.$ $\left. + \tilde{P}\left(\sqrt{\Phi(q-1)} - \Phi_1\right)\right) \Big/ \Gamma^2$ не превосходит h. Тогда утверждение теоремы следует из леммы 2 и результатов работы [14].◄

В таблице для проверки справедливости модели Кокса для трех выборок представлены точные вероятности $P(\tilde{T}^2 < h)$, вычисленные согласно соотношению (4), в случае равных объемов выборок при $m_1 = m_2 = m_3 = 2$.

Точные вероятности $P(\tilde{T}^2 < h)$ для трех выборок при $h = 3,186$

n	$h = 3,186$	
	$k_1 = 1;\ k_2 = 2;\ k_3 = 3$	$k_1 = 1;\ k_2 = 1,5;\ k_3 = 2$
100	0,9842	0,9852
200	0,9853	0,9862
300	0,9853	0,9861
400	0,9854	0,9860
500	0,9852	0,9859
600	0,9852	0,9859
700	0,9851	0,9858
∞	0,9848	0,9848

Аргумент $h = 3,186$ является квантилем уровня 0,9848 предельного распределения Кифера — Гихмана. Таким образом, для практических целей можно использовать асимптотическое распределение, начиная с объемов выборок больше 100.

Оценки параметров Кокса. Полученный критерий типа Кифера — Гихмана позволяет проверять гипотезы о возможных значениях параметров модели Кокса. Однако во многих случаях наиболее важна оценка данных параметров.

Предположим, что значения параметров $k_i, i = \overline{2,q}$ неизвестны. Их оценки могут быть получены минимизацией статистики (2).

Методами статистического моделирования проведено исследование точности получаемых оценок с помощью алгоритма, представленного ниже. Алгоритм моделирования является обобщением алгоритма, представленного в работах [14, 15].

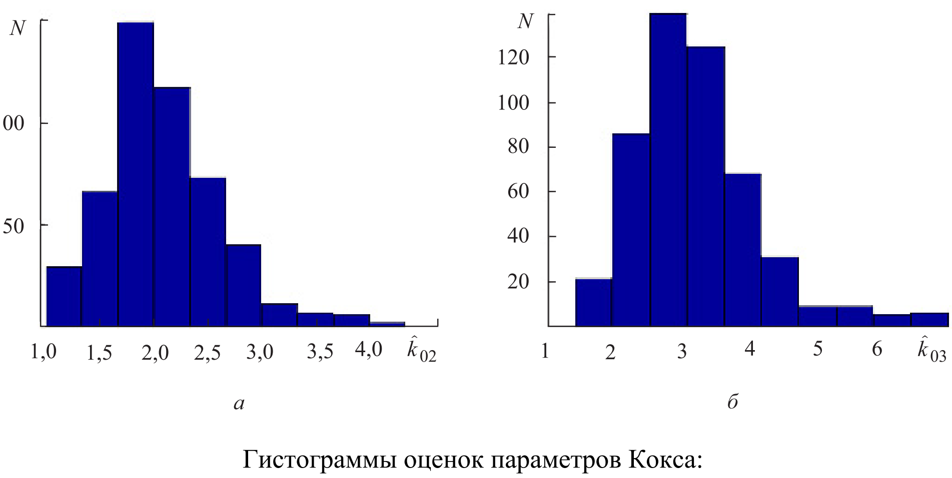
1. Моделируем $n_1 m_1$ наработок до отказа с функцией распределения $F_0(t)$.

2. Наработки случайным образом разбиваем на n_1 групп по m_1 величин в каждой. В результате $\left(\xi_1^{1,j},...,\xi_{m_1}^{1,j}\right), j=\overline{1,n_1}$ — элементы j-й группы. Определяем $\theta_1^j=\min\left(\xi_1^{1,j},...,\xi_{m_1}^{1,j}\right)$.

3. Для $i=\overline{2,q}$ аналогичным образом моделируем $n_i m_i$ одинаково распределенных случайных величин $(\xi_1^i,...,\xi_{m_i n_i}^i)$ с функцией распределения $\left(1-\left(1-\left(F_0(t)\right)^{k_i}\right)^{1/k_i}\right)$, где k_i — некоторое заданное значение параметра. Наработки $(\xi_1^i,...,\xi_{m_i n_i}^i)$ случайным образом разбиваем на n_i групп по m_i величин в каждой. Элементы j-й группы обозначаем $\left(\xi_1^{i,j}, ..., \xi_{m_i}^{i,j}\right)$, $j=\overline{1,n_i}$. Определяем $\theta_i^j=\min\left(\xi_1^{i,j},...,\xi_{m_i}^{i,j}\right)$.

4. В результате имеем q выборок $\vec{\Theta}_1=\left(\theta_1^1, ...,\theta_1^{n_1}\right)$, $\vec{\Theta}_2=\left(\theta_2^1, ...,\theta_1^{n_2}\right)$, ..., $\vec{\Theta}_q=\left(\theta_q^1, ...,\theta_q^{n_q}\right)$. Для заданных значений параметров $k_{0i}, 1\le k_{0i}\le K_i,\ i=\overline{2,q}$ по полученным выборкам вычисляем значение статистики типа Кифера — Гихмана $\tilde{T}^2\left(k_{02}, k_{03}, ..., k_{0q}\right)$.

5. В качестве оценок $\hat{k}_i, i=\overline{2, q}$ определяется набор значений $k_{0i}, i=\overline{2, q}$, минимизирующий значение статистики $\tilde{T}^2\left(k_{02},k_{03}, ..., k_{0q}\right)$.



Гистограммы оценок параметров Кокса:
$а$ — $k_2=2$; $б$ — $k_3=3$

При численных расчетах для определения статистических свойств предложенных оценок параметров п. 1–5 повторялись 500 раз. В качестве распределения $P_0(t)$ рассмотрено экспоненциальное распределение с параметром $\beta = 0,001$. Гистограммы полученных оценок $\widehat{k}_{02}, \widehat{k}_{03}$ для $m_1 = m_2 = m_3 = 2$, $n_1 = n_2 = n_3 = 50$, $k_2 = 2$ и $k_3 = 3$ приведены на рисунке.

Для полученных оценок были рассчитаны средние значения, составившие $\widehat{k}_{02} = 2{,}1048$, $\widehat{k}_{03} = 3{,}2297$. Данные результаты свидетельствуют о допустимости применения предложенного метода оценки.

Заключение. В работе предложен критерий для проверки справедливости модели Кокса для нескольких независимых прогрессивно цензурированных выборок. Для нахождения оценок функций надежности использованы оценки Каплана — Мейра. Показано, что в качестве приближения асимптотического распределения может быть использовано распределение Кифера — Гихмана. Предложен метод вычисления точных распределений предлагаемой статистики. Методом Монте-Карло получены оценки параметров модели Кокса. Полученные результаты находят широкое применение при проведении форсированных испытаний сложных систем в случае чрезмерной продолжительности обычных испытаний.

ЛИТЕРАТУРА

[1] Гнеденко Б.В., Беляев Ю.К., Соловьев А.Д. *Математические методы в теории надежности. Основные характеристики надежности и их статистический анализ*. Москва, УРСС Либроком, 2013, 584 с.

[2] Садыхов Г.С., Крапоткин В.Г., Казакова О.И. Расчет и оценка показателей ресурса изделий с использованием модели аддитивного накопления повреждений. *Математическое моделирование и численные методы*, 2014, № 1 (1), с. 82–98.

[3] Зарубин В.С., Кувыркин Г.Н. Особенности математического моделирования технических устройств. *Математическое моделирование и численные методы*, 2014, № 1 (1), с. 5–17.

[4] Kiefer J. K-sample analogues of the Kolmogorov-Smirnov and Cramer-V. mises tests. *The Annals of Mathematical statistics*, 1959, vol. 30, no. 2, pp. 420–447.

[5] Гихман И.И. Об одном непараметрическом критерии однородности k выборок. *Теория вероятностей и ее применения*, 1957, № 2, с. 380–384.

[6] Тимонин В.И., Ермолаева М.А. Многовыборочный аналог критерия Смирнова проверок степенных гипотез Лемана. *Электромагнитные волны и электронные системы*, 2011, № 11, с. 6–11.

[7] Тимонин В.И., Тянникова Н.Д. Прогрессивное цензурирование — проверка однородности нескольких независимых выборок. *Физические основы приборостроения*, 2016, т. 5, № 2 (19), с. 80–87.

[8] Balakrishnan N., Cramer E. *The art of progressive censoring. Applications to reliability and quality*. New York, Springer, 2014, 645 p.

[9] Bagdonavicius V., Kruopis J., Nikulin M.S. *Nonparametric tests for censored data*. London, Wiley, 2011, 233 p.

[10] Ng N., Balakrishnan N. Precedence-type test based on Kaplan-Meier estimator of cumulative distribution function. *Journal of Statistical Planning and Inference*, 2010, vol. 140, no. 8, pp. 2295–2311.

[11] 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.

[12] Bordes L. Non-parametric estimation under progressive censoring. *Journal of Statistical Planning and Inference*, 2004, no. 119, pp. 171–189.

[13] Большев Л.Н., Смирнов Н.В. *Таблицы математической статистики.* Москва, Наука, 1983, 416 с.

[14] Тянникова Н.Д., Тимонин В.И. Метод вычисления точных распределений статистик типа Колмогорова‒Смирнова в случае нарушения однородности и независимости анализируемых выборок. *Наука и образование: электронное научное издание*, 2014, № 11.
URL: http://technomag.bmstu.ru/doc/740251.html (дата обращения 23.05.2017).

[15] Тимонин В.И., Тянникова Н.Д. Сравнение прогрессивно цензурированных выборок — численные методы табулирования распределений статистик однородности и исследование оценки параметров связи их распределений методом Монте-Карло. *Математическое моделирование и численные методы*, 2015, № 3 (7), с. 89–100.

Статья поступила в редакцию 19.04.2017

Ссылку на эту статью просим оформлять следующим образом:
Тимонин В.И., Тянникова Н.Д. Проверка справедливости модели Кокса по нескольким прогрессивно цензурированным выборкам. *Математическое моделирование и численные методы*, 2017, № 1, с. 102–117.

Тимонин Владимир Иванович — д-р физ.-мат. наук, профессор кафедры «Высшая математика» МГТУ им. Н.Э. Баумана, автор более 40 работ в теории надежности и математической статистике. e-mail: timoninmgtu52@mail.ru

Тянникова Нина Дмитриевна — канд. физ.-мат. наук, ассистент кафедры «Высшая математика» МГТУ им. Н.Э. Баумана. Область научных интересов: теория надежности, математическая статистика. e-mail: tiannikova@yandex.ru

Cox model validity checking for several progressively censored samples

© V.I. Timonin, N.D. Tyannikova

Bauman Moscow State Technical University, Moscow, 105005, Russia

The article proposes a nonparametric criterion of the Kiefer — Gihman type to test the Cox model validity for several progressively censored samples. As estimates of the reliability function for each sample we are using the Kaplan — Meyer ones.
The paper proves that if the hypothesis is valid, the Kiefer — Gihman distribution can be used as an approximation of the asymptotic distribution of the criterionstatistics. Based on the particle random walk model over a multidimensional cells array, the paper has developed the method for calculating the exact statistics distributions. The article pre-

sents obtained probability values tables of the proposed statistics exact distributions for a wide range of samples possible values. Statistical modeling methods show Cox parameters estimating method consistency, based on the statistics minimization. We present the obtained estimates histograms for the developments exponential distribution to failure.
The research results are used when analyzing the redundant technical systems of different multiplicity tests results operating in different operating conditions.
Analyzed systems find applications in all industries — from machine building to radio electronic.

Keywords: *nonparametric statistics, Kiefer — Gihman type criterion, Kaplan — Meyer estimates, progressive censoring, Cox model*

REFERENCES

[1] Gnedenko B.V., Belyaev Yu.K., Solovev A.D. *Matematicheskie metody v teorii nadezhnosti. Osnovnye kharakteristiki nadezhnosti i ikh statisticheskiy analiz* [Mathematical methods in reliability theory. The main characteristics of reliability and their statistical analysis]. Moscow, Librokom Publ., 2013, 584 p.

[2] Sadykhov G.S., Krapotkin V.G., Kazakova O.I. *Matematicheskoe modelirovanie i chislennye metody — Mathematical Modeling and Computational Methods*, 2014, no. 1 (1), pp. 82–98.

[3] Zarubin V.S., Kuvyrkin G.N. *Matematicheskoe modelirovanie i chislennye metody — Mathematical Modeling and Computational Methods,* 2014, no. 1 (1), pp. 5–17.

[4] Kiefer J. *The Annals of Mathematical statistics*, 1959, vol. 30, no. 2, pp. 420–447.

[5] Gihman I.I. *Teoriya veroyatnostey i ee primeneniya — Theory of Probability and its Applications*, 1957, no. 2, pp. 380–384.

[6] Timonin V.I., Ermolaeva M.A. *Elektromagnitnye volny i elektronnye sistemy — Electromagnetic Waves and Electronic Systems*, 2011, no. 11, pp. 6–11.

[7] Timonin V.I., Tyannikova N.D. *Fizicheskie osnovy priborostroeniya — Physical Bases of Instrumentation*, 2016, vol. 5, no. 2 (19), pp. 80–87.

[8] Balakrishnan N., Cramer E. *The art of progressive censoring. Applications to reliability and quality*. New York, Springer, 2014, 645 p.

[9] Bagdonavicius V., Kruopis J., Nikulin M.S. *Nonparametric tests for censored data*. London, Wiley, 2011, 233 p.

[10] Ng N., Balakrishnan N. *Journal of Statistical Planning and Inference*, 2010, vol. 140, no. 8, pp. 2295–2311.

[11] Timonin V.I., Tyannikova N.D. *Vestnik MGTU im. N.E. Baumana. Ser. Estestvennye nauki — Herald of Bauman Moscow State Technical University. Series Natural Sciences*, 2015, no. 6, pp. 68–84.

[12] Bordes L. *Journal of Statistical Planning and Inference*, 2004, no. 119, pp. 171–189.

[13] Bolshev L.N., Smirnov N.V. *Tablitsy matematicheskoy statistiki* [Tables of mathematical statistics]. *Moscow*, Nauka Publ., 1983, 416 p.

[14] Tyannikova N.D., Timonin V.I. *Nauka i obrazovanie — Science and Education*, 2014, no. 11. Available at: http://technomag.bmstu.ru/doc/740251.html (accessed May 23, 2017).

[15] Timonin V.I., Tyannikova N.D. *Matematicheskoe modelirovanie i chislennye metody — Mathematical Modeling and Computational Methods*, 2015, no. 3 (7), pp. 89–100.

Timonin V.I., Dr. Sc. (Phys.-Math.), Professor of Higher Mathematics Department, Bauman Moscow State Technical University. Author of more than 40 research works in the field of theory of reliability and mathematical statistics.
e-mail: timoninmgtu52@mail.ru

Tyannikova N.D., Cand. Sc. (Phys.-Math.), Assistant of Higher Mathematics Department, Bauman Moscow State Technical University. Science research interests include: reliability theory, mathematical statistics. e-mail: tiannikova@yandex.ru