УДК 519.248: 62–192

Критерии типа Колмогорова – Смирнова проверки гипотез Кокса – Лемана в случае прогрессивно цензурированных выборок — о возможности использования оценок Каплана – Мейера в статистиках критериев

В.И. Тимонин[1], **Н.Д. Тянникова**[1]

[1] *Кафедра высшей математики,*
Московский государственный технический университет
имени Н.Э. Баумана,
ул. 2-ая Бауманская, д.5, стр.1, Москва, Россия, 105005

Аннотация. В работе показана возможность применения оценок Каплана – Мейера при проверке гипотез однородности (в общем случае — гипотез Кокса – Лемана) статистиками типа Колмогорова – Смирнова, если выборки являются прогрессивно цензурированными. Доказано, что распределение статистик при справедливости проверяемых гипотез не зависит от распределения элементов выборок. Разработан метод вычисления точных распределений статистик при справедливости проверяемых гипотез. Для частного случая получен критерий типа Колмогорова – Смирнова проверки гипотез Кокса – Лемана о распределении наработок элементов сложных систем. Он основан на сравнении оценок Каплана – Мейера функции надежности элементов систем по результатам испытаний в нескольких режимах последовательных (или параллельных) систем различной кратности. Показано, что предельное распределение его статистики в случае двух выборок является классическим распределением Колмогорова – Смирнова, в случае нескольких выборок — может быть приближено распределением Кифера — Гихмана. Проведено сравнение точных и асимптотических квантилей распределений статистик. Показано, что асимптотическим распределением можно пользоваться, начиная с объемов выборок порядка 70 – 80 единиц.

Ключевые слова: прогрессивное цензурирование, критерий типа Колмогорова – Смирнова, гипотезы Кокса – Лемана, оценка Каплана – Мейера.

1. Введение

В работе используется терминология теории надежности, которая наиболее удобна для изложения результатов статьи.

Пусть в некотором режиме испытывается N изделий. Механизм прогрессивного цензурирования описывается следующим образом [1,2]. При очередном i-м отказе $\xi_i(\xi_{i-1} < \xi_i, i = \overline{1, s})$ с испытаний снимаются (цензурируются) r_i случайным образом отобранных изделий. Уточним, что параметры r_i, s известны заранее и не являются случайными величинами. В этом случае непараметрическая оценка Каплана –

Мейера функции надежности $P_\xi(t) = P(\xi > t)$ наработок до отказа изделий имеет вид [3]

$$\widehat{P}_{\text{КМ}}(t) = \prod_{\xi_i < t}\left(1 - \frac{1}{R_i}\right), R_i = N - \sum_{j=1}^{i-1}(r_j+1), i = \overline{2, s-1}, R_1 = N;$$
$$R_s = 1 \text{ при } t > \xi_s; \widehat{P}_{\text{КМ}}(t) = 1 \text{ при } t < \xi_1.$$

R_i называется объемом множества риска перед отказом ξ_i.

Частным случаем рассмотренной схемы являются испытания n последовательных систем, составленных из одного и того же количества m одинаковых элементов, $r_i = r = m-1, N = nm$. Число m будем называть кратностью системы.

2. Основная часть

Обозначим $\lambda_i(t), i = \overline{1,q}$ — интенсивность отказов элемента в режиме эксплуатации (испытаний) ε_i. Требуется проверить гипотезу Кокса

$$H_0^1 : k_1\lambda_1(t) = k_2\lambda_2(t) = ... = k_q\lambda_q(t), \tag{1}$$

где $k_i \geq 1, i = \overline{1,q}$ — известные фиксированные числа, вид $\lambda_i(t)$ полностью неизвестен. Без ограничения общности будем считать, что $k_1 = 1$, $k_1 \leq k_2 \leq ... \leq k_q$. В дальнейшем для уменьшения объема изложения описывается случай $q = 2$. В этом случае гипотеза (1) имеет вид

$$H_0 : \lambda_1(t) = k\lambda_2(t) \Leftrightarrow P_1(t) = \left(P_2(t)\right)^k. \tag{2}$$

Общий случай дан ссылками на работы авторов [4,5]. Гипотеза (2) проверяется по двум прогрессивно цензурированным выборкам из наработок до отказа элементов в последовательных системах.

Испытываются n_1 систем кратности m_1 в режиме ε_1 и n_2 систем кратности m_2 в режиме ε_2. При отказе одного из элементов системы, оставшиеся $(m_j - 1), j = 1,2$ наработки до отказа элементов системы цензурируются. Тогда имеются две выборки $\Theta_1 = \left(\theta_1^1, ..., \theta_1^{n_1}\right), \Theta_2 = \left(\theta_2^1, ..., \theta_2^{n_2}\right)$ где $\theta_1^i = \min\{\xi_1^i, \xi_2^i, ..., \xi_{m_1}^i\}, i = \overline{1, n_1}, \theta_2^j = \min\{\xi_1^j, \xi_2^j, ..., \xi_{m_2}^j\}$, $j = \overline{1, n_2}$ — минимумы наработок до отказа элементов систем, работающих в режимах ε_1 и ε_2 соответственно.

Для проверки гипотезы (2) предлагается статистика вида

$$T = C\max_t \frac{\left(k_2\left(1-\widehat{F}^1\right)^{\frac{1}{m_1}} + k_1\left(1-\widehat{F}^2\right)^{\frac{k}{m_2}}\right)^{\frac{m_2}{k}-1}}{k_2\left(k_2\left(1-\widehat{F}^1\right)^{\frac{1}{m_1}} + k_1\left(1-\widehat{F}^2\right)^{\frac{k}{m_2}}\right)^{\frac{m_2}{k}-1} + k_1} \cdot \tag{3}$$
$$\cdot \left|\widehat{P}_{\theta_1}(t) - \left(\widehat{P}_{\theta_2}(t)\right)^k\right|.$$

Здесь $C = \dfrac{m_1 m_2 \sqrt{\rho n_2}}{\sqrt{k^2 \rho m_1^2 + m_2^2}}, \widehat{P}_{\theta_1}(t), \widehat{P}_{\theta_2}(t)$ — оценки Каплана – Мейера функций надежности элементов систем первой и второй выборок; $F^1(t) = 1 - (P_1(t))^{m_1}, F^2(t) = 1 - (P_2(t))^{m_2}$ — функции распределения наработок до отказа систем, $\widehat{F}^1(t), \widehat{F}^2(t)$ — эмпирические функции распределения выборок $\Theta_1, \Theta_2, \rho = \dfrac{n_1}{n_2}, k_2 = \dfrac{m_2^2}{\rho m_1^2 k^2 + m_2^2}, k_1 = 1 - k_2$.

Статистика (3) при $m_1 = m_2 = k = 1$ совпадает с классической статистикой Смирнова проверки однородности двух нецензурированных выборок.

Теорема 1. Точные вероятности $P\{T < h\}$ равны величине $\pi^h_{n_1,n_2}$, которую можно получить повторным применением соотношения

$$\pi^h_{ij} = \begin{cases} 1, \text{если } i = 0, j = 0; \\ \left(\dfrac{m_2(n_2 - j + 1)}{km_1(n_1 - i) + m_2(n_2 - j + 1)} \pi^h_{i,j-1}\right) \cdot \chi^h_{ij}, i = 0, j = \overline{1, n_2}; \\ \left(\dfrac{km_1(n_1 - i + 1)}{km_1(n_1 - i + 1) + m_2(n_2 - j)} \pi^h_{i-1,j}\right) \cdot \chi^h_{ij}, i = \overline{1, n_1}, j = 0; \\ \left(\begin{array}{c} \dfrac{km_1(n_1 - i + 1)}{km_1(n_1 - i + 1) + m_2(n_2 - j)} \pi^h_{i-1,j} + \\ + \dfrac{m_2(n_2 - j + 1)}{km_1(n_1 - i) + m_2(n_2 - j + 1)} \pi^h_{i,j-1} \end{array}\right) \cdot \chi^h_{ij}, \begin{array}{c} i = \overline{1, n_1}, \\ j = \overline{1, n_2}. \end{array} \end{cases}$$

$$\chi^h_{ij} = \begin{cases} 1, t_{ij} < h \\ 0, t_{ij} \geq h \end{cases};$$

$$t_{ij} = C \cdot \frac{\left(k_2 \left(1 - \frac{i}{n_1}\right)^{\frac{1}{m_1}} + k_1 \left(1 - \frac{j}{n_2}\right)^{\frac{k}{m_2}}\right)^{\frac{m_2}{k} - 1}}{k_2 \left(k_2 \left(1 - \frac{i}{n_1}\right)^{\frac{1}{m_1}} + k_1 \left(1 - \frac{j}{n_2}\right)^{\frac{k}{m_2}}\right)^{\frac{m_2}{k} - m_1} + k_1} \cdot \Delta_{ij};$$

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

Теорема 2. Асимптотическое распределение T при справедливости (2) является распределением Колмогорова – Смирнова с функцией распределения

$$K(x) = \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2x^2}.$$

В Таблице 1 приведены рассчитанные точные вероятности $P\{T < h\}$ и их предельные значения для некоторых значений n_1, n_2, k и $m_1 = m_2 = 2$ и квантилей $h = 1,22; 1,36; 1,63$.

Таблица 1

Точные и предельные вероятности $P\{T < h\}$ при $m_1 = 2, m_2 = 2$

$n_1 = n_2$	$P\{T < h\}$					
	$h = 1,22$		$h = 1,36$		$h = 1,63$	
	$k = 1,5$	$k = 3$	$k = 1,5$	$k = 3$	$k = 1,5$	$k = 3$
100	0,9108	0,8916	0,9572	0,9442	0,9913	0,9864
700	0,9041	0,9028	0,9536	0,9530	0,9908	0,9906
1100	0,9029	0,9021	0,9530	0,9528	0,9907	0,9906
1500	0,9020	0,9020	0,9527	0,9525	0,9906	0,9906
∞	0,8981	0,8981	0,9505	0,9505	0,9901	0,9901

В работе [5] для аналогичного вида испытаний систем в q режимах предложен критерий типа Кифера – Гихмана проверки (1). Предельное распределение его статистики может быть приближено распределением статистики Кифера для проверки однородности q полных выборок, точные распределения вычисляются согласно несложному обобщению приведенного в работе алгоритма.

3. Заключение

По результатам испытаний в различных режимах нескольких выборок последовательных систем различной кратности, составленных из однотипных элементов, решена задача проверки непараметрических гипотез (Кокса – Лемана) о распределениях наработок элементов. Для этого разработаны критерии типа Колмогорова – Смирнова и Кифера – Гихмана. Предложен общий метод вычисления точных и асимптотических распределений статистик.

Литература

1. *Balakrishnan N., Cramer E.* The Art of Progressive Censoring. Applications to Reliability and Quality. — New York, Springer, 2014.

2. *Bordes L.* Non-parametric estimation under progressive censoring // Journal of Statistical Planning and Inference. — 2004. — Vol. 119, no. 1. — P. 171–189.
3. *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. — P. 2295–2311.
4. *Тимонин В. И., Тянникова Н. Д.* Прогрессивное цензурирование — проверка однородности нескольких независимых выборок // Физические основы приборостроения. — 2016. — Т. 5, No. 2 (19). — С. 80–87.
5. *Тимонин В. И., Тянникова Н. Д.* Проверка справедливости модели Кокса по нескольким прогрессивно цензурированным выборкам // Математическое моделирование и численные методы. — 2017. – No. 1 (13). – С. 102–117.

UDC 519.248: 62–192

The Kolmogorov – Smirnov type tests for the Lehmann – Cox hypotheses in the case of progressively censored samples — about the possibility to use the Kaplan – Meier estimates in test statistics

V.I. Timonin[1], N.D. Tyannikova[1]

[1] *Department of Higher Mathematics,*
Bauman Moscow State Technical University,
ul. Baumanskaya 2-ya, 5, Moscow, 105005, Russia

The paper shows the possibility of using the Kaplan – Meier estimates to test homogeneity hypotheses (in general case — the Lehmann – Cox hypotheses) by the Kolmogorov – Smirnov statistics, when the samples are progressive censored. It is proved that under justice of hypotheses the statistics distribution does not depend on the sample elements distribution. A method for calculating the exact statistics distribution is developed. For the particular case the Kolmogorov – Smirnov type test for the Cox – Lehmann hypotheses about the complex systems elements distribution is presented. It is based on a comparison of the Kaplan – Meier estimates of the system elements reliability function according to the results of the experimental testing of different multiplicity serial (or parallel) systems in several modes. It is shown that when there are two samples the asymptotic distribution of its statistic is a classical Kolmogorov – Smirnov distribution, in the case of multiple samples, its distribution can be approximated by the distribution of the Kiefer – Gihman. The exact and asymptotic quintiles of the statistics distributions were compared. It is shown that the asymptotic distribution can be used since the sample sizes of the order of 70 – 80 units.

Keywords: progressive censoring, the Kolmogorov – Smirnov type test, the Cox – Lehmann hypotheses, the Kaplan — Meier estimates.