УДК 519.718

М.А. Ермолаева

ИССЛЕДОВАНИЕ МИНИМАКСНЫХ ОЦЕНОК ПАРАМЕТРОВ МОДЕЛИ КОКСА — ЛЕМАНА МЕТОДАМИ СТАТИСТИЧЕСКОГО МОДЕЛИРОВАНИЯ

С помощью статистического моделирования исследована точность оценок параметров степенной модели Кокса — Лемана. Разработан метод, основанный на минимизации статистики типа Кифера — Гихмана.

E-mail: ermolaevama@yandex.ru

Ключевые слова: *непараметрические методы, степенная модель Кокса — Лемана, критерий типа Колмогорова — Смирнова, точные и асимптотические распределения статистик, метод Монте-Карло.*

Оценивание связей между функциями распределения различных совокупностей является одной из основных задач прикладной математической статистики. Все большее распространение получают методы оценки непараметрической регрессии, основанные на моделях Кокса — Лемана [1]. Суть их заключается в анализе степенных связей между функциями распределения или функциями надежности наработок изделий в разных режимах. В 1983 г. В.И. Тимониным [2] была предложена статистика типа Колмогорова — Смирнова для проверки гипотезы

$$H_0 : F(t) = [G(t)]^k, \tag{1}$$

где $F(t), G(t)$ — функции распределения двух совокупностей.

Однако на практике часто необходимо проверить наличие степенной зависимости функций распределения наработок до отказа для нескольких режимов, т. е. для многовыборочного случая.

Пусть $\bar{\xi}_i = \left(\xi_1^i, \ldots, \xi_{n_i}^i\right)$ — наработки до отказа в режиме ε^i, $i = \overline{1, q}$; $n = \sum_{i=1}^{q} n_i$. Таким образом, обобщением гипотезы (1) на многовыборочный случай является гипотеза

$$H_0 : F_1^{k_1}(t) = ... = F_q^{k_q}(t). \tag{2}$$

Без ограничения общности можно считать, что $k_1 = 1$, $k_1 \le k_2 \le \ldots \le k_q$. В случае $k_i = 1$ для любого i гипотеза является классической гипотезой однородности.

Критерий проверки основной гипотезы. Для проверки гипотезы (2) предлагается использовать статистику вида

$$T_k{}^2 = \max_t \frac{\sum_{i=1}^{q} n_i \left(\widehat{F}_i^{k_i} - \overline{F}\right)^2 + \overline{F}\left(\Phi\sqrt{q-1} - \Phi_1\right)}{\Gamma^2}. \qquad (3)$$

Здесь $\widehat{F}_i(t)$ — эмпирическая функция распределения i-й выборки;

$$\overline{F} = \sum_{i=1}^{q} \rho_i \widehat{F}_i^{k_i}(t); \quad \rho_i = n_i / n;$$

$$\Phi = \left[\left(\sum_{i=1}^{q} k_i^2 \rho_i \overline{F}^{1-k_{0i}} \left(1 - \overline{F}_i^{k_{0i}}\right)\right)^2 + \sum_{i=1}^{q} k_i^4 \overline{F}^{2(1-k_{0i})} \left(1 - \overline{F}^{k_{0i}}\right)^2 (1 - 2\rho_i)\right]^{1/2};$$

$$\Phi_1 = \sum_{i=1}^{q} k_i^2 (1-\rho_i) \overline{F}^{1-k_{0i}} \left(1 - \overline{F}^{k_{0i}}\right); \; \Gamma = \overline{F} + \Phi / \sqrt{q-1}.$$

Доказано, что основанный на этой статистике критерий является состоятельным, получен метод вычисления точных вероятностей статистики $T_k{}^2$ при справедливости проверяемой гипотезы и найдено приближенное асимптотическое распределение $T_k{}^2$ [3].

Алгоритм оценивания параметров степенной модели. Для оценки неизвестных параметров k_i, $i = 2, \ldots, q$, воспользуемся статистикой $T_k{}^2$. В качестве оценки выберем значения k_i, минимизирующие значение $T_k{}^2$ (минимаксная оценка).

Алгоритм статистического моделирования с целью определения точности предлагаемого метода оценивания и условий его применения (минимальный объем выборок, вычисление поправочных коэффициентов, определение интервальных оценок с использованием нормального приближения) включал в себя следующие шаги:

1) выбор значений q, $k_1 = 1$, $k_2, ..., k_q$;

2) моделирование q выборок с функциями распределения F_i, $i = 1, \ldots, q$, по n_i элементов в каждой, $F_1 = F_2^{k_2} = \ldots = F_q^{k_q}$;

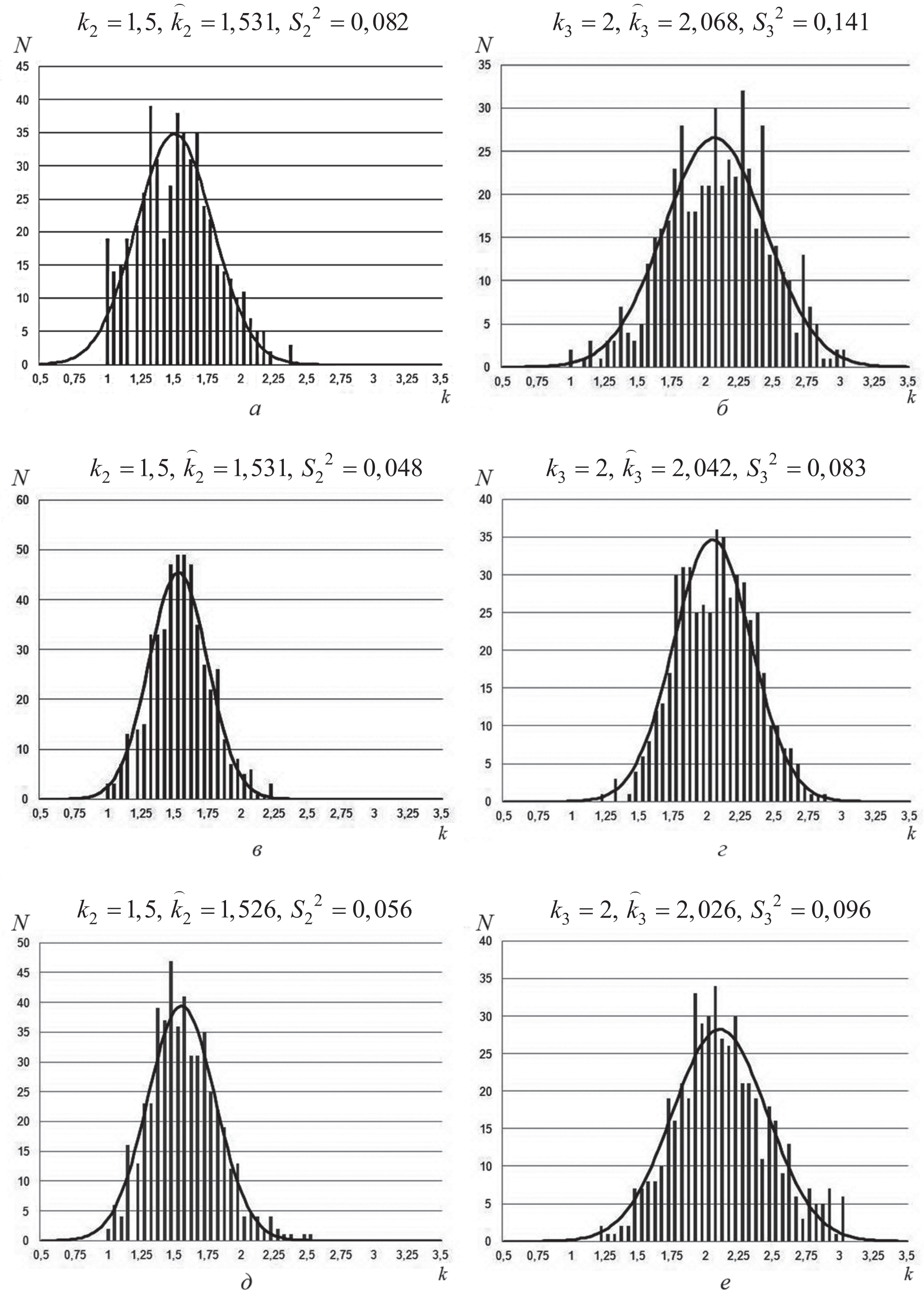
3) определение значений $\widehat{k}_i$, минимизирующих значение $T_k{}^2$;

4) расчет оценок $\widehat{k}_2, \ldots, \widehat{k}_q$, согласно п. 2 и 3, 500 раз;

5) построение гистограммы оценок, вычисление эмпирического среднего оценки $\widehat{k}_i$ и ее дисперсии $S_i{}^2$.

С помощью данного алгоритма было проведено обширное статистическое моделирование для различных объемов выборок и значений параметров. В качестве закона распределения элементов в выборках был выбран экспоненциальный вида $F_1 = 1 - e^{-\lambda x}$ с параметром $\lambda = 0{,}001$, $F_1 = F_2^{k_2} = F_3^{k_3}$.

Результаты моделирования. На рисунке показаны гистограммы оценок параметров k_2, k_3 для трех выборок объемом от 50 до 120 элементов. Здесь n_1, n_2, n_3 — объемы первой, второй и третьей вы-



Результаты моделирования:

а, б – $n_1 = n_2 = n_3 = 50$; *в, г* – $n_1 = n_2 = n_3 = 100$; *д, е* – $n_1 = 60$, $n_2 = 90$, $n_3 = 120$

борок соответственно; k_i — параметры, при которых проводилось моделирование; $\widehat{k}_i$ — полученные оценки параметров моделирования; S_i^2 — оценки дисперсии $\widehat{k}_i$.

Анализ результатов моделирования показывает, что при одинаковых объемах выборок точность оценки становится удовлетворительной при объемах не менее 50 элементов, а при неравных объемах в выборках — не менее 60 элементов. Для конечных объемов выборок наблюдается определенное систематическое смещение оценок вправо, которое уменьшается с увеличением объемов выборок (как следствие состоятельности используемого критерия). Методом Монте-Карло можно определять поправочные коэффициенты для устранения смещения оценок.

Таким образом, показано, что предлагаемый метод оценивания является состоятельным и пригодным для использования на практике.

СПИСОК ЛИТЕРАТУРЫ

1. Кокс Д., Оукс Д. Анализ данных типа времени жизни. – М.: Финансы и статистика, 1988. – 191 с.
2. Тимонин В.И. О предельном распределении статистики одного непараметрического критерия // Теория вероятностей и ее применение. – 1987. – Т. 32. – № 4. – С. 790–792.
3. Тимонин В.И., Ермолаева М.А. Многовыборочный аналог критерия Смирнова проверок степенных гипотез Лемана // Электромагнитные волны и электронные системы. – 2011. – № 11. – С. 6–11.

Статья поступила в редакцию 28.09.2012