УДК 519.254

С. В. Антоненко, А. Е. Полонская

Днепропетровский национальный университет имени Олеся Гончара

ВЫЧИСЛИТЕЛЬНЫЕ СХЕМЫ ВОССТАНОВЛЕНИЯ СПЛАЙН-ЭКСПОНЕНЦИАЛЬНОГО РАСПРЕДЕЛЕНИЯ С ДВУМЯ УЗЛАМИ

Предложен метод и вычислительная схема восстановления сплайн-экспоненциального распределения с двумя узлами.

Ключевые слова: *сплайн-экспоненциальное распределение; итерационный метод; вычислительная схема.*

Запропоновано метод і обчислювальну схему відновлення сплайн-експоненціального розподілу з двома вузлами.

Ключові слова: *сплайн-експоненціальний розподіл; ітераційний метод; обчислювальна схема.*

A method and a computational scheme for recovering a spline-exponential distribution with two nodes are proposed.

Keywords: *spline-exponential distribution; iterative method; computational scheme.*

Введение. Практически любой вид анализа данных, основанный на аппарате математической статистики, опирается на процедуры восстановления распределений. Поэтому актуальной является алгоритмизация и разработка современных методов восстановления распределений. При оценке показателей надежности и эффективности функционирования технических изделий используются различные методы обработки массивов данных об особых случаях, возникающих в процессе эксплуатации. К таким особым случаям относятся, например, отказы каких-либо систем наблюдаемой техники. Описание массивов отказов классическими законами распределений зачастую приводит к заниженным оценкам. Поэтому были предложены другие законы распределений, в частности, сплайны [1].

Анализ литературных данных и постановка проблемы. Для параметрических сплайн-распределений с одним узлом склеивания реализованы эффективные процедуры их восстановления по статическим данным [1, 2]. Они реализуют классические методы оценки параметров распределения, при этом узлы склеивания находятся путем перебора вариант. Последующее наращивание кусков распределения требует введения новых вычислительных процедур. Одна из таких вычислительных процедур для сплайн-экспоненциального распределения с двумя узлами рассмотрена в работе [3].

Постановка задачи. В настоящей работе предложена разработка нового метода восстановления сплайн-экспоненциального распределения с двумя узлами, который обладает несколькими преимуществами:

– простой с вычислительной точки зрения;
– быстрый;
– итерационный.

Вследствие своей итерационности предложенный метод позволяет наращивать число узлов склеивания без разработки новых процедур.

Основной материал. Рассмотрим предложенный метод восстановления сплайн-экспоненциального распределения с двумя узлами.

Пусть имеем выборку независимых наблюдений $\Omega_n = \{x_1, x_2, ..., x_n\}$ случайной величины ξ, которая подчиняется сплайн-экспоненциальному распределению с двумя узлами

$$f(x,\overline{\theta}) = \begin{cases} \lambda_1 \exp(-\lambda_1 x), & \mathrm{x} \in (0, x_1]; \\ \lambda_2 \exp(-\lambda_2 x - (\lambda_1 - \lambda_2)\,\mathrm{x}_1), & \mathrm{x} \in (x_1, x_2]; \\ \lambda_3 \exp(-\lambda_3 x - (\lambda_1 - \lambda_2)\,\mathrm{x}_1 - (\lambda_2 - \lambda_3)\,\mathrm{x}_2), & \mathrm{x} \in (x_2, \infty); \end{cases}$$

где $\overline{\theta} = \{\lambda_1, \lambda_2, \lambda_3, x_1, x_2\}$.

Для восстановления сплайн-экспоненциального распределения с двумя узлами предлагается реализовать итерационную процедуру на основе алгоритма восстановления сплайн-экспоненциального распределения с одним узлом. Причем такая процедура может использоваться для восстановления сплайн-экспонециального распределения с произвольным числом узлов склеивания.

Итак, приведем процедуру восстановления сплайн-экспоненциального распределения с одним узлом.

Пусть имеем выборку независимых наблюдений $\Omega_n = \{x_1, x_2, ..., x_n\}$ случайной величины ξ, которая подчиняется сплайн-экспоненциальному распределению с одним узлом

$$f(x,\overline{\theta}) = \begin{cases} \lambda_1 \exp(-\lambda_1 x), & \mathrm{x} \in (0, x_1]; \\ \lambda_2 \exp(-\lambda_2 x - (\lambda_1 - \lambda_2)\,\mathrm{x}_1), & \mathrm{x} \in (x_1, \infty); \end{cases}$$

$$\text{где } \overline{\theta} = \{\lambda_1, \lambda_2, x_1\}.$$

Предполагается поиск оценок параметров распределения и местоположения узла склеивания из условия

$$\rho = \max_{\substack{\mathrm{x}_k, \\ \overline{k=2,n-3}}} \max_{\underline{\Theta}_k} L\left(x_1,\ldots,x_n;\overline{\theta_k}\right),$$

где функция правдоподобия

$$L\left(x_1,\ldots,x_n;\overline{\theta}\right) = \prod_{i=1}^{k} f_1\left(x_i,\overline{\theta}\right) \prod_{i=k+1}^{n} f_2\left(x_i,\overline{\theta}\right).$$

Логарифм функции правдоподобия имеет вид

$$\Lambda = \ln L\left(x_1,\ldots,x_n;\overline{\theta}\right) = \sum_{i=1}^{k} \ln f_1\left(x_i,\overline{\theta}\right) + \sum_{i=k+1}^{n} \ln f_2\left(x_i,\overline{\theta}\right) =$$
$$= L_1\left(x_1,\ldots,x_k\right) + L_2\left(x_{k+1},\ldots,x_n\right) =$$
$$= k\ln\lambda_1 + (n-k)\ln\lambda_2 - (n-k)\left(\lambda_1 - \lambda_2\right)x_0 -$$
$$-\lambda_1 \sum_{i=1}^{k} x_i - \lambda_2 \sum_{i=k+1}^{n} x_i \qquad (1)$$

Решающее правило запишем в виде

$$\rho = \max_{\substack{x_k, \\ \overline{k=2,n-3}}} \max_{\underline{\Theta}_k} \ln L\left(x_1,\ldots,x_n;\overline{\theta_k}\right). \qquad (2)$$

Вектор параметров $\overline{\theta}_k$ находим из решения системы уравнений:

$$\begin{cases} \dfrac{\partial\Lambda}{\partial\lambda_1} = 0; \\ \dfrac{\partial\Lambda}{\partial\lambda_2} = 0, \end{cases}$$

или

$$\begin{cases} \dfrac{\partial \Lambda}{\partial \lambda_1} = \dfrac{k}{\lambda_1} - (n-k)x_0 - \sum\limits_{i=1}^{k} x_i = 0; \\ \dfrac{\partial \Lambda}{\partial \lambda_2} = \dfrac{n-k}{\lambda_2} + (n-k)x_0 - \sum\limits_{i=k+1}^{n} x_i = 0, \end{cases}$$

Решение системы уравнений имеет вид

$$\lambda_1 = \frac{k}{(n-k)x_0 + \sum\limits_{i=1}^{k} x_i}, \; \lambda_2 = \frac{n-k}{-(n-k)x_0 + \sum\limits_{i=k+1}^{n} x_i},$$

или

$$\lambda_1 = \frac{p_k}{(1-p_k)x_0 + \overline{x_k}}, \; \lambda_2 = \frac{1}{\overline{x_{n-k}} - x_0}, \tag{3}$$

где $p_k = \dfrac{k}{n}$, $\overline{x_k} = \sum\limits_{i=1}^{k} x_i$, $\overline{x_{n-k}} = \sum\limits_{i=k+1}^{n} x_i$.

Таким образом, основные этапы вычислительной процедуры восстановления сплайн-экспоненциального распределения состоят в следующем:

1. Считая, что узел склеивания функции распределения x_0 совпадает с одной из вариант x_k , $k = \overline{2, n-3}$, вычисляют для каждого $x_0 = x_k$ оценки параметров $\overline{\hat{\theta}_k} = \{\hat{\lambda}_{1k}, \hat{\lambda}_{2k}\}$ по формуле (3).

2. Для каждого узла $x_0 = x_k$, $k = \overline{2, n-3}$ вычисляют значение функции $\Lambda = \ln L\left(x_1, ..., x_n; \overline{\theta_k}\right)$ по формуле (1).

3. Реализуя решающее правило, основанное на максимизации функционала (2), определяют значения узла $\hat{x}_0 = x_r$ склеивания $r \in \overline{2, n-3}$, и оценок параметров $\hat{\lambda}_1 = \hat{\lambda}_{1r}$, $\hat{\lambda}_2 = \hat{\lambda}_{2r}$ распределения.

Для оценки найденных параметров распределения найдем дисперсионно-ковариационную матрицу

$$DC = \begin{Vmatrix} D\{\lambda_1\} & \operatorname{cov}(\lambda_1, \lambda_2) \\ \operatorname{cov}(\lambda_2, \lambda_1) & D\{\lambda_2\} \end{Vmatrix}.$$

Для ее нахождения используем информационную матрицу

$$I = \begin{Vmatrix} \dfrac{\partial^2 \Lambda}{\partial \lambda_1{}^2} & \dfrac{\partial^2 \Lambda}{\partial \lambda_1 \lambda_2} \\ \dfrac{\partial^2 \Lambda}{\partial \lambda_2 \lambda_1} & \dfrac{\partial^2 \Lambda}{\partial \lambda_2{}^2} \end{Vmatrix} = \begin{Vmatrix} -\dfrac{k}{\lambda_1{}^2} & 0 \\ 0 & -\dfrac{n-k}{\lambda_2{}^2} \end{Vmatrix}.$$

Тогда

$$DC = -I^{-1} = \begin{Vmatrix} \dfrac{\lambda_1{}^2}{k} & 0 \\ 0 & \dfrac{\lambda_2{}^2}{n-k} \end{Vmatrix}.$$

Таким образом, дисперсии оценок параметров равны

$$D\{\lambda_1\} = \frac{\lambda_1{}^2}{k}, \; D\{\lambda_2\} = \frac{\lambda_2{}^2}{n-k}.$$

Тогда доверительные границы для параметров λ_1, λ_2 с доверительной вероятностью $1-\alpha$ имеют вид

$$\lambda_{н,в}(\lambda_1) = \hat{\lambda}_1 \mp t_{\upsilon,\alpha/2}\sqrt{D(\hat{\lambda}_1)}, \; \lambda_{н,в}(\lambda_2) = \hat{\lambda}_2 \mp t_{\upsilon,\alpha/2}\sqrt{D(\hat{\lambda}_2)},$$

где $t_{\upsilon,\alpha/2}$ – квантиль распределения Стьюдента с числом степеней свободы $\nu = n-2$.

Если количество вариант n > 30, то можно квантиль $t_{\upsilon,\alpha/2}$ заменить квантилем нормального распределения $u_{\alpha/2}$.

Теперь перейдем к восстановлению сплайн-экспоненциального распределения с двумя узлами.

1. Считая, что узел склеивания функции распределения x_0 совпадает с одной из вариант x_k, $k = \overline{2, n-3}$, вычисляют для каждого $x_0 = x_k$ оценки параметров $\overline{\hat{\theta}_k} = \{\hat{\lambda}_{1k}, \hat{\lambda}_{2k}\}$ по формуле (3).

2. Для каждого узла $x_0 = x_k$, $k = \overline{2, n-3}$ вычисляют значение функции $\Lambda = \ln L(x_1,..., x_n; \overline{\theta_k})$ по формуле (1).

3. Реализуя решающее правило, основанное на максимизации функционала (2), определяют значения узла $\hat{x}_0 = x_r$ склеивания $r \in \overline{2, n-3}$, и оценок параметров $\hat{\lambda}_1 = \hat{\lambda}_{1r}$, $\hat{\lambda}_2 = \hat{\lambda}_{2r}$ распределения.

4. Для найденных параметров сплайн-экспоненциального

распределения с одним узлом вычисляют значения функций $L_1(x_1,...,x_k)$ и $L_2(x_{k+1},...,x_n)$.

5. Находят $\min\{L_1(x_1,...,x_k), L_2(x_{k+1},...,x_n)\}$.

6. Если $\min\{L_1(x_1,...,x_k), L_2(x_{k+1},...,x_n)\} = L_1(x_1,...,x_k)$, то п.1–3 повторяют для выборки $\{x_1,...,x_k\}$. В противном случае – для выборки $\{x_{k+1},...,x_n\}$.

Для оценки найденных параметров распределения найдем дисперсионно-ковариационную матрицу

$$DC = \begin{Vmatrix} D\{\lambda_1\} & \operatorname{cov}(\lambda_1,\lambda_2) & \operatorname{cov}(\lambda_1,\lambda_3) \\ \operatorname{cov}(\lambda_2,\lambda_1) & D\{\lambda_2\} & \operatorname{cov}(\lambda_2,\lambda_3) \\ \operatorname{cov}(\lambda_3,\lambda_1) & \operatorname{cov}(\lambda_3,\lambda_2) & D\{\lambda_3\} \end{Vmatrix}.$$

Для ее нахождения используем информационную матрицу

$$I = \begin{Vmatrix} \dfrac{\partial^2\Lambda}{\partial\lambda_1^2} & \dfrac{\partial^2\Lambda}{\partial\lambda_1\lambda_2} & \dfrac{\partial^2\Lambda}{\partial\lambda_1\lambda_3} \\ \dfrac{\partial^2\Lambda}{\partial\lambda_2\lambda_1} & \dfrac{\partial^2\Lambda}{\partial\lambda_2^2} & \dfrac{\partial^2\Lambda}{\partial\lambda_2\lambda_3} \\ \dfrac{\partial^2\Lambda}{\partial\lambda_3\lambda_1} & \dfrac{\partial^2\Lambda}{\partial\lambda_3\lambda_2} & \dfrac{\partial^2\Lambda}{\partial\lambda_3^2} \end{Vmatrix} = \begin{Vmatrix} -\dfrac{s}{\lambda_1^2} & 0 & 0 \\ 0 & -\dfrac{k-s}{\lambda_2^2} & 0 \\ 0 & 0 & -\dfrac{n-k}{\lambda_3^2} \end{Vmatrix},$$

где

$$\Lambda = \ln L\left(x_1,...,x_n;\overline{\theta}\right) =$$

$$= \sum_{i=1}^{s} \ln f_1\left(x_i,\overline{\theta}\right) + \sum_{i=s+1}^{k} \ln f_2\left(x_i,\overline{\theta}\right) + \sum_{i=k+1}^{n} \ln f_3\left(x_i,\overline{\theta}\right) =$$

$$= s\ln\lambda_1 + (k-s)\ln\lambda_2 + (n-k)\ln\lambda_3 - (n-s)\left(\lambda_1 - \lambda_2\right)x_0 -$$

$$-(n-k)\left(\lambda_2 - \lambda_3\right)x_1 - \lambda_1\sum_{i=1}^{s} x_i - \lambda_2\sum_{i=s+1}^{k} x_i - \lambda_3\sum_{i=k+1}^{n} x_i .$$

Тогда

$$DC = -I^{-1} = \begin{Vmatrix} \frac{\lambda_1^2}{s} & 0 & 0 \\ 0 & \frac{\lambda_2^2}{k-s} & 0 \\ 0 & 0 & \frac{\lambda_3^2}{n-k} \end{Vmatrix}.$$

Таким образом, дисперсии оценок параметров равны

$$D\{\lambda_1\} = \frac{\lambda_1^2}{s},\ D\{\lambda_2\} = \frac{\lambda_2^2}{k-s},\ D\{\lambda_3\} = \frac{\lambda_3^2}{n-k}.$$

Тогда доверительные границы для параметров λ_1, λ_2, λ_3 с доверительной вероятностью $1-\alpha$ имеют вид

$$\lambda_{н,в}(\lambda_1) = \hat{\lambda}_1 \mp t_{\upsilon,\alpha/2}\sqrt{D(\hat{\lambda}_1)},\ \lambda_{н,в}(\lambda_2) = \hat{\lambda}_2 \mp t_{\upsilon,\alpha/2}\sqrt{D(\hat{\lambda}_2)},$$

$$\lambda_{н,в}(\lambda_3) = \hat{\lambda}_3 \mp t_{\upsilon,\alpha/2}\sqrt{D(\hat{\lambda}_3)},$$

где $t_{\upsilon,\alpha/2}$ – квантиль распределения Стьюдента с числом степеней свободы $\nu = n-2$.

Если количество вариант n > 30, то можно квантиль $t_{\upsilon,\alpha/2}$ заменить квантилем нормального распределения $u_{\alpha/2}$.

Описанная вычислительная схема определяет итерационную процедуру восстановления сплайн-распределения, которая была реализована к оценке надежности систем гидропривода, установленных на зерноуборочных комбайнах.

На рис. 1 представлена эмпирическая и теоретическая функции распределения времени непрерывной работы до отказа агрегата.

Реализация описанной вычислительной процедуры к данным по отказам позволяет получить, что с достоверностью 0,92 справедливо сплайн-экспоненциальное распределение с такими параметрами:

$\lambda_1 = 0{,}5777 \cdot 10^{-3}$; $\lambda_2 = 0{,}4147 \cdot 10^{-3}$; $\lambda_3 = 1{,}7512 \cdot 10^{-3}$ и узлами склеивания $T[час]$: $T_1 = 590$; $T_2 = 1020$, где λ [1/час].

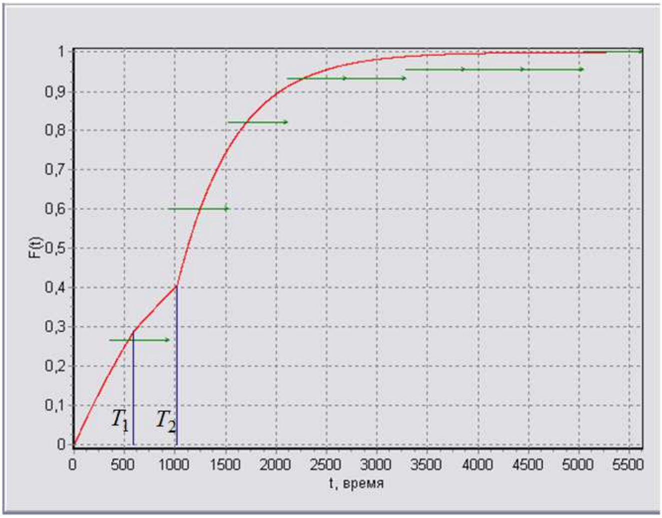

Рисунок 1 – Эмпирическая и теоретическая функции распределения времени до отказа агрегата

Выводы. Предложенная процедура позволяет осуществлять восстановление сплайн-экспоненциального распределения с двумя узлами. При этом за счет итеративного характера та же процедура может быть использована для восстановления сплайн-экспоненциального распределения с k узлами.

Библиографические ссылки

1. Приставка А. Ф. Сплайн-распределения в статистическом анализе. Д.: ДГУ, 1990. 150 с.

2. Приставка А. Ф., Рожков А. И., Малаховская Н. Л., Лойко Л. Л. ОСТ 5430011. Правила определения оценок, доверительных границ параметров и теоретической функции распределения Вебулла с одним узлом. Москва: МГА, 1983. Т. 59. 44 с.

3. Приставка А. Ф., Виниченко Л. Ф. Восстановление сплайн-экспоненциального распределения с двумя узлами. Депропетровск: ДГУ, 1987. С.100–104.

Надійшла до редколегії 30.06.2017.