

TÁBIYIY HÁM TEXNIKALÍQ ILIMLER

Fizika.Matematika.Texnika.Informatika.

ОЦЕНИВАНИЕ СРЕДНЕГО ОСТАТОЧНОГО ВРЕМЕНИ ЖИЗНИ В СПЕЦИАЛЬНОЙ РЕГРЕССИОННОЙ МОДЕЛИ СЛУЧАЙНОГО ЦЕНЗУРИРОВАНИЯ С ДВУХ СТОРОН

Ф.Абдикаликов – доктор философии по физико-математическим наукам

Каракалпакский государственный университет имени Бердаха

Институт математики имени В.И.Романовского Академии наук Республики Узбекистан

Р.Сейтахов – стажер преподаватель

Ж.Жартыбаев – магистрант

Каракалпакский государственный университет имени Бердаха

Таянч сўзлар: даражали баҳо, информатив цензурираниш, регрессия, ўртача колдик умр, ядровий баҳо.

Ключевые слова: степенная оценка, информативное цензурирование, регрессия, средние остаточное время жизни, ядерная оценка.

Key words: power estimator, informative censoring, regression, mean residual lifetime, kernel estimator.

Введение. На практике, в экспериментальных случаях интересующая нас случайная величина (с.в.) Z могут оказаться под влиянием другой сопровождающей величины-ковариаты (фактора влияния, регрессора) X . В медицине доза лекарства, в инженерии такие факторы как температура, давление, напряжение тока могут оказать влияние на продолжительность жизни индивидуума или безотказной работы технического устройства. Статистическая задача состоит в оценивании условной функции выживания или ее среднего остаточного времени жизни в моделях регрессии с цензурированными наблюдениями. Характерная особенность этих задач состоит в том, что для одной и той же оцениваемой характеристики могут быть предложены различные оценки в зависимости от структуры рассматриваемой модели цензурированных наблюдений. Такую задачу впервые при случайном цензурировании справа рассматривал Беран [3], предложив в качестве оценки для функции выживания обобщенный аналог известной множительной оценки Каплана-Мейера. В работах [5,6] исследована в основном оценка множительной структуры. В [1] были предложены и исследованы оценки степенной структуры в различных моделях регрессии случайного цензурирования. В настоящей работе рассматривается специальная модель регрессии при случайном цензурировании с двух сторон и в ней так же исследуется оценки степенной структуры.

Характеризация модели и вспомогательные результаты. Рассмотрим ситуацию, когда с.в. Z подвергается случайному цензурированию с двух сторон парой с.в. (L, Y) . Предполагается, что компоненты вектора (L, Z, Y) независимые с.в. при заданной ковариате

$X = x \in [0, 1]$. Пусть $\{(L_i, Z_i, Y_i), i = \overline{1, n}\}$ - независимые реализации вектора (L, Z, Y) в n экспериментах и $0 \leq x_1 \leq \dots \leq x_n \leq 1$ - фиксированные точки дизайна ковариаты X . Наблюдается следующая выборка объема n :

$$S^{(n)} = \{(\xi_i, \delta_{0i}, \delta_{1i}, \delta_{2i}), i = \overline{1, n}\}, \text{ где}$$

$\xi_i = \max\{L_i, \min(Z_i, Y_i)\}$, $\delta_{0i} = I(\min(Z_i, Y_i) < L_i)$, $\delta_{1i} = I(L_i \leq Z_i \leq Y_i)$, $\delta_{2i} = I(L_i \leq Y_i < Z_i)$ и $I(A)$ - индикатор события A . Пусть F_{x_i} , G_{x_i} , K_{x_i} и H_{x_i} - условные ф.р. с.в. Z_i , Y_i , L_i и ξ_i при заданной ковариате $X = x_i$ являются непрерывными. Далее для упрощения записей x_i обозначим через x . Заметим, что интересующие нас с.в. Z_i наблюдаемы лишь при $\delta_{1i} = 1$. Здесь рассматривается специальная модель, т.е.

F_x , G_x и K_x удовлетворяют следующим степенным представлениям при всех $t \geq 0$:

$$1 - G_x(t) = (1 - F_x(t))^{\theta_x+1}, \quad 1 - K_x(t) = (1 - F_x(t))^{\theta_x},$$

где θ_x - положительный неизвестный мешающий параметр.

Тогда оценка для условной функции выживания определяется как [2]:

$$1 - F_{xh}(t) = \left[1 - (H_{xh}(t))^{1/2} \right]^{\gamma_{xh}}, \quad t \geq 0, \quad (1)$$

где

$$H_{xh}(t) = \sum_{i=1}^n \omega_{ni}(x; h_n) I(\xi_i \leq t),$$

$p_{mxh} = \sum_{i=1}^n \omega_{ni}(x; h_n) I(\delta_{mi} = 1)$, $m = 1, 2$, $\gamma_{xh} = 2p_{1xh}$ явля-

ются сглаженными оценками для $H_x(t)$,

$p_{mx} = P(\delta_{mx} = 1)$, $m = 1, 2$ и $\gamma_x = 2p_{1x}$ используя веса Гессера-Мюллера [4]:

$$\omega_{ni}(x; h_n) = \int_{x_{i-1}}^{x_i} \frac{1}{h_n} \pi\left(\frac{x-y}{h_n}\right) dy \left(\int_0^{x_n} \frac{1}{h_n} \pi\left(\frac{x-y}{h_n}\right) dy \right)^{-1}, \quad i = \overline{1, n}.$$

Здесь $x_0 = 0$, π - известная функция плотности и $\{h_n\}$ - последовательность положительных постоянных такое, что $h_n \rightarrow 0$ при $n \rightarrow \infty$.

Далее нами будут использованы обозначения:

$$\underline{\Delta}_n = \min_{1 \leq i \leq n} (x_i - x_{i-1}), \quad \bar{\Delta}_n = \max_{1 \leq i \leq n} (x_i - x_{i-1}),$$

$$\|\pi\|_2^2 = \int_{-\infty}^{\infty} \pi^2(y) dy, \quad m_v(\pi) = \int_{-\infty}^{\infty} y^v \pi(y) dy, \quad v = 1, 2.$$

Легко видеть, что ввиду условий (1) в рассматриваемой модели

$$T_{F_x} = T_{H_x} = \inf\{t \geq 0: H_x(t) = 1\},$$

$$\tau_{F_x} = \tau_{H_x} = \sup\{t \geq 0: H_x(t) = 0\}.$$

Рассмотрим условия:

$$(Y1) \quad \text{При } n \rightarrow \infty, \quad x_n \rightarrow 1, \quad \bar{\Delta}_n = O(n^{-1}),$$

$$\bar{\Delta}_n - \underline{\Delta}_n = o(n^{-1});$$

(Y2) $\pi(y)$ - плотность с носителем $[-M, M]$, при некотором $M > 0$, $m_1(\pi) = 0$ и π удовлетворяет условию Липшица первого порядка;

$$(Y3) \quad \ddot{H}_x(t) = \frac{\partial^2}{\partial x^2} H_x(t), \quad H_x''(t) = \frac{\partial^2}{\partial t^2} H_x(t),$$

$\dot{H}_x'(t) = \frac{\partial^2}{\partial x \partial t} H_x(t)$ существуют и непрерывны для $(t, x) \in [\tau, T] \times [0, 1]$, где $\tau_{H_x} < \tau < T < T_{H_x}$;

$$(Y4) \quad \ddot{p}_{mx} = \frac{d^2}{dx^2} p_{mx}, \quad m=1, 2 \text{ существуют и непре-}$$

рывны для $x \in [0, 1]$.

Далее используем также обозначение $r^{-1} = \sup_{\tau \leq t \leq T} [(H_x(t))^{1/2} - H_x(t)]^{-1}$.

В [2] проведено исследование оценки (1). Мы представляем некоторые из полученных результатов.

Теорема 1. [2]. Пусть выполнены условия (Y1), (Y2), а для функций H_x , p_{0x} , p_{1x} условия (Y3), (Y4), $r > 0$ и при $n \rightarrow \infty$, $h_n \rightarrow 0$, $\frac{\ln n}{nh_n} = o(1)$,

$\frac{nh_n^5}{\ln n} = O(1)$. Тогда для $t \in [\tau, T]$:

$$F_{xh}(t) - F_x(t) = \sum_{i=1}^n \omega_{ni}(x; h_n) \Psi_{ix}(\xi_i, \delta_{1i}, \delta_{2i}) + r_n(t, x),$$

где

$$\begin{aligned} \Psi_{ix}(\xi_i, \delta_{1i}, \delta_{2i}) &= \\ &= (1 - F_x(t)) \left\{ \frac{p_{1x}}{(H_x(t))^{1/2} - H_x(t)} (I(\xi_i \leq t) - H_x(t)) - \right. \\ &\quad \left. - 2 \log[1 - (H_x(t))^{1/2}] (\delta_{1i} - p_{1x}) \right\}, \\ \sup_{\tau \leq t \leq T} |r_n(t, x)| &= O((nh_n)^{-1} \ln n). \end{aligned}$$

Теорема 2. [2]. В условиях теоремы 1 следует равномерно сильная состоятельность оценки (1) при $n \rightarrow \infty$:

$$\sup_{\tau \leq t \leq T} |F_{xh}(t) - F_x(t)| = O((nh_n)^{-1/2} (\ln n)^{1/2}).$$

Теорема 3. [2]. Пусть выполнены условия (Y1), (Y2), а функции H_x , p_{0x} , p_{1x} удовлетворяют условиям (Y3), (Y4) и $r > 0$. Тогда

(А) Если $nh_n^5 = o(1)$ и $(nh_n)^{-1/2} \ln n = o(1)$, тогда при $n \rightarrow \infty$ и $t \in [\tau, T]$

$$(nh_n)^{1/2} (F_{xh}(t) - F_x(t)) \xrightarrow{d} N(0, \sigma_x^2(t)).$$

(В) Если $h_n = Cn^{-1/5}$ при некотором $C > 0$, тогда при $n \rightarrow \infty$ и $t \in [\tau, T]$

$$(nh_n)^{1/2} (F_{xh}(t) - F_x(t)) \xrightarrow{d} N(a_x(t), \sigma_x^2(t)),$$

где

$$a_x(t) = \frac{1}{2} (1 - F_x(t)) \{A_x(t) \ddot{H}_x(t) - B_x(t) \ddot{p}_{1x}\} m_2(\pi) C^{5/2},$$

$$\sigma_x^2(t) = \|\pi\|_2^2 (1 - F_x(t))^2$$

$$\{A_x^2(t) H_x(t) (1 - H_x(t)) + B_x^2(t) p_{1x} (1 - p_{1x})\},$$

$$A_x(t) = \frac{p_{1x}}{(H_x(t))^{1/2} - H_x(t)},$$

$$B_x(t) = 2 \log[1 - (H_x(t))^{1/2}].$$

В условиях теоремы 3 установлена слабая сходимость процесса $\{W_{nx}(t) = (nh_n)^{1/2} (F_{xh}(t) - F_x(t)), \tau \leq t \leq T\}$, в пространстве $D[\tau, T]$ к гауссовскому процессу. Рассмотрим $W_x(t)$ гауссовский процесс с нулевым средним и ковариационной функций при $s, t \in [\tau, T]$:

$$Q_{st} = \|\pi\|_2^2 (1 - F_x(s))(1 - F_x(t)) \cdot$$

$$\cdot \{A_x(s) A_x(t) [H_x(\min(s, t)) - H_x(s) H_x(t)] + B_x(s) B_x(t) p_{1x} (1 - p_{1x})\},$$

где функций $A_x(t)$ и $B_x(t)$ определены в теореме 3, а гауссовский процесс $\tilde{W}_x(t)$ имеет среднее $a_x(t)$ и ковариацию Q_{st} .

Теорема 4. [2]. Пусть выполнены условия теоремы 3. Тогда

(А)' В условиях (А): $W_{nx}(t) \rightarrow W_x(\cdot)$ в $D[\tau, T]$;

(В)' В условиях (В): $W_{nx}(\cdot) \rightarrow \tilde{W}_x(\cdot)$ в $D[\tau, T]$.

Основные результаты. В данной работе рассматривается задача оценивания условного среднего остаточного времени жизни

$$\mu_x(t) = M[(Z_x - t)/Z_x > t] = (1 - F_x(t))^{-1} \cdot \int_t^{+\infty} (1 - F_x(s)) ds, \quad (2)$$

Оценку для $\mu_x(t)$ можем построить методом подстановки оценки (1) вместо $1 - F_x(t)$ в формуле (2). Однако, из теорем 2 следует, что оценка (1) является равномерно состоятельной в некотором $[\tau, T]$. Поэтому сперва введём следующую урезанную версию функции (2):

$$\mu_x^T(t) = (1 - F_x(t))^{-1} \cdot \int_t^T (1 - F_x(s)) ds, \quad \tau < t < T, \quad (3)$$

Теперь функцию (3) оценим статистикой

$$\mu_{xh}^T(t) = (1 - F_{xh}(t))^{-1} \cdot \int_t^T (1 - F_{xh}(s)) ds, \quad \tau < t < T. \quad (4)$$

Справедливо следующее утверждение об асимптотической нормальности оценки (1), т.е. аналог теоремы 3.

Теорема 5. Пусть выполнены условия теоремы 3.

(С) Если $nh_n^5 = o(1)$ и $(nh_n)^{-1/2} \ln n = o(1)$ при $n \rightarrow \infty$, то для $t \in [\tau, T]$

$$(nh_n)^{1/2} (\mu_{xh}^T(t) - \mu_x^T(t)) \xrightarrow{d} N(0, \beta_x^2(t));$$

(D) Если $h_n = Cn^{-1/5}$ при некотором $C > 0$, то при $n \rightarrow \infty$ и $t \in [\tau, T]$

$$(nh_n)^{1/2}(\mu_{xh}^T(t) - \mu_x^T(t)) \xrightarrow{d} N(\alpha_x(t), \beta_x^2(t)),$$

где $\alpha_x(t)$ и $\beta_x^2(t)$ определяются формулами

$$\alpha_x(t) = \frac{1}{2(1-F_x(t))} \left\{ \int_t^T a_x(s) ds - a_x(t) \int_t^T (1-F_x(s)) ds \right\} m_2(\pi) C^{5/2},$$

$$\beta_x^2(t) = \|\pi\|_2^2 \frac{1}{(1-F_x(t))^2} \int_t^T \left(\int_t^T (1-F_x(s)) ds \right)^2 d\gamma_x(s),$$

$$\gamma_x(t) = A_x^2(t) H_x(t) (1-H_x(t)) + B_x^2(t) p_{1x} (1-p_{1x}).$$

Доказательство. При помощи нетрудных выкладок и теоремы 1, имеем

$$\mu_{xh}^T(t) - \mu_x^T(t) = \int_t^T \left[\frac{1-F_{xh}(s)}{1-F_{xh}(t)} - \frac{1-F_x(s)}{1-F_x(t)} \right] ds = \quad (5)$$

$$= M_{nx}(t) + \sum_{k=1}^4 Q_{nx}^{(k)}(t),$$

где

$$M_{nx}(t) = \sum_{i=1}^n \omega_{ni}(x; h_n)$$

$$\left[-\frac{1}{1-F_x(t)} \int_t^T \Psi_{sx}(\xi_i, \delta_{1i}, \delta_{2i}) ds + \frac{\Psi_{tx}(\xi_i, \delta_{1i}, \delta_{2i})}{(1-F_x(t))^2} \int_t^T (1-F_x(s)) ds \right],$$

$$Q_{nx}^{(1)}(t) = -\frac{1}{1-F_x(t)} \int_t^T r_n(s, x) ds,$$

$$Q_{nx}^{(2)}(t) = \frac{r_n(t; x)}{(1-F_x(t))^2} \int_t^T (1-F_x(s)) ds,$$

$$Q_{nx}^{(3)}(t) = \frac{(F_{xh}(t) - F_x(t))}{(1-F_x(t))^2} \int_t^T (F_{xh}(s) - F_x(s)) ds,$$

$$Q_{nx}^{(4)}(t) = \frac{(F_{xh}(t) - F_x(t))^2}{(1-F_x(t))^3} \int_t^T (1-F_{xh}(s)) ds.$$

Для $Q_{nx}^{(1)}$ и $Q_{nx}^{(2)}$ используем теорему 1, для $Q_{nx}^{(3)}$ и $Q_{nx}^{(4)}$ теорему 2. Тогда видим, что все эти остаточные члены представления (5) с вероятностью 1 равномерно по отрезку $[\tau, T]$ имеют порядок $O((nh_n)^{-1} \ln n)$. Теперь утверждения (C) и (D) теоремы следует из соответствующих утверждений теоремы 3. Теорема 5 доказана.

В условиях теоремы 5 также возможно установить слабую сходимости процесса $\{V_{nx}(t) = (nh_n)^{1/2}(\mu_{xh}^T(t) - \mu_x^T(t)), \tau \leq t \leq T\}$, в пространстве $D[\tau, T]$ к гауссовскому процессу. Рассмотрим $V_x(t)$ гауссовский процесс с нулевым средним и ковариационной функцией при $s, t \in [\tau, T]$:

$$\tilde{Q}_{st} = \|\pi\|_2^2 \frac{1}{(1-F_x(s))(1-F_x(t))} \int_{\min(s,t)}^T \left(\int_t^T (1-F_x(s)) ds \right)^2 d\gamma_x(u),$$

где функция $\gamma_x(t)$ определена в теореме 5, а гауссовский процесс $\tilde{V}_x(t)$ имеет среднее $\alpha_x(t)$ и ковариацию Q_{st} .

Теорема 6. Пусть выполнены условия теоремы 5. Тогда

(C)' В условиях (C): $V_{nx}(t) \rightarrow V_x(\cdot)$ в $D[\tau, T]$;

(D)' В условиях (D): $V_{nx}(\cdot) \rightarrow \tilde{V}_x(\cdot)$ в $D[\tau, T]$.

Доказательство. Нетрудно видеть, что процесс $V_{nx}(t)$ является линейным функционалом процесса $W_{nx}(t)$:

$$V_{nx}(t) = -\frac{1}{1-F_x(t)} \int_t^T W_{nx}(s) ds + \frac{W_{nx}(t)}{(1-F_{xh}(t))(1-F_x(t))} \int_t^T (1-F_{xh}(s)) ds.$$

Следовательно, утверждения теоремы 6 вытекают из соответствующих утверждений теоремы 4. Детали опускается. Теорема 6 доказана.

Литература

1. Абдушукуров А., Абдикаликов Ф. Исследования степенных оценок функций выживания в моделях регрессии. LAMBERT Academic Publishing. 2012. -С.89.
2. Абдикаликов Ф.А., Эрисбаев С., Сейтахов Р. Свойства оценок в специальной регрессионной модели случайного цензурирования с двух сторон. Международной научно-практической конференции «Актуальные задачи математического моделирования и информационных технологий». -Нукус: Май 2-3, 2023.
3. Beran R. Nonparametric regression with randomly censored survival data. Technical Report. Univ. California. Berkeley. 1981. -P. 19.
4. Gasser T., Müller H. Kernel estimation of regression function. Lecture Notes in Mathematics. Springer. -New York: 1977.
5. Van Keilegom I., Veraverbeke N. Uniform strong convergence for the conditional Kaplan-Meier estimator and its quantiles. // Commun. Statist.: Th.&Meth. 1996. Vol. 25. - P. 2251-2265.
6. Van Keilegom I., Veraverbeke N. Estimation and bootstrap with censored data in fixed design nonparametric regression. // Ann. Inst. Statist. Math. 1997. Vol. 49. - №3. - P.467-491. <http://dx.doi.org/10.1023/A:1003166728321>.

РЕЗЮМЕ. Ушбу маколада умр давомийлиги ҳар икки тала томонидан тасодифий цензурланганда махсус регрессия моделида ўртача қолдиқ умри баҳолашни ўрганамиз. Баҳонинг асимптотик нормаллиги ва эмперик тасодифий жараёнинг Скороход фазосида Гаусс жараёнига суғ яқинлашишлари исботланган.

РЕЗЮМЕ. В статье мы исследуем оценку среднего остаточного времени жизни в специальной регрессионной модели, когда время жизни подвергается случайной цензурированию с двух сторон. Доказана асимптотическая нормальность оценок и слабая сходимости эмпирического случайного процесса к гауссовскому процессу в пространстве Скорохода.

SUMMARY. In this article, we investigate the estimator of the mean residual lifetime in special regression model, when lifetimes are randomly censored on both sides. The asymptotic normality are proved of the estimators and the weakly convergence of an empirical random process to a Gaussian process in Skorokhod space.