

TÁBIIYÚ HÁM TEXNIKALÍQ ILIMLER

Texnikalíq ilimler

АСИМПТОТИЧЕСКАЯ НОРМАЛЬНОСТЬ ОЦЕНОК ФУНКЦИИ СРЕДНЕГО ОСТАТОЧНОГО
ВРЕМЕНИ БЕЗОТКАЗНОЙ РАБОТЫ В РЕГРЕССИОННОЙ МОДЕЛИ СЛУЧАЙНОГО
ЦЕНЗУРИРОВАНИЯ СПРАВА

Ф.Абдикаликов – доктор философии по физико-математическим наукам

П.Насыров – ассистент преподаватель

Каракалпакский государственный университет имени Бердаха

Таянч сўзлар: тасодифий цензурланиш, даражали баҳо, регрессия, колдик умр давомийлиги функцияси.

Ключевые слова: случайное цензурирование, степенная оценка, регрессия, функция среднего остаточного времени жизни.

Key words: random censoring, power estimator, regression, mean residual life function.

Функция среднего остаточного времени безотказной работы (ФСОВБР) является одной из важных характеристик в анализе данных типа времени жизни. Подробный обзор литературы и ряд новых результатов для различных оценок ФСОВБР в случае полной и цензурированной данных имеется в [3]. Значительный интерес представляет оценивание этой характеристики для зависимых цензурированных данных. В данной работе приведены асимптотические свойства оценки условной ФСОВБР в модели случайного цензурирования справа, в которой зависимость между наблюдениями выражается через регрессию с фиксированными ковариатами.

Пусть случайная величина (с.в.) Z зависит от реализации с.в. X . Например с.в. Z может соответствовать времени жизни пациента, а X дозе лекарства или Z соответствует времени безотказной работы некоторого технического устройства, которое может подвергаться давлению, температуре или каким-нибудь другим факторам и их можно принять за реализации X . Однако, как правило в таких случаях с.в. Z полностью ненаблюдаема и может быть цензурирована. В подобных ситуациях возникает задача оценивания условной функции выживания и её характеристик, скажем ФСОВБР при заданном значении фактора влияния (

ковариаты X). Рассмотрим ситуацию, когда с.в. Z подвергается случайному цензурированию справа с.в. Y . Пусть $\{(Z_i, Y_i), i = \overline{1, n}\}$ - независимые реализации случайного вектора (Z, Y) в n экспериментах и $0 = x_0 \leq x_1 \leq \dots \leq x_n \leq 1$ - фиксированные точки дизайна ковариаты X . Наблюдается следующая выборка объема n : $S^{(n)} = \{(\xi_1, \delta_1, X_1), \dots, (\xi_n, \delta_n, X_n)\}$, где $\xi_i = \min(Z_i, Y_i)$ и $\delta_i = I(Z_i \leq Y_i)$ индикаторы, указанных в скобках событий. Пусть F_{x_i} , G_{x_i} и H_{x_i} означают условные функции распределения (ф.р.) с.в. Z_i , Y_i и ξ_i при заданном $X = x_i$ соответственно. Далее, для любой точки $x \in [0, 1]$, обозначим через F_x , G_x и H_x соответственно ф.р. с.в. Z_x , Y_x и $\xi_x = \min(Z_x, Y_x)$ и пусть $\delta_x = I(Z_x \leq Y_x)$. Легко видеть, что $H_x(t) = 1 - (1 - F_x(t))(1 - G_x(t))$, $t \geq 0$. Условная ФСОВБР определяется так

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

т.е. является средним значением по условному распределению

$$F_x(s/t) = P(Z_x - t \leq s/Z_x > t) = \frac{F_x(t+s) - F_x(t)}{1 - F_x(t)}, \quad t, s \geq 0.$$

Очевидно, функцию (1) можно оценить методом подстановки вместо F_x её оценки. В данной работе используется оценка для F_x построенная и исследованная авторами [1,2]. Определим веса Гессера-Мюллера [4] формулами

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

Здесь весовая функция (ядро) π - заданная плотность вероятности и $\{h_n\}$ - последовательность положительных чисел, «ширина окна» стремящаяся к нулю при $n \rightarrow \infty$. Оценка F_{xh} для F_x определяется как [1,2]:

$$1 - F_{xh}(t) = [1 - H_{xh}(t)]^{\frac{1}{\tilde{H}_{xh}(t)}}, \quad t \geq 0, \quad (2)$$

где

$$\mathbb{R}_{xh}(t) = \frac{\tilde{\Lambda}_{xh}(t)}{\Lambda_{xh}(t)}, \quad \tilde{\Lambda}_{xh}(t) = \int_0^t \frac{d\tilde{H}_{xh}(u)}{1 - H_{xh}(u-)}, \quad \tilde{\Lambda}_{xh}(t) = \int_0^t \frac{d\tilde{H}_{xh}(u)}{1 - H_{xh}(u-)},$$

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

Далее нами будут использованы обозначения:

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

$$(II) \|\pi\|_2^2 = \int_{-\infty}^{\infty} \pi^2(y) dy, \quad m_1(\pi) = \int_{-\infty}^{\infty} y\pi(y) dy, \quad m_2(\pi) = \int_{-\infty}^{\infty} y^2\pi(y) dy.$$

В дальнейшем предполагаем выполненными условия:

$$(Y1) x_n \rightarrow 1, \quad \bar{\Delta}_n = O\left(\frac{1}{n}\right), \quad \bar{\Delta}_n - \underline{\Delta}_n = o\left(\frac{1}{n}\right) \text{ при } n \rightarrow \infty;$$

(Y2) π - плотность с компактным носителем $[-M, M]$, $M > 0$. А также $m_1(\pi) = 0$ и π удовлетворяет условию Липшица 1-порядка;

(Y3) Производные второго порядка $\ddot{H}_x(t) = \frac{\partial^2}{\partial x^2} H_x(t)$, $H_x''(t) = \frac{\partial^2}{\partial t^2} H_x(t)$, $\dot{H}_x'(t) = \frac{\partial^2}{\partial x \partial t} H_x(t)$ существуют и непрерывны для $(x, t) \in [0, 1] \times [\tau, T]$, где

$$\sup\{t \geq 0 : H_x(t) = 0\} = \tau_{H_x} < \tau < T < T_{H_x} = \inf\{t \geq 0 : H_x(t) = 1\}.$$

В работах [1,2] установлены следующие свойства оценки (2):

Теорема 1. [1,2] (равномерная сильная состоятельность). Пусть выполнены условия (Y1)-(Y3) и при $n \rightarrow \infty$, $\frac{\log n}{nh_n} = o(1)$, $\frac{nh_n^5}{\log n} = O(1)$. Тогда

$$\sup_{\tau \leq t \leq T} |F_{xh}(t) - F_x(t)| \stackrel{\text{п.н.}}{=} O\left(\left(\frac{\log n}{nh_n}\right)^{1/2}\right).$$

Теорема 2. [1,2] (асимптотическая нормальность). Пусть выполнены условия (Y1)-(Y3). Тогда

(A) если $nh_n^5 = o(1)$ и $(nh_n)^{-1}(\log n)^3 = o(1)$ при $n \rightarrow \infty$, то

$$(nh_n)^{1/2} (F_{xh}(t) - F_x(t)) \stackrel{D}{\Rightarrow} N(0, \sigma_x^2(t)),$$

(B) если $h_n = Cn^{-1/5}$ при некотором $C > 0$, то при $n \rightarrow \infty$

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

где

$$a_x(t) = \frac{1}{2}(1 - F_x(t)) \left\{ \int_0^t \frac{\ddot{H}_x(u) d\tilde{H}_x(u)}{(1 - H_x(u))^2} + \int_0^t \frac{d\ddot{H}_x(u)}{1 - H_x(u)} \right\} m_2(\pi) C^{5/2},$$

$$\sigma_x^2(t) = \|\pi\|_2^2 (1 - F_x(t))^2 \int_0^t \frac{d\tilde{H}_x(u)}{(1 - H_x(u))^2}, \quad \tilde{H}_x(t) = \int_0^t (1 - G_x(u-)) dF_x(u).$$

Оценку для $\mu_x(t)$ построим подстановкой оценки (2) вместо F_x в формуле (1). Однако, из теорем 1 следует, что оценка (2) является равномерно состоятельной в некотором замкнутом интервале $[\tau, T]$, где $\tau_{H_x} < \tau < T < T_{H_x}$. Поэтому сперва введём следующую “урезанную” версию функции (1):

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

Теперь функцию (3) оценим статистикой

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

Имеет место следующее утверждение об асимптотической нормальности оценки (4), т.е. аналог теоремы 2.

Теорема 3. Пусть выполнены условия теоремы 2.

(А) Если $nh_n^5 = o(1)$ и $(nh_n)^{-1}(\log n)^3 = o(1)$ при $n \rightarrow \infty$, то для $\tau < t < T$

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

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

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

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

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

$$\beta_x^2(t) = \|\pi\|_2^2 \frac{1}{(1 - F_x(t))^2} \int_t^T \left(\int_u^T (1 - F_x(s)) ds \right)^2 d\gamma_x(u), \quad \gamma_x(t) = \int_0^t \frac{d\tilde{H}_x(u)}{(1 - H_x(u))^2}.$$

Литература

1. Абдушукуров А.А., Абдикаликов Ф.А. Оценивание условной функции распределения при непараметрической регрессии с фиксированным дизайном. // Сб. "Статистические методы оценивания и проверка гипотез". – Пермь: Пермский госуниверситет. 2011. вып. 23. – С. 181-189.
2. Абдикаликов Ф.А. Непараметрическое оценивание условной функции выживания в регрессионной модели случайного цензурирования справа. // Вестник НУУз. 2013г №1. – С.55-57.
3. Сагидуллаев К.С., Абдушукуров А.А. Асимптотический анализ ядерных оценок, связанных со средним значением. LAP.Lambert Academic Publishing. 2011. – С. 132.
4. Gasser T., Müller H.G. Kernel estimation of regression function. Lecture Notes in Mathematics. Springer. -New York: 1977.

РЕЗЮМЕ

Мақолада ўртача қолдиқ умр давомийлиги функциясининг ўнг томондан тасодифий цензурланиш регрессия моделидаги баҳосининг асимптотик нормаллиги исботланган.

РЕЗЮМЕ

В статье доказана асимптотическая нормальность оценки для функции среднего остаточного времени жизни в модели регрессии со случайным цензурированием справа.

SUMMARY

In the article is proved an asymptotic normality of estimator for mean residual life function in regression model under random censoring from the right.

BERILGANLARNI QAYTA ISHLASH ALGORITMLARIDA O'LCHOV SHKALALARI VA TANLANMA FAYLLARINING O'RNI

E.A.Eshboyev - katta o'qituvchi

F.Y.Shodiyev - o'qituvchi

A.A.Bozorov - o'qituvchi

Qarshi davlat universiteti

Tayanch so'zlar: alomat, obyekt, sinf, munosabat, miqdoriy va nominal alomatlar, nominal shkala, tartib shkala, interval shkala, nisbiy shkala, tanlanma, klasterizatsiya, klassifikatsiya, identifikatsiya.

Ключевые слова: признак, объект, класс, отношение, количественные и номинальные признаки, номинальная шкала, порядковая шкала, интервальная шкала, относительная шкала, выборка, кластеризация, классификация, идентификация.

Key words: feature, object, class, relationship, quantitative and nominal features, nominal scale, ordinal scale, interval scale, relative scale, sampling, clustering, classification, identification.

Ushbu maqolada katta hajmdagi berilganlar bazasining tashkil etuvchilari bo'lgan berilgan (alamat) larning o'lchovlari va ularga mos shkalalarning qo'llanilishi aniq masalalar misolida bayon etiladi.

Ko'p hollarda tadqiqotchi o'zi tadqiq qilayotgan parametrlarni sonli o'lchash imkoniyatiga ega bo'lmaydi. Masalan, insonning biror narsaga munosabati, nimanidir afzal ko'rish darajasi kabilar. Bunday hollarda o'lchash usullari an'anaviy usullardan farq qiladi va obyektlarning sifat xususiyatlarini aks ettiruvchi belgilarga son qiymatlarni mos qo'yadigan ixtiyoriy usulni o'lchash deb qarash mumkin bo'ladi. Shuningdek belgilar va ularni aks ettiruvchi sifatlar o'rtasida o'zaro turg'un bog'liqliklar bo'lishi talab etiladi. Boshqacha qilib aytganda, sifat xususiyatli obyekt klasterizatsiyasini amalga oshirish uchun

shkalalash texnikasi usullaridan foydalanish zarur bo'ladi [1].

Bu kabi ishlar bilan olimlardan S.Rassel, P.Norvig, N.Ignatev, Sh.Madraximov, R.Usmanov, D.Saidovlarning ilmiy ishlarida batafsil tanishish mumkin.

Masalan, N.Ignatevning 2015-yil chop etilgan «Обобщенные оценки и локальные метрики объектов в интеллектуальном анализ данных» nomli monografiyasida: berilganlar tahlilida umumlashgan baholash, klassifikatsiya masalalarida lokal metrika, alomatlarning shajaraviy guruhlanishi kabi ishlar ilmiy asoslanib yoritilgan. N.Ignatev, R.Usmanov, Sh.Madraximovlar tomonidan 2018-yilda nashr ettirilgan «Berilganlarning intellektual tahlili» nomli o'quv qo'llanmada: o'lchov shkalalari, alomatlarning vaznlari,