Оценивание квантилей функции распределения с использованием полиномов Бернштейна

Тихов М.С.

Нижний Новгород, Россия
Нижегородский Государственный Университет им. Н.И.Лобачевского,
Институт Информационных Технологий, Математики и Механики,
Кафедра Теории Вероятностей и Анализа Данных
E-mail: tikhovm@mail.ru

Пусть $X_1, X_2, \ldots, X_n$ есть латентные (ненаблюдаемые) независимые и одинаково распределенные случайные величины с неизвестной непрерывной функцией распределения $F(x)$ и плотностью распределения $f(x) > 0, x \in (0,1)$, отрезок $[0,1]$ – носитель этого распределения, $u_i = i/n, i = 0, 1, \ldots, n$ – точки деления интервала $[0,1]$ и $W_i = I(X_i < u_i)$ есть индикатор события:$\{\ X_i < u_i\}$. Рассматривается задача оценивания квантиля порядка $0 < \lambda < 1$ функции распределения $F(x)$ на основе данных $(W_i, u_i, i = 0, \ldots, n)$. Такая задача возникает в биологии и называется зависимостью «доза-эффект» (см. [1], [2], [3]). Заметим, что $\mathbf{E}(W_i) = F(u_i)$.

Обычная практика оценивания $F(x)$ и ее квантилей состоит в использовании ядерных оценок [2]. В последнее время появилось большое количество работ, посвященных оцениванию функции распределения с использованием для этой цели полиномов $b_k(n,x) = C_n^k x^k (1-x)^{n-k}, k = 0, 1, \ldots, n, x \in (0,1)$. Так, в работах [4], [5] для полной выборки $X_1, X_2, \ldots, X_n$ предлагается использовать статистику $\hat{F}_n(x) = \sum\limits_{k=0}^{n} F_n(u_k) b_k(n,x)$ в качестве оценки функции распределения $F(x)$, где $F_n(x)$ есть эмпирическая функция распределения. Для бернуллиевской функции регрессии оценки с использованием полиномов Бернштейна изучены в работе [6], именно, в качестве оценки для $F(x)$ предложена статистика

$$F_n^*(x) = \sum_{k=0}^{n} W_k b_k(n,x).$$

Так как $\mathbf{E}(F_n^*(x)) = B_n(x) = \sum\limits_{k=0}^{n} F_n(u_k) b_k(n,x)$ является полиномом Бернштейна порядка n и $\sqrt{4\pi n x(1-x)} \sum_{k=0}^{n} b_k^2(n,x) = 1 + o(1)$ [4], то эти оценки состоятельны и в [6] также показано, что они асимптотически нормальны.

Теперь для заданного $0 < \lambda < 1$ определим

$$x_\lambda = \inf x : F(x) \geqslant \lambda, \quad \hat{x}_{n,\lambda} = \inf x : F_n^*(x) \geqslant \lambda.$$

В настоящем сообщении мы рассматриваем статистику $\hat{x}_{n,\lambda}$ в качестве оценки квантиля x_λ порядка $0 < \lambda < 1$. Имеют место следующие результаты.

Теорема 1. *Пусть $F(x)$ имеет третью ограниченную производную и $0 < \lambda < 1$ задано. Тогда*

$$\hat{x}_{n,\lambda} \xrightarrow[n\to\infty]{p} x_\lambda.$$

Теорема 2. *При условиях предыдущей теоремы*

$$\sqrt{n}(x_{n,\lambda} - x_\lambda) \xrightarrow[n\to\infty]{d} N(0, \sigma^2), \quad \text{где} \quad \sigma^2 = \frac{\lambda(1-\lambda)}{4\pi f^2(x_\lambda) x_\lambda (1 - x_\lambda)}.$$

ЛИТЕРАТУРА

1. M.S. Tikhov. *Statistical Estimation Based on Interval Censored Data* // Parametric and Semiparametric Models with Applications to Reliability, Survival Analysis, and Quality of Life (ed. N.Balakrishnan etc.). Springer, NY. 555 p., 211-218 (2004).

2. С.В. Криштопенко, М.С. Тихов, Е.Б. Попова. *Доза-эффект*, М.: Медицина. 2008. 288 с.

3. М.С. Тихов, Д.С. Криштопенко. *Оценивание распределений в зависимости доза-эффект при фиксированном плане эксперимента* // Статистические методы оценивания и проверки гипотез: межвуз. сб. науч. тр., Перм. ун-т, Пермь, 66-77 (2006).

4. G.J. Basu, A.J. Canty, Y.P. Chaubey. *Application of Bernstein polynomials for smooth estimation of a distribution and density function* // Journal of Statistical Planning and Inference, **105**, 377-392 (2002).

5. A. Leblanc. *On estimating distribution functions using Bernstein polynomials* // Ann. Inst. Statist. Math., **64**:5, 919-943 (2012)

6. П. Бабилуа, Е. Надарая. *Об одном новом методе оценки бернуллиевской функции регрессии* // Теория вероятн. и её примен., **67**:2, 209-222 (2022).