

Д. С. Криштопенко (Нижний Новгород, ННГУ). **Асимптотические распределения интегрированных квадратичных ошибок оценок Надарая–Ватсона в зависимости доза–эффект.**

Пусть $X_1, X_2, \dots, X_n$ — независимые и одинаково распределенные случайные величины с неизвестной непрерывной функцией распределения $F(x)$. Мы вводим случайные дозы U_i , которые наблюдаются с ошибкой, накладываемой аддитивно, и наблюдаем эффект от введенных доз $Y_i = u_i + \varepsilon_i$, т. е. мы наблюдаем выборку $\{(W_i, Y_i), i = 1, \dots, n\}$, $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ — независимые и одинаково распределенные случайные величины с плотностью распределения $q(x) > 0$, $x \in \mathbf{R}^1$, не зависящие от $\{X_i, 1 \leq i \leq n\}$; $W_i = I\{X_i < U_i\}$ — индикатор события $\{X_i < U_i\}$. Рассматриваются задачи оценивания функции распределения $F(x)$ по выборкам $\{(W_i, U_i), 1 \leq i \leq n\}$ (прямые наблюдения) и $\{(W_i, Y_i), i = 1, \dots, n\}$ (непрямые наблюдения), а также задачи проверки гипотез согласия и однородности. Данная задача решается для случая, когда вводимые дозы u_i выбираются заранее (фиксированный план), а также для случайных планов эксперимента.

В токсикометрии (см. [1]) рассматриваемая модель интерпретируется как зависимость «доза–эффект», где X — минимальная доза, с которой начинается реакция организма, U_i (u_i) — вводимые в организм случайные (неслучайные) дозы, а Y_i — измеренные с неизвестной ошибкой дозы.

Для статистической оценки $F(x)$ (см. [2] и [3]) при прямых и непрямых наблюдениях используют следующие статистики: а) для фиксированного плана (без ошибок — $F_n(x)$, с ошибками — $F_n^*(x)$):

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n W_i K_h(x - u_i), \quad F_n^*(x) = \frac{1}{n} \sum_{i=1}^n W_i K_h(x - Y_i);$$

б) для случайного плана: $F_n(x) = S_{2n}(x)/S_{1n}(x)$, $F_n^*(x) = S_{2n}^*(x)/S_{1n}^*(x)$, где

$$S_{1n}(x) = \frac{1}{n} \sum_{i=1}^n K_h(U_i - x), \quad S_{2n}(x) = \frac{1}{n} \sum_{i=1}^n W_i K_h(U_i - x),$$

$$S_{1n}^*(x) = \frac{1}{n} \sum_{i=1}^n K_h(Y_i - x), \quad S_{2n}^*(x) = \frac{1}{n} \sum_{i=1}^n W_i K_h(Y_i - x),$$

$K(\cdot) \geq 0$ — ядерная функция, $h = h(n) > 0$ — неслучайная последовательность, $K_h(x) = (1/h)K(x/h)$.

Положим $\|K\|^2 = \int K^2(x) dx$, $\nu^2 = \int x^2 K(x) dx$, $R(x) = \int F(u)g(u|x) du$, $m(y) = \int F(u)g(y, u) du$, $q(y) = \int g(y, u) du > 0$, $g(u) = \int g(y, u) dy > 0$, $\Phi(x) = \int F(t)q(x - t) dt$. Заметим, что при случайном плане $F(x) = m(x)g(x)$ в случае прямых наблюдений и $R(x) = m(x)q(x)$ в случае непрямых наблюдений.

Предположения (К).

(К0) $\max_i |u_n - u_{i-1}| = o(n^{-1})$ при $n \rightarrow \infty$.

(К1) $K(x) \geq 0$ является четной ограниченной непрерывной функцией на $\mathbf{R}$ и $\|K\|^2 < \infty$.

(К2) $\int K(x) dx = 1$, $\nu^2 = \int x^2 K(x) dx < \infty$.

(К3) $K(x) = 0$ для $x \notin [-1, 1]$.

(К4) Последовательность $h = h(n)$ такова, что $h \rightarrow 0$, $nh \rightarrow \infty$ при $n \rightarrow \infty$.

(К5) $f'(x)$ — непрерывная функция, причем $\int (f'(x))^2 dx < \infty$, а $f''(x)$ ограничена на $\mathbf{R}$.

(К6) $f(x)/F(x)$, $f'(x)f(x)/F(x)$ — ограниченные интегрируемые функции и $\int (f'(x))^4 dx < \infty$.

(К7) $g(y, u)$ имеет вторые непрерывные ограниченные производные на $\mathbf{R}^2$.

(К8) $g(y, u)$ имеет третьи ограниченные производные на $\mathbf{R}^2$.

Предположения (В).

(В1) $w(x) \geq 0$ — ограниченная функция на $\mathbf{R}$.

(В2) $\int w^2(x) dx < \infty$.

Асимптотическая нормальность рассматриваемых оценок доказывается в [2] и [3]. В качестве меры отклонения данных статистик от своих оцениваемых значений мы берем интегрированную квадратичную ошибку (ИКО):

$$J_n(x) = \int (F_n(x) - F(x))^2 dx, \quad J_n^* = \int (F_n^*(x) - \Phi(x))^2 dx.$$

Мы рассматриваем задачу нахождения предельных распределений ИКО. Пусть $c(n) = \int \mathbf{E} (F_n(x) - F(x))^2 w(x) dx$ и $\mu_m(n) = \int \mathbf{E} (F_n(x) - R(x))^2 w(x) dx$. Следующие теоремы устанавливают асимптотическую нормальность интегрированных квадратичных ошибок при фиксированном плане эксперимента.

Теорема 1. При указанных условиях на (К) имеем:

- 1) если $nh^5 \rightarrow \infty$, то $n^{1/2}h^{-2}(J_n - c(n)) \xrightarrow{d} N(0, \sigma_1^2)$, $n \rightarrow \infty$;
- 2) если $nh^5 \rightarrow 0$, то $nh^{-1/2}(J_n - c(n)) \xrightarrow{d} N(0, 2\sigma_3^2)$, $n \rightarrow \infty$;
- 3) если $nh^5 \rightarrow \lambda$, то $n^{9/10}(J_n - c(n)) \xrightarrow{d} N(0, \lambda^{4/5}\sigma_1^2 + 2\lambda^{-1/5}\sigma_3^2)$, $n \rightarrow \infty$, где

$$\sigma_1^2 = \frac{\nu^4}{4} \int F(t)(1 - F(t))(f'(t))^2 dt,$$

$$\sigma_3^2 = \frac{1}{2} \int F^2(t)(1 - F(t))^2 dt \int du \left(\int K(\nu)K(\nu + u) d\nu \right)^2.$$

Теорема 2. При указанных условиях на (К) имеем:

- 1) если $nh^5 \rightarrow \infty$, то $n^{1/2}h^{-2}(J_n^* - \mu(n)) \xrightarrow{d} N(0, (\sigma_1^*)^2)$, $n \rightarrow \infty$;
- 2) если $nh^5 \rightarrow 0$, то $nh^{-1/2}(J_n^* - \mu(n)) \xrightarrow{d} N(0, 2(\sigma_3^*)^2)$, $n \rightarrow \infty$;
- 3) если $nh^5 \rightarrow \lambda$, то $n^{9/10}(J_n^* - \mu(n)) \xrightarrow{d} N(0, \lambda^{4/5}(\sigma_1^*)^2 + 2\lambda^{-1/5}(\sigma_3^*)^2)$, $n \rightarrow \infty$, где

$$(\sigma_1^*)^2 = \frac{\nu^4}{4} \left\{ \int \Phi(u)(\Phi^*(u))^2 du - \int F^2(t) \left\{ \int \Phi''(t+y)q(y) dy \right\}^2 dt \right\},$$

$$(\sigma_3^*)^2 = \frac{1}{2} \int dx \int F^2(t)q^2(x-t) \int du \left(\int K(\nu)K(\nu+u) d\nu \right)^2.$$

Аналогичные результаты имеют место и для случайного плана. Рассмотрим

$$J_n = \int (F_n(x) - F(x))^2 w(x) dx, \quad J_n^* = \int (F_n^*(x) - R(x))^2 w(x) dx.$$

Теорема 3. Пусть выполнены условия (К) и (В). Тогда:

- 1) если $nh^5 \rightarrow \infty$, то $n^{1/2}h^{-2}(J_n^* - \mathbf{E} J_m) \xrightarrow{d} N(0, \nu^4 \sigma_I^2)$, $n \rightarrow \infty$;
- 2) если $nh^5 \rightarrow 0$, то $nh^{-1/2}(J_n^* - \mathbf{E} J_m) \xrightarrow{d} N(0, 2\sigma_{II}^2)$, $n \rightarrow \infty$;
- 3) если $nh^5 \rightarrow \lambda$, то $n^{9/10}(J_n^* - \mathbf{E} J_m) \xrightarrow{d} N(0, \nu^4 \lambda^{4/5} \sigma_I^2 + 2\lambda^{-1/5} \sigma_{II}^2)$, $n \rightarrow \infty$, где

$$\sigma_I^2 = \int (g''(x))^2 \frac{F^4(x)w^2(x)}{g^3(x)} dx + \int (m''(x))^2 \frac{F(x)w^2(x)}{g^3(x)} dx$$

$$+ 2 \int (g''(x))^2 \frac{F^3(x)w^2(x)}{g^3(x)} dx - 3 \left\{ \int g''(x) \frac{F^2(x)w(x)}{g(x)} dx \right\}^2$$

$$+ 2 \int (m''(x))^2 \frac{F^2(x)w^2(x)}{g^3(x)} dx - 3 \left\{ \int m''(x) \frac{F(x)w(x)}{g(x)} dx \right\}^2,$$

$$\sigma_{II}^2 = 3 \int F^2(y)w(y) dy \int du \left(\int K(u + \nu)K(\nu) d\nu \right)^2.$$

Теорема 4. Пусть выполнены условия **(K)** и **(B)**. Тогда:

- 1) если $nh^5 \rightarrow \infty$, то $n^{1/2}h^{-2}(J_n^* - \mathbf{E} J_n^*) \xrightarrow{d} N(0, \nu^4 \sigma_{*I}^2)$, $n \rightarrow \infty$;
- 2) если $nh^5 \rightarrow 0$, то $nh^{1/2}(J_n^* - \mathbf{E} J_n^*) \xrightarrow{d} N(0, 2\sigma_{*II}^2)$, $n \rightarrow \infty$;
- 3) если $nh^5 \rightarrow \lambda$, то $n^{9/10}(J_n^* - \mathbf{E} J_n^*) \xrightarrow{d} N(0, \nu^4 \lambda^{4/5} \sigma_{*I}^2 + 2\lambda^{-1/5} \sigma_{*II}^2)$, $n \rightarrow \infty$, где

$$\begin{aligned} \sigma_{*I}^2 &= \int (q''(x))^2 \frac{R^4(x)w^2(x)}{q^3(x)} dx + \int (m''(x))^2 \frac{R(x)w^2(x)}{q^3(x)} dx \\ &\quad + 2 \int (q''(x))^2 \frac{R^3(x)w^2(x)}{q^3(x)} dx - 3 \left\{ \int q''(x) \frac{R^2(x)w(x)}{q(x)} dx \right\}^2 \\ &\quad + 2 \int (m''(x))^2 \frac{R^2(x)w^2(x)}{q^3(x)} dx - 3 \left\{ \int m''(x) \frac{R(x)w(x)}{q(x)} dx \right\}^2, \\ \sigma_{*II}^2 &= 3 \int R^2(y)w(y) dy \int du \left(\int K(u + \nu)K(nu) d\nu \right)^2. \end{aligned}$$

Полученные результаты об асимптотической нормальности интегрированных квадратичных ошибок позволяют построить критерии для проверки гипотез согласия для функции распределения и однородности двух выборок в зависимости «доза–эффект».

СПИСОК ЛИТЕРАТУРЫ

1. Криштопенко С. В., Тихов М. С. Токсикометрия эффективных доз. Нижний Новгород: Изд-во ННГУ, 1997, 156 с.
2. Tikhov M. S. Statistical estimation on the basis of interval-censored data. — J. Math. Sci., 2004, v. 119, № 3, p. 321–335.
3. Тихов М. С., Криштопенко Д. С., Ярощук М. В. Оценивание распределений в зависимости доза–эффект при фиксированном плане эксперимента. — В сб.: Статистические методы оценивания и проверки гипотез. Пермь: Изд-во Пермского ун-та, 2006, с. 66–67.