М. С. Т и х о в, **Д. С. К р и ш т о п е н к о** (Нижний Новгород, ННГУ). **Асимптотические распределения суммируемых квадратичных уклонений оценок функции распределения в зависимости доза–эффект.**

Целью данного сообщения является изложение результатов асимптотического поведения L_2-уклонений ядерных оценок $F_N(x)$ функции распределения (ф. р.), определяемого как $I_N = \int (F_N(x) - F(x))^2 \omega(x)\,dx$, где $F(x)$ — неизвестная функция распределения случайной величины (с. в.) X, $\omega(x)$ — весовая функция, в зависимости «доза–эффект» (см. [1]) по выборке $\mathcal{U}^{(N)} = \{(W_i, U_i),\ 1 \leqslant i \leqslant N\}$, $W_i = I\{X_i < U_i\}$ есть индикатор события $\{X_i < U_i\}$, $\{U_i,\ 1 \leqslant i \leqslant N\}$ — набор одинаково $G(x)$-распределенных случайных величин, независимых от $\{X_i,\ 1 \leqslant i \leqslant N\}$. Полученные результаты можно использовать для проверки гипотез согласия в зависимости «доза–эффект». Мы рассматривали kNN-оценки функции распределения. Для них доказана асимптотическая нормальность статистик I_N при $N \to \infty$. Вывод результатов использует асимптотическую нормальность сумм (см. [2])

$$S_N = \frac{1}{N} \sum_{i=1}^{N} A_N\left(\frac{i}{N}, U_N^{(i)}\right), \qquad \widetilde{S}_N = \frac{1}{N} \sum_{i=1}^{N} A_N\left(\frac{i}{N}, U_N^{(i)}, W_N^{[i]}\right)$$

и

$$T_N = \frac{1}{N^2} \sum_{1 \leqslant i < j \leqslant N} B_N\left(\frac{i}{N}, \frac{i}{N}, U_N^{(i)}, U_N^{(j)}\right),$$
$$\widetilde{T}_N = \frac{1}{N^2} \sum_{1 \leqslant i < j \leqslant N} B_N\left(\frac{i}{N}, \frac{i}{N}, U_N^{(i)}, U_N^{(j)}, W_N^{[i]}, W_N^{[j]}\right),$$

где $U_N^{(i)}$ — i-я порядковая, а $W_N^{[i]}$ — i-я индуцированная порядковая статистика для выборки $\mathcal{U}^{(N)}$.

При компьютерной проверке гипотез более удобным является рассмотрение сумм квадратов уклонений ядерных оценок ф. р. $F_n(x)$ от $F(x)$ в заданных точках $\{x_j\}$, определяемых как

$$J_{nm} = \frac{1}{m} \sum_{j=1}^{m} (F_n(x_j) - F(x_j))^2$$

по повторной выборке $\mathcal{Y}^{(n)} = \{(W_i, Y_i),\ 1 \leqslant i \leqslant n\}$, где $Y_i = u_i + \varepsilon_i$, u_i могут быть как случайными, так и неслучайными величинами, ε_i — независимые с. в., одинаково распределенные с плотностью $q(x) > 0$. Наблюдению доступна выборка $\mathcal{Y}^{(n)} = \{(W_i, Y_i),\ 1 \leqslant i \leqslant n\}$. В качестве непараметрической оценки по $\mathcal{Y}^{(n)}$ берется статистика $F_n(x) = S_{2n}(x)/S_{1n}(x)$, где $S_{1n}(x) = (1/n)\sum_{i=1}^{n} K_h(Y_i - x)$, $S_{2n}(x) = (1/n)\sum_{i=1}^{n} W_i K_h(Y_i - x)$, $K(.) \geqslant 0$ есть *ядерная функция*, $h > 0$ — ширина окна просмотра данных, и $K_h(x) = (1/h)K(x/h)$.

В работах [3], [4] переменная U трактовалась как вводимая в организм наблюдаемая доза, а X — как минимальная доза, с которой начинается реакция организма. Здесь мы рассматриваем проблему, в основном, для фиксированных планов эксперимента (см. [5], [6]), когда u есть истинная доза, а Y — наблюдаемая со случайной ошибкой доза.

Определим $\|K\|^2 = \int K^2(x)\,dx$, $\nu^2 = \int x^2 K(x)\,dx$.

Условия (**A**).

(**A**1) $K(x) \geqslant 0$ является ограниченной четной непрерывной функцией с носителем на $[-1, 1]$.

(**A**2) $\int K(x)\,dx = 1$, $\|K\|^2 < \infty$, $\nu^2 = \int x^2 K(x)\,dx < \infty$, $d^4 = \int x^4 K(x)\,dx < \infty$.

(**A**3) Последовательность $h = h(n)$ такова, что $h \to 0$, $nh^3 \to \infty$ при $n \to \infty$.

($\mathbf{A}4$) Производная $f'(x)$ есть непрывная и ограниченная функция на $\mathbf{R}$, а $\int (f'(x))^2\,dx < \infty$.

($\mathbf{A}5$) Функции $f(x)/F(x)$, $f'(x)f(x)/F(x)$ ограниченны и интегрируемы, а $\int (f'(x))^4\,dx < \infty$.

Пусть $\mu(n) = \mathbf{E}\,(J_{nm})$. Следующая теорема устанавливает асимптотическую нормальность случайных величин J_{nm} при $n \to \infty$ для фиксированных планов эксперимента при фиксированном m, когда дозы наблюдаются без ошибок.

Теорема 1. *Пусть выполнены условия* ($\mathbf{A}$). *Тогда*:

1) *если* $nh^5 \to \infty$, *то* $n^{1/2}h^{-3/2}(J_{nm} - \mu(n)) \xrightarrow{d} \zeta_1 \in N(0, \sigma_1^2)$, $n \to \infty$;

2) *если* $nh^5 \to 0$, *то* $nh\,(J_{nm} - \mu(n)) \xrightarrow{d} \zeta_2 \in N(0, \sigma_2^2)$, $n \to \infty$;

3) *если* $nh^5 \to \lambda \in (0, \infty)$, *то* $n^{4/5}\,(J_{nm} - \mu(n)) \xrightarrow{d} \zeta_\lambda \in N(0, \lambda^{4/5}\sigma_1^2 + 2\lambda^{-1/5}\sigma_2^2)$, $n \to \infty$, *где*

$$\sigma_1^2 = \frac{\nu^4\,\|K\|^2}{4m^2}\sum_{j=1}^m F(x_j)(1-F(x_j))(f'(x_j))^2, \quad \sigma_2^2 = \frac{\beta_2}{2m^2}\sum_{j=1}^m F^2(x_j)(1-F(x_j))^2,$$

$$\beta_2 = \iint K^2(u)K^2(u+z)\,du\,dz.$$

Если доза наблюдается с ошибкой и мы имеем выборку $\mathcal{Y}^{(n)}$, то определим следующие величины:

$$\Phi(x) = F(t)q(x-t)\,dt,$$

$$J_{nm}^* = \frac{1}{m}\sum_{j=1}^m (F_n^*(x_j) - \Phi(x_j))^2, \quad \text{где} \quad F_n^*(x) = \frac{1}{n}\sum_{i=1}^n W_i K_h(x - Y_i).$$

Положим

$$\rho_1^2 = \frac{\nu^4\|K\|^2}{4m^2}\sum_{j=1}^m \Phi(x_j)(\Phi''(x_j))^2, \quad \rho_2^2 = \frac{1}{2m^2}\sum_{i=1}^m\sum_{j=1}^m K_2^2(x_i - x_j)\int F^2(u)q^2(x_j - u)\,du,$$

$$K_2(x) = \int K(z)K(x+z)\,dz, \quad \mu_m^*(n) = \mathbf{E}\,(J_{nm}^*).$$

В следующей теореме установлена асимптотическая нормальность суммируемых квадратичных уклонений когда дозы наблюдаются с ошибкой.

Теорема 2. *При указанных условиях* ($\mathbf{A}$) *и при* $n \to \infty$ *имеем*:

1) *если* $nh^5 \to \infty$, *то* $n^{1/2}h^{-3/2}\,(J_{nm}^* - \mu_m^*(n)) \xrightarrow{d} \zeta_1 \in N(0, \rho_1^2)$;

2) *если* $nh^5 \to 0$, *то* $nh(J_{nm}^* - \mu_m^*(n)) \xrightarrow{d} \zeta_2 \in N(0, \rho_2^2)$;

3) *если* $nh^5 \to \lambda \in (0, \infty)$, *то* $n^{4/5}(J_{nm}^* - \mu_m^*(n)) \xrightarrow{d} \zeta_\lambda \in N(0, \lambda^{4/5}\rho_1^2 + 2\lambda^{-1/5}\rho_2^2)$.

Результаты теорем 1, 2 сохраняются и когда $m = m(n)$, $m \to \infty$ ($n \to \infty$), но $m = o\,(h)$, при этом нормирующий множитель меняется. Так, в пукте 1) теоремы 1 он равен $n^{1/2}h^{-2}m^{-1/2}$. Если же при $n \to \infty$, $h = o\,(m)$, то, например, в теореме 1 3) номировка будет с помощью $n^{9/10}$, а не с помощью $n^{4/5}$, и

$$\sigma_1^2 = \frac{\nu^4\|K\|^2}{4}\int F(x)(1-F(x))(f'(x))^2\,dx, \quad \sigma_2^2 = \|K_2\|^2\int F^2(x)(1-F(x))^2\,dx.$$

Аналогичные результаты имеют место и для случайных планов эксперимента как для прямых, так и для непрямых наблюдений.

Рассмотрен также дискретный аналог статистики Андерсона–Дарлинга вида

$$AD_{nm} = \frac{1}{m}\sum_{j=1}^m \frac{(F_n(_j) - F(x_j))^2}{F_n(x_j)(1-F_n(x_j))}$$

и найдено ее асимптотическое распределение при фиксированном m и $n \to \infty$.

СПИСОК ЛИТЕРАТУРЫ

1. *Криштопенко С. В.*, *Тихов М. С.* Токсикометрия эффективных доз. Нижний Новгород: Изд-во ННГУ, 1997, 156 с.

2. *Тихов М. С.* Линейные функции индуцированных порядковых статистик и непараметрическое оценивание распределений в зависимости доза–эффект. — Обозрение прикл. и промышл. матем., 1999, т. 6, № 1, с. 244.

3. *Tikhov M. S.* Statistical estimation on the basis of interval-censored Data. — J. Math. Sci., 2004, v. 119, № 3, p. 321–335.

4. *Tikhov M.S.* Statistical estimation based on interval censored data. — In: Parametric and Semiparametric Models with Applications to Reliability, Survival Analysis, and Quality of Life Series: Statistics for Industry and Technology./ Ed. by M. S. Nikulin, N. Balakrishnan, M. Mesbah, N. Limnios. Heidelberg etc.: Springer, 2004, p. 209–215.

5. *Тихов М. С.*, *Криштопенко Д. С.*, *Ярощук М. В.* Оценивание распределений в зависимости доза–эффект при фиксированном плане эксперимента. — В сб.: Стат. методы оценивания и проверки гипотез. Пермь: Перм. ун-т, 2006, с. 66–77.

6. *Tikhov M. S.*, *Krishtopenko D. S.* Asymptotic normality of the integrated square error of distribution function estimators in dependence dose–response from indirect observations. — In: Abstract of the 12th International Conference on ASMDA, Greece, 2007.
7. *Тихов М. С.*, *Криштопенко Д. С.* Оценивание распределений в зависимости доза–эффект при фиксированном плане эксперимента в случае непрямых наблюдений. — Вестник Нижегородского университета им. Н. И. Лобачевского, 2007, № 2, с. 158–166.