М. С. Тихов, В. М. Кочеганов (Нижний Новгород, ННГУ). **Проверка гипотез согласия в зависимости доза–эффект по выборкам конечного объема.**

Рассматривается пороговая модель зависимости доза–эффект, когда в биообъект вводится случайная доза U и наблюдается альтернативный ответ W. Основу модели составляет предположение, что существует латентная величина X — пороговая доза, являющаяся самой большой дозой, при которой не наблюдается эффекта в эксперименте. Величина $W = I\{U > X\}$ есть индикатор события $\{U > X\}$. Биологическая суть вышеизложенного состоит в следующем. Предположим, что речь идет о яде, который попадает в биообъект (все яд и все лекарство, лишь дозы разделяют их). Например, при биопробе воды речь обычно идет о *Ceriodaphnia*, которые чувствительны к загрязнению воды веществами и которые для них являются ядом. Для каждого яда теоретически существует минимальная доза, которая вызывает у тест-объекта их гибель. Оценить эту дозу для каждого биообъекта довольно трудоемко. Если биообъект в эксперименте выжил, то он получил дозу, заведомо меньшую минимальной летальной дозы. Для каждого биообъекта эта доза будет различной, что определяется индивидуальной чувствительностью особей биологического вида к тестируемому препарату. Однако в однородной массе величина X будет случайной величиной с неизвестной функцией распределения $F(x) = \mathbf{P}\{X < x\}$. Такая модель рассмотрена в [1].

Пусть $\{(X_i, U_i),\, 1 \leqslant i \leqslant n\}$ — стационарная последовательность независимых пар случайных величин (X, U), где U имеет неизвестную функцию распределения $G(u)$ и плотность $g(u) > 0$ на $\mathbf{R}^1$. Мы наблюдаем выборку $\mathcal{U}^{(n)} = \{(U_i, W_i),\, 1 \leqslant i \leqslant n\}$.

В работах [2, 3–6] соответственно рассматривались задачи оценивания и проверки гипотез согласия H_0: $F(x) = F_0(x)$ против альтернативы H_1: $F(x) \neq F_0(x)$ по выборке $\mathcal{U}^{(n)}$. Для оценки $F(x)$ использовалось отношение

$$\widetilde{F}_n(x) = \sum_{i=1}^{n} W_i K\biggl(\frac{x - U_i}{h}\biggr) \bigg/ \sum_{i=1}^{n} K\biggl(\frac{x - U_i}{h}\biggr),$$

где $K(x) \geqslant 0$ — ядерная функция (ядро). В качестве ширины окна просмотра h рассматривались неслучайные последовательности $h = h(n)$, сходящиеся к нулю при $n \to \infty$. Было показано, что если $h = n^{-1/5}$, то

$$n^{2/5}(\widetilde{F}_n(x) - F(x)) \underset{n\to\infty}{\overset{\mathcal{D}}{\to}} \mathbf{N}(a(x), \sigma_0^2(x)),$$

$$a(x) = \frac{f'(x)g(x) + 2g'(x)f(x)}{g(x)}, \quad \sigma_0^2(x) = \frac{F(x)(1 - F(x))\|K\|^2}{g(x)}, \quad \|K\|^2 = \int K^2(x)\,dx.$$

При конкретной реализации критериев и оценок по выборке $\mathcal{U}^{(n)}$ конечного объема n оказалось, что мы получаем оценку, которая «ближе» не к функции $F(x)$, а к функции $F * K_h(x)$, где $F * K_h$ есть свертка функции $F(x)$ с функцией $K_h(x)$, а $K_h(x) = h^{-1}K(xh^{-1})$. Свертка $F * K_h(x)$ отличается от функции $F(x)$, особенно это отличие заметно на хвостах распределения $F(x)$. Так, например, если $F(x) = 1 - e^{-x}$, $x > 0$ (показательное распределение) и $K_h(x) = 1/(2h)$, $-h < x < h$, то $F * K_h(x) = (x + h + 1 - e^{-(x+h)}/(2h)$ для $-h < x < h$ и $F * K_h(x) = 1 - e^{-x}\operatorname{sh} h/h$. Таким образом, если исходное распределение $F(x)$ показательное с носителем на интервале $(0, \infty)$, то распределение $F * K_h(x)$ имеет носитель на интервале $(-h, \infty)$ и не является показательным. При этом $n^{2/5}\,(\widetilde{F}_n(x) - F * K_h(x)) \to^{\mathcal{D}}_{n\to\infty} \mathbf{N}(0, \sigma_0^2(x))$.

Отсюда мы заключаем, что при конечных выборках нам необходимо проверять гипотезу, касающуюся не функции $F_0(x)$, а гипотезу относительно функции $F_0 * K_h(x)$. Для этого мы можем использовать либо интегрированные квадратичные уклонения $I_n = \int (\widehat{F}_n(x) - F_0 * K_h(x))^2\,\omega(x)\,dx$, либо суммируемые квадратичные уклонения $S_{nm} = \sum_{j=1}^{m} (\widehat{F}_n(x_j) - F_0 * K_h(x_j))^2$.

Показано, что предельные распределения этих статистик являются нормальными распределениями. Что касается оценивания функции распределения $F(x)$, то мы предлагаем следующую процедуру. Пусть $\varphi_K(t)$ есть характеристическая функция плотности $K(x)$, а $\widehat{\varphi}(t)$ — состоятельная оценка характеристической функции, построенная по $\widehat{F}_n(x)$. В качестве оценки функции распределения при конечном n предлагается функция

$$\widehat{F}_n(\beta) - \widehat{F}_n(\alpha) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-it\alpha} - e^{-it\beta}}{it} \frac{\widehat{\varphi}(t)}{\varphi_K(t/h)}\, dt.$$

Так, для ядра Епанечникова $K(x) = (3/4)(1-x^2)$ характеристическая функция равна $\varphi_K(t) = (3\sin t - t\cos t)/t^3$, а для косинус-ядра $K(x) = (\pi/4)\cos(\pi x/2)$ — равна $\varphi_K(t) = \pi^2 \cos t/(\pi^2 - 4t^2)$. Оба ядра имеют носитель на отрезке $[-1, 1]$.

При условиях регулярности эти оценки $\widehat{F}_n(x)$ асимптотически нормальны. Именно,

$$\frac{n^{2/5}(\widehat{F}_n(x) - \mathbf{E}\,(\widehat{F}_n(x)))}{\sigma_n(x)} \underset{n\to\infty}{\overset{\mathcal{D}}{\to}} \mathbf{N}(0,1), \quad \sigma_n^2 = \mathbf{D}\,(\widehat{F}_n(x)).$$

СПИСОК ЛИТЕРАТУРЫ

1. *Криштопенко С. В.*, *Тихов М. С.*, *Попова Е. Б.* Доза–эффект. М.: Медицина, 2008, 288 с.
2. *Tikhov M. S.* Statistical Estimation based on Interval Censored Data. – In: Param. and Semiparam. Models with Appl. to Rel., Surv. Analisys, and Qual. of Life: Springer-Verlag: Theor.& Meth., 2004, p. 209–215.
3. *Тихов М. С.*, *Криштопенко Д. С.* Асимптотичесикие распределения интегрированных квадратичных ошибок функции распределения в зависимости доза–эффект. — Обозрение прикл. и промышл. матем., 2005, т. 12, в. 3, с. 665–667.
4. *Криштопенко Д. С.*, *Тихов М. С.* Оценивание распределений в зависимости доза–эффект при фиксированном плане эксперимента в случае непрямых наблюдений. — Вестник Нижегородского ун-та, 2007, № 2, с. 158–164.
5. *Тихов М. С.*, *Криштопенко Д. С.* Асимптотичесикие распределения суммируемых квадратичных уклонений оценок функции распределения в зависимости доза–эффект. — Обозрение прикл. и промышл. матем., 2009, т. 16, в. 5, с. 772–786.
6. *Тихов М. С.*, *Криштопенко Д. С.* Распределение интегрированных квадратичных ошибок несмещенных ядерных оценок функции распределения в зависимости доза–эффект. — Нелинейный мир, 2007, т. 5, № 1-2, с. 20–30.