УДК 519.2

# ОБ УМЕНЬШЕНИИ ПОГРЕШНОСТИ НАБЛЮДЕНИЙ ПРИ ОЦЕНИВАНИИ ЗАВИСИМОСТИ ДОЗА – ЭФФЕКТ



***М.В. Ярощук***

Нижегородский госуниверситет им. Н.И. Лобачевского

marina.yaroschuk@yandex.ru

*Поступила в редакцию 26.03.2012*

Рассматривается математическая модель зависимости доза–эффект в случае непрямых наблюдений. Предлагаются и исследуются способы уменьшения погрешности наблюдений для оценивания неизвестной функции распределения.

*Ключевые слова:* зависимость доза–эффект, непрямые наблюдения, погрешности наблюдений, регрессия.

## Введение

В работе рассматривается математическая модель зависимости доза–эффект (см. [1]) для схемы непрямых наблюдений, т.е. когда вводимая в организм доза измеряется с некоторой ошибкой, а реакция организма (эффект) идет на «чистую» вводимую дозу. Рассмотрен случайный план эксперимента, когда вводимая доза является случайной величиной. Данная модель при наличии погрешности измерения в основном отражает применяемую методику проведения клинических испытаний лекарственных препаратов. В схеме непрямых наблюдений, если распределение ошибки неизвестно, устранить ее влияние в определенных случаях можно либо при помощи специальных ядер (см. [2]), либо используя априорные сведения о распределении ошибок. Однако, как показали результаты численного моделирования, указанные методы устранения ошибок работают для оценки плотности, а для оценки функции распределения дают неудовлетворительные результаты. Мы предлагаем и исследуем способы уменьшения погрешности для оценки функции распределения в некоторых случаях, имеющих важное значение для приложений. Так, если распределение с.в. $X$ является нормальным $N(a, \sigma^2)$ с неизвестными параметрами, распределение ошибки $\varepsilon$ также подчиняется нормальному закону $N(0,\sigma_0^2)$ с известной дисперсией $\sigma_0^2$, где дисперсия $\sigma_0^2$ много меньше дисперсии $\sigma^2$, мы строим оценку, предельное распределение которой мало зависит от погрешности измерения.

В задаче доза–эффект для случайных планов эксперимента математическая модель в схеме непрямых наблюдений имеет следующий вид. Измерения вводимой дозы $U$ осуществляются с погрешностью $\varepsilon$, имеющей плотность $g(x)$, то есть вместо с.в. $U$ наблюдается с.в. $Y$. Эта ошибка может накладываться аддитивно, тогда $Y = U + \varepsilon$, при фиксированном значении $U = u$ распределение величины $Y$ имеет плотность $q(y - u)$. В общем случае распределение ошибки описывается условной плотностью $q(y|u)$.

Имеем: $X_1, X_2,\ldots,X_n$ – независимые одинаково распределенные случайные величины (н.о.р.с.в.) с функцией распределения (ф.р.) $F(x)$, $U_1, U_2,\ldots,U_n$ – независимые между собой и одинаково распределенные с.в., независимые от $\{X_i, 1 \le i \le n\}$, с неизвестной ф.р. $G(x)$, $Y_1, Y_2,\ldots,Y_n$ – н.о.р.с.в. с неизвестной ф.р. $Q(y)$. Мы наблюдаем повторную выборку $\mathbf{Y}^{(n)} = \{(Y_i, W_i), 1 \le i \le n\}$, где $W_i = I(U_i > X_i)$ есть индикатор события $\{U_i > X_i\}$, т.е. $Y_i$ – наблюдаемое значение, а реакция организма осуществляется на величину $U_i$.

Определим статистики

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

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

где $h = cn^{-1/5}$, $c$ – некоторая заданная положительная константа, функция $K(x)$ – ядерная функция (ядро), $K_h(x) = \frac{1}{h}K\left(\frac{x}{h}\right)$.

По выборке $\{(Y_i, W_i)\}_{i=1}^n$ построим оценку для $F(x)$ как отношение статистики $S_{2n}^*(x)$ к $S_{1n}^*(x)$

$$F_n^*(x) = \frac{S_{2n}^*(x)}{S_{1n}^*(x)}, \qquad (1)$$

полагая $F_n^*(x) = 0$, если $S_{1n}^*(x) = 0$.

Пусть $m(x) = \int F(y)g(x-y)\,dy$.

Известно (см. [3, 4]), что при некоторых условиях регулярности оценки (1) являются состоятельными и асимптотически нормальными оценками неизвестной функции распределения $F(x)$ и кроме того

$$\sqrt{nh}\left(S_{1n}^{*}(x)-1\right)\xrightarrow[n\to\infty]{d}N\left(0,\|K\|^{2}\right), \qquad (2)$$

$$\sqrt{nh}\left(S_{2n}^{*}(x)-m(x)\right)\xrightarrow[n\to\infty]{d}N\left(\frac{\nu^{2}}{2}m''(x),\|K\|^{2}m(x)\right), \quad (3)$$

$$\sqrt{nh}\left(F_{n}^{*}(x)-m(x)\right)\xrightarrow[n\to\infty]{d}N\left(\frac{\nu^{2}}{2}m''(x),\ \|K\|^{2}\sigma^{2}(x)\right), \quad (4)$$

где $\sigma^{2}(x)=m(x)\left(1-m(x)\right)$, $\nu^{2}=\int\limits_{-\infty}^{\infty}x^{2}K(x)dx$,

$\|K\|^{2}=\int\limits_{-\infty}^{\infty}K^{2}(x)\,dx$.

## Результаты и их обсуждение

Мы рассмотрим задачу устранения погрешности измерений, когда:

1) погрешность $\varepsilon$ имеет известное нормальное распределение $N\left(0,\sigma_{0}^{2}\right)$, распределение случайной величины $X$ – нормальное $N(a, \sigma^2)$, параметры $(a, \sigma^2)$ неизвестны;

2) погрешность $\varepsilon$ имеет известное нормальное распределение $N\left(0,\sigma_{0}^{2}\right)$, функция распределения $F(x)$ случайной величины $X$ неизвестна;

3) погрешность $\varepsilon$ имеет известное нормальное распределение $N\left(0,\sigma_{0}^{2}\right)$, распределение случайной величины $U$ – нормальное $N(a, \sigma^2)$, параметры $(a, \sigma^2)$ неизвестны, распределение случайной величины $X$ неизвестно.

**1.** Если ошибка $\varepsilon$ имеет нормальное распределение $N\left(0,\sigma_{0}^{2}\right)$, то из соотношений (2) и (3) заключаем, что нормированные разности $\sqrt{nh}\left(S_{1n}^{*}(x)-1\right)$ и $\sqrt{nh}\left(S_{2n}^{*}(x)-F(x)\right)$ имеют асимптотически нормальные распределения $N\left(0,\|K\|^{2}\right)$ и $N\left(\frac{\nu^{2}}{2}f'(x),F(x)\|K\|^{2}\right)$ соответственно.

Пусть распределение случайной величины $U$ имеет плотность

$$g(x)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^{2}}{2\sigma_{0}^{2}}\right).$$

Пусть с.в. $X\in N\left(a,\sigma^{2}\right)$, т.е. $F(t) = P(X < t) = =\Phi\left(\frac{x-a}{\sigma}\right)$, $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x}e^{-\frac{t^{2}}{2}}dt$.

Тогда $m(x)=\int F(t)g(x-t)dt-$ свертка двух нормальных распределений. Применяя теорему Фубини [5], заключаем, что

$$m'(x)=\int\limits_{-\infty}^{\infty}f(t)g(x-t)dt=$$

$$=\int\limits_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(t-a)^{2}}{2\sigma^{2}}}\cdot\frac{1}{\sqrt{2\pi}\sigma_{0}}e^{-\frac{(x-t)^{2}}{2\sigma_{0}^{2}}}dt=$$

$$=\frac{1}{\sqrt{2\pi\left(\sigma_{0}^{2}+\sigma^{2}\right)}}e^{-\frac{(x-a)^{2}}{2\left(\sigma_{0}^{2}+\sigma^{2}\right)}}.$$

Следовательно,

$$m(x)=\int\limits_{-\infty}^{x}\frac{1}{\sqrt{2\pi\left(\sigma_{0}^{2}+\sigma^{2}\right)}}e^{-\frac{(t-a)^{2}}{2\left(\sigma_{0}^{2}+\sigma^{2}\right)}}dt=\Phi\left(\frac{x-a}{\sqrt{\sigma_{0}^{2}+\sigma^{2}}}\right)$$

и

$$\sqrt{nh}\left(S_{2n}^{*}(x)-F(x)\right)\xrightarrow[n\to\infty]{d}N\left(\frac{\nu^{2}}{2}m''(x),m(x)\|K\|^{2}\right).$$

Теперь из (4) заключаем, что

$$\sqrt{nh}\left(\widehat{m}(x)-F(x)\right)\xrightarrow[n\to\infty]{d}$$

$$\xrightarrow[n\to\infty]{d}N\left(\frac{\nu^{2}}{2}m''(x),\ m(x)\left(1-m(x)\right)\|K\|^{2}\right),$$

где $\widehat{m}(x)=F_{n}^{*}(x)=\frac{S_{2n}^{*}(x)}{S_{1n}^{*}(x)}$.

Таким образом, установлено, что при нормальном распределении с.в. $X$ $N(a, \sigma^2)$ и известной дисперсии $\sigma_{0}^{2}$ с.в. $\varepsilon\in N\left(0,\sigma_{0}^{2}\right)$ предельное распределение оценок получилось аналогичным тому, что и в схеме прямых наблюдений. Отличие состоит в том, что дисперсии распределений $m(x)$ и $F(x)$ различны.

Чтобы найти оценку $\widehat{m}(x)$, вычислим первый и второй начальные моменты $\widehat{\mu}_{1}$ и $\widehat{\mu}_{2}$:

$$\begin{cases}\int\limits_{-\infty}^{\infty}x\,d\widehat{m}(x)\equiv\int\limits_{0}^{\infty}(1-\widehat{m}(x))\,dx-\int\limits_{-\infty}^{0}\widehat{m}(x)\,dx=\widehat{\mu}_{1},\\ \int\limits_{-\infty}^{\infty}x^{2}d\widehat{m}(x)=\widehat{\mu}_{2}.\end{cases}$$

Положим $\widehat{\rho}^{2}=\widehat{\mu}_{2}-\widehat{\mu}_{1}^{2}$. Тогда неизвестные параметры $a$ и $\sigma^2$ найдем из системы уравнений

$$\begin{cases}a+\frac{\nu^{2}}{2}+\widehat{m}''(x)=\widehat{\mu}_{1},\\ \widehat{\rho}^{2}=\sigma^{2}+\sigma_{0}^{2},\end{cases}\Rightarrow\begin{cases}\widehat{a}=\widehat{\mu}_{1}-\frac{\nu^{2}}{2}\widehat{m}''(x),\\ \widehat{\sigma}^{2}=\widehat{\rho}^{2}-\sigma_{0}^{2}.\end{cases}$$

Функцию $\psi_n(x)$ определим соотношением

$$\psi_{n}(x)=\widehat{m}''(x)=\frac{\widehat{m}(x+h)-2\widehat{m}(x)+\widehat{m}(x-h)}{h^{2}}.$$

Заметим, что если с.в. $X$ имеет функцию распределения $F(x)$ и функцию времени жизни $S(x)=1-F(x)$, то $\int_0^\infty x\,dF(x)=\int_0^\infty S(x)\,dx$.

Далее,

$$\int_0^\infty\left(\int_x^\infty S(u)\,du\right)^2\frac{dF(x)}{S^2(x)}=$$

$$=-\int_0^\infty\left(\int_x^\infty S(u)\,du\right)^2\frac{dS(x)}{S^2(x)}=\int_0^\infty\left(\int_x^\infty S(u)\,du\right)^2 d\left(\frac{1}{S(x)}\right)=$$

$$=\left(\int_x^\infty S(u)\,du\right)^2\cdot\frac{1}{S(x)}\Bigg|_0^\infty+\int_0^\infty\frac{2}{S(x)}\int_x^\infty S(u)\,du\cdot S(x)\,dx=$$

$$=-\mu^2-2x\int_x^\infty S(u)\,du\Bigg|_0^\infty+2\int_0^\infty xS(x)\,dx=$$

$$=-\mu^2+\int_0^\infty S(x)\,d(x^2)=-\mu^2+\int_0^\infty x^2\,dF(x).$$

Поэтому эмпирические моменты можно считать используя приведенные соотношения и подставляя вместо теоретической функции распределения эмпирическую.

Найденные значения $\hat{a}$ и $\hat{\sigma}^2$ подставим в функцию распределения с.в. $X$ и получим оценку этой функции распределения.

**2.** Пусть $Y=U+\varepsilon$ – сумма независимых случайных величин $U$ и $\varepsilon$, причем распределение $\varepsilon$ – нормально с нулевым математическим ожиданием и известной дисперсией $\sigma_0^2$, а случайная величина $U$ имеет неизвестную плотность $g(u)>0$. Для устранения погрешности $\varepsilon$ мы воспользуемся следующей идеей работы Э. Надарая [6]. Кривая регрессии $U$ по $Y$ имеет вид:

$$u(x)=E(U\mid Y=x)=\frac{\rho(x)}{q(x)},$$

где $\rho(x)=\int_{-\infty}^{\infty}u\,g(u)\frac{1}{\sqrt{2\pi}\sigma_0}e^{-\frac{(x-u)^2}{2\sigma_0^2}}du$, $q(x)=\int_{-\infty}^{\infty}g(u)\times$

$$\times\frac{1}{\sqrt{2\pi}\sigma_0}e^{-\frac{(x-u)^2}{2\sigma_0^2}}du.$$

Найдем условное распределение $Y$ при условии $U=u$. Имеем

$$P(Y<y|U=u)=P(U+\varepsilon<y|U=u)=$$

$$=P(u+\varepsilon<y|U=u)=P(\varepsilon<y-u),$$

поэтому условная плотность $Y$ на $U$ равна

$$q_{Y|U}(y\mid u)=q(y\mid u)=\frac{1}{\sqrt{2\pi}\sigma_0}e^{-\frac{(y-u)^2}{2\sigma_0^2}}.$$

Совместная плотность пары $(Y,\ U)$ будет равна

$$g(y,u)=q(y\mid u)\cdot g(u)=\frac{1}{\sqrt{2\pi}\sigma_0}e^{-\frac{(y-u)^2}{2\sigma_0^2}}\cdot g(u),$$

а маргинальная плотность распределения величины $Y$ будет равна

$$q(y)=\int_{-\infty}^{\infty}g(y,u)\,du=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_0}e^{-\frac{(y-u)^2}{2\sigma_0^2}}g(u)\,du.$$

Кривая регрессии $U$ по $Y$ имеет вид:

$$u(x)=E(U\mid Y=x)=\frac{\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_0}e^{-\frac{(y-u)^2}{2\sigma_0^2}}g(u)du}{q(x)}=\frac{\rho(x)}{q(x)}.$$

С помощью этой кривой регрессии мы будем «исправлять» выборочные данные и использовать их для оценки распределения величины $X$ «без ошибки».

Представим последнее соотношение в другом виде. Для этого рассмотрим производную $q'(x)$. Она равна

$$q'(x)=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_0}\left(-\frac{x-u}{\sigma_0^2}\right)e^{-\frac{(x-u)^2}{2\sigma_0^2}}g(u)\,du=$$

$$=-x\,q(x)+\rho(x).$$

Отсюда

$$\frac{q'(x)}{q(x)}=-\frac{x}{\sigma_0^2}+\frac{\rho(x)}{\sigma_0^2q(x)},$$

значит,

$$\frac{\rho(x)}{q(x)}=\sigma_0^2\frac{q'(x)}{q(x)}+x.$$

Пример 1. Пусть случайная величина $Y$ имеет нормальное распределение с неизвестным математическим ожиданием $a$ и неизвестной дисперсией $\sigma^2$, ее плотность равна

$$q(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Тогда

$$\frac{q'(x)}{q(x)}=\left(\ln q(x)\right)'=-\frac{x-a}{\sigma^2}$$

$$\text{и }\ \sigma_0^2\frac{q'(x)}{q(x)}+x=\frac{\sigma^2-\sigma_0^2}{\sigma^2}x+a\frac{\sigma_0^2}{\sigma^2}.$$

Поскольку $a$ и $\sigma^2$ неизвестны, то мы оценим их по выборке $y_1, y_2,\ldots,y_n$ с помощью следующих статистик:

$$\hat{a}=\bar{y}=\frac{1}{n}\sum_{i=1}^n y_i\ \text{ и }\ \hat{\sigma}^2=s^2=\frac{1}{n}\sum_{i=1}^n(y_i-\bar{y})^2.$$

В качестве приближения по выборочным данным мы рассмотрим статистику

$$u_n(x)=\sigma_0^2\frac{\rho_n(x)}{q_n(x)}+x,$$

где $$q_n(x)=\frac{1}{n}\sum_{i=1}^n K_h(y_i-x)$$ и

$$\rho_n(x)=\frac{q_n(x+h)-q_n(x-h)}{2h}.$$

Имеет место следующий результат.

**Теорема 1.** *Пусть $I$ – произвольный конечный интервал $[a, b]$ на прямой и выполнены следующие условия: $K(x)$ – неотрицательная, нормированная, четная, финитная функция с ограниченной вариацией $\mu = V(K)$, $nh^2 \to \infty$, при $n \to \infty$, плотность $q(x)$ имеет до третьего порядка непрерывные и ограниченные производные.*

*Тогда*

$$\sup_{x\in I} |u_n(x) - u(x)| \underset{n\to\infty}{\overset{p}{\to}} 0.$$

**Доказательство** повторяет с небольшим отличием идею доказательства работы [6].

Пусть $\varepsilon > 0$ и $V_{2n} = \sup_x |\rho_n(x) - E(\rho_n(x))|$.

Имеем:

$$V_{2n} = \sup_x |\rho_n(x) - E(\rho_n(x))| =$$

$$\sup_x \left| \frac{q_n(x+h) - q_n(x-h)}{2h} - E\left( \frac{q_n(x+h) - q_n(x-h)}{2h} \right) \right| =$$

$$= \sup_x \left| \frac{1}{2h^2} \int K\left( \frac{u-x-h}{h} \right) dF_n(u) - \right.$$

$$- \frac{1}{2h^2} \int K\left( \frac{u-x+h}{h} \right) dF_n(u) -$$

$$\left. - \frac{1}{2h^2} \int K\left( \frac{u-x-h}{h} \right) dF(u) + \frac{1}{2h^2} \int K\left( \frac{u-x+h}{h} \right) dF(u) \right| =$$

$$= \sup_x \left| \frac{1}{2h} \int (F_n(u) - F(u)) dK\left( \frac{u-x+h}{h} \right) + \right.$$

$$\left. + \frac{1}{2h} \int (F_n(u) - F(u)) dK\left( \frac{u-x-h}{h} \right) \right| \le$$

$$\le \sup_x |F_n(x) - F(x)| \cdot \frac{\mu}{2h}.$$

Пусть $D_n(x) = \sup_x |F_n(x) - F(x)|$, где $F(x)$ – ф.р. с.в. $X$, а $F_n(x)$ – эмпирическая ф.р. с.в. $X$. Как показано в [7], для $n \ge N_0$ существует такая универсальная константа $C_1$, что

$$P\left( D_n \ge \frac{\lambda}{\sqrt{n}} \right) \le C_1 e^{-2\lambda^2}.$$

Принимая во внимание эти неравенства, заключаем, что для $\varepsilon > 0$,

$$P(\sup_x |\rho_n(x) - E(\rho_n(x))| > \varepsilon) \le$$

$$\le P(\sup_x |F_n(x) - F(x)| > \varepsilon \frac{h}{\mu}) \le C_1 e^{-4\beta n h^2},$$

где $\beta = 2\left( \frac{\varepsilon}{\mu} \right)^2$.

Ввиду того, что плотность $q(x)$ имеет непрерывные и ограниченные производные до третьего порядка включительно, причем $\sup_x |q'''(x)| \le L_3$,

$$\sup_x |E(\rho_n(x)) - q'(x)| =$$

$$= \sup_x \left| \frac{1}{2h^2} \int K\left( \frac{y-x-h}{h} \right) q(y)\, dy - \right.$$

$$\left. - \frac{1}{2h^2} \int K\left( \frac{y-x+h}{h} \right) q(y)\, dy - q'(x) \right| =$$

$$= \sup_x \left| \frac{1}{2h} \int (q(x + (z+1)h) - \right.$$

$$\left. - q(x + (z-1)h)) K(z)\, dz - q'(x) \right| =$$

$$= \sup_x \left| \frac{1}{2h} \int (2hq'(x) + \frac{h^2 z}{2} q''(x) + \frac{2(z^2+2)}{6} h^3 \times \right.$$

$$\left. \times (q'''(\xi_1) + q'''(\xi_2)) K(z)\, dz - q'(x) \right| \le \frac{(\nu^2+2) L_3}{3} h^2.$$

Поэтому при $n \ge N_1(\varepsilon)$ $P(\sup_x |\rho_n(x) - q'(x)| > \varepsilon) < C_1 e^{-\beta_1 n h^2}$, где $\beta_1 = \frac{1}{2}\left( \frac{\varepsilon}{\mu} \right)^2$.

Обозначим $(\ln q(x))' = V(x)$, и пусть $\eta_1 = \min_{x\in I} q(x) \ne 0$, $\eta_2 = \max_{x\in I} |V(x)|$.

Тогда

$$P\left( \sup_{x\in I} \left| \frac{\rho_n(x)}{q_n(x)} - V(x) \right| > \varepsilon \right) \le$$

$$\le P(\sup_{x\in I} |\rho_n(x) - V(x) q_n(x)| >$$

$$> \varepsilon(\beta_1 - \varepsilon)) + P(\sup_{x\in I} |q_n(x) - q(x)| > \varepsilon) \le$$

$$\le P(\sup_{x\in I} |\rho_n(x) - V(x) q_n(x)| >$$

$$> \varepsilon(\beta_1 - \varepsilon)) + P(\sup_{x\in I} |q_n(x) - q(x)| > \varepsilon) \le$$

$$\le P(\sup_{x\in I} |\rho_n(x) - q'(x)| + \sup_{x\in I} |V(x)(q_n(x) - q(x))| >$$

$$> \varepsilon(\mu_1 - \varepsilon)) + P(\sup_{x\in I} |q_n(x) - q(x)| > \varepsilon) \le$$

$$\le P(\sup_{x\in I} |\rho_n(x) - q'(x)| > \varepsilon(\eta_1 - \varepsilon)/2) +$$

$$+ P(\sup_{x\in I} |q_n(x) - q(x)| > \varepsilon) +$$

$$+ P(\sup_{x\in I} |(q_n(x) - q(x))| > \varepsilon(\eta_1 - \varepsilon)/(2\eta_2)) \le$$

$$\le C_4 \left( e^{-\beta_1 n h^2} + e^{-\beta_2 n h^2} + e^{-\beta_3 n h^2} \right),$$

где $\beta_2 = \frac{1}{8}\left( \frac{\varepsilon(\eta_1 - \varepsilon)}{\mu} \right)^2$, $\beta_3 = \beta_2 / \eta_2^2$.

По условию теоремы $nh^2 \to \infty$, поэтому $\sup_{x\in I} |u_n(x) - u(x)| \underset{n\to\infty}{\overset{p}{\to}} 0$.

Теорема доказана.

**3.** Рассмотрим случай, когда имеются наблюдения $Y = U + \varepsilon$; $U$, $\varepsilon$ независимы и имеют нормальные распределения соответственно $N(a, \sigma^2)$ и $N(0,\sigma_0^2)$, $\sigma_0^2$ известна. Здесь условная плотность распределения $q(u|y)$ имеет нормальное распределение $N(\mu_1(y), \rho_1^2)$, где $\mu_1(y) = \frac{a\sigma_0^2}{\sigma_0^2+\sigma^2} + \frac{y\sigma^2}{\sigma_0^2+\sigma^2}$, $\rho_1^2 = \frac{\sigma^2\sigma_0^2}{\sigma_0^2+\sigma^2}$, т.е. $q(u\,|\,y) = \frac{1}{\rho_1}\varphi\left(\frac{u-\mu_1(y)}{\rho_1}\right)$, где $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$.

Пусть $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$, $s^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2$.

Предположим, что случайная величина $Z$ имеет нормальное распределение $N(\mu,\rho^2)$ с функцией распределения $\Phi((x-\mu)/\rho)$, $\Phi(x) = \int_{-\infty}^{x}\varphi(t)\,dt$.

При $n \to \infty$ последовательность $\{\widehat{F}_n(v), n \geq 1\}$ сходится к среднему

$$R(v) = \int_{-\infty}^{v}\Phi\left(\frac{u-\mu}{\rho}\right)\frac{1}{\rho_1}\varphi\left(\frac{u-\mu_1(v)}{\rho_1}\right)du = \Phi\left(\frac{\mu_1(v)-\mu}{\sqrt{\rho^2+\rho_1^2}}\right).$$

Если $\frac{\sigma_0^2}{\sigma^2} << 1$, то заменим $\sqrt{1+\frac{\sigma_0^2}{\sigma^2}}$ на 1 и потребуем, чтобы

$$\Phi\left(\frac{\frac{a\sigma_0^2}{\sigma_0^2+\sigma^2} + \frac{v\sigma^2}{\sigma_0^2+\sigma^2} - a}{\sigma}\right) = \Phi\left(\frac{x-a}{\sigma}\right).$$

Отсюда $\frac{a\sigma_0^2}{\sigma_0^2+\sigma^2} + \frac{v\sigma^2}{\sigma_0^2+\sigma^2} = x$, значит, $v = x\cdot\frac{\sigma_0^2+\sigma^2}{\sigma^2} - a$.

Поскольку $a$ и $\sigma^2$ неизвестны, то возьмем их оценки: $\hat{a} = \bar{y}, \hat{\sigma}^2 = s^2$. Поэтому

$$v^* = x\frac{s^2}{s^2-\sigma_0^2} - \bar{y} = x + (x-\bar{y})\frac{\sigma_0^2}{s^2-\sigma_0^2}.$$

Значение $\widehat{F}_n(v^*)$ будем использовать в качестве оценки $\Phi\left(\frac{x-a}{\sigma}\right)$ в точке $x$.

Если условие $\frac{\sigma_0^2}{\sigma^2} << 1$ не выполнено, то $v^*$ найдем из уравнения

$$v^* = \left(\sqrt{1+\frac{\sigma_0^2}{s^2-\sigma_0^2}}\right)^3(x-\bar{y}) - \bar{y},$$

где $x$ – заданная точка функции распределения $F(x)$.

*Список литературы*


1. Тихов М.С. *PC*-оценки функции распределения в модели доза–эффект по случайным планам эксперимента // Вестник Нижегородского университета. 2012. Вып.1(1). С. 138–143.
2. Fan J., Truong Y. Nonparametric regression with errors in variables // The Annals of Statistics. 1993. V. 21, № 4. P. 1900–1925.
3. Тихов М.С., Ярощук М.В. Статистическое оценивание распределений по интервально цензурированным выборкам в схеме непрямых наблюдений // Нелинейный мир. 2007. Т. 5, №1. 2. С. 4–8.
4. Ярощук М.В. Имитационное моделирование зависимости доза–эффект и статистический анализ оценок функции эффективности // Обозрение прикл. и пром. математики. 2009. Т. 16. В. 6. С. 1148–1150.
5. Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа. М.: Наука, 1976. 542 с.
6. Надарая Э.А. О непараметрических оценках плотности вероятности и регрессии // Теория вероятн. и ее примен. 1965. Т. 10. В.1. С. 199–203.
7. Dvoretzky A., Kiefer J. and Wolfowitz J. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator // Annals of Mathematical Statistics. 1956. V. 27. P. 642–669.


## ON REDUCTION OF OBSERVATION ERRORS IN ESTIMATING THE DOSE-EFFECT DEPENDENCE

***M.V. Yaroshchuk***

A mathematical model of doze-effect dependence in case of indirect observations is considered. The ways to reduce observational errors for the estimation of an unknown distribution function are proposed and studied.

*Keywords:* dose-effect dependence, indirect observations, observation errors, regression.