Министерство образования и науки Российской Федерации
Нижегородский государственный университет им. Н.И. Лобачевского
Факультет вычислительной математики и кибернетики
Кафедра прикладной теории вероятностей

Выпускная квалификационная работа бакалавра
«Оценивание эффективных доз в зависимости «доза-эффект»

Заведующий кафедрой ПТВ
д.ф.-м.н., профессор
Федоткин М.А.

Научный руководитель
д.ф.-м.н., профессор
Тихов М.С.

Выполнил
студент группы 8403
Кочеганов В.М.

Нижний Новгород
2011

Содержание

Аннотация

В работе предложены оценки категорий эффективных доз в зависимости «доза-эффект». Доказана их состоятельность и асимптотическая нормальность. Для выборок конечного объема проведен компьютерный анализ рассматриваемых оценок.

Результаты работы опубликованы в журнале «Обозрение прикладной и промышленной математики» (см. [?],[?]).

Введение

Во многих областях медицины и биологии (фармакологии, токсикологии, радиобиологии, биохимии и др.) фундаментальной проблемой является изучение механизмов действия лекарственных средств, токсических веществ, ионизирующей радиации на биологические, экологические объекты. Наиболее востребовано решение данной проблемы в фармакологии при создании новых лекарственных средств (т.е. фармакологических средств, прошедших клинические испытания), поэтому при разработке новых лекарств анализ связи между дозой и эффектом и ее количественное определение имеет большое значение для практики. Несмотря на то, что спектр проявлений токсического процесса определяется строением токсиканта, тем не менее, выраженность развивающегося эффекта является функцией количества действующего агента. Для обозначения количества вещества, воздействующего на биологический объект, используется понятие доза. Под дозой понимается некоторое количественное значение агента (фактора), изменяющее состояние исследуемого объекта на введенную дозу.

В некоторых случаях (например, при статистической оценке возраста менархе или возраста менопаузы, см. [**?**], когда фиксируется возраст пациента и наличие или нет интересующего нас события, точный момент наступления которого неизвестен) отсутствует понятие дозы в вышеопределенном смысле, тем не менее, задача оценки возраста менархе или оценки возраста менопаузы укладывается в рассматриваемую в работе модель доза-эффект.

На современном этапе в токсикометрии востребованными являются величины доз, которые вызывают появление эффекта, учитываемого в экспериментальной группе тест-объектов с заданной вероятностью 0.01; 0.05; 0.1; 0.16; 0.5; 0.84; 0.9; 0.95; 0.99. Такие дозы получили название доз ED_1, ED_5, ED_{10}, ED_{16}, ED_{50}, ED_{84}, ED_{90}, ED_{95}, ED_{99}. Поэтому как среднеэффективная доза ED_{50}, так и другие категории доз: малые $(ED_1, ED_5, ED_{10}, ED_{16})$ и большие $(ED_{84}, ED_{90}, ED_{95}, ED_{99})$ дозы , должны в полной мере отвечать максимально жестким критериям корректности, надежности, адекватности и состоятельности.

В предложенной работе нас интересует проблема нахождения граничных (малых и больших) эффективных доз, для решения которой использованы оценки квантилей распределения, основанные на непараметрической оценке Надарая-Ватсона.

1. Постановка задачи

Рассмотрим следующую математическую модель зависимости «доза-эффект».

Пусть в биообъект вводится 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. Для каждого биообъекта эта доза будет различной, что определяется индивидуальной чувствительностью особей биологического вида к тестируемому препарату. Однако в однородно массе величина X будет случайной величиной.

Пусть $\{(X_i, U_i), 1 \leq i \leq n\}$ — потенциальная повторная выборка, вместо которой наблюдается выборка $\mathcal{U}^{(n)} = \{(U_i, W_i),\ 1 \leq i \leq n\}$, где $W_i = I(X_i < U_i)$ есть индикатор события $\{X_i < U_i\}$. Здесь U_i рассматриваются как вводимые дозы, а W_i — как эффект от воздействия этой дозы (U_i) на биообъект.

Как правило, модель зависимости «доза-эффект» предполагает величины U_i случайными (случайный план эксперимента), однако в некоторых случаях вводимые дозы известны заранее и, следовательно, могут рассматриваться как детерминированные величины. В данной работе рассматривается последний случай, называемый также фиксированным планом эксперимента.

Случайные величины X_i имеют функцию распределения $F(x)$, которую будем предполагать абсолютно непрерывной: $F(x) = \int_{-\infty}^{x} f(t)\,dt$, причем $f(t) > 0$.

2. Исходные предположения

В данном пункте определены оценки квантилей распределения, анализ асимптотического поведения которых составляет содержание данной работы. Здесь также опишем условия на введенные в модели математические объекты, в которых будем проводить дальнейшие выкладки.

Пусть $\{X_i,\ i = 1, \ldots, n\}$ — последовательность независимых одинаково распределенных на отрезке $[0, 1]$ случайных величин с функцией распределения $F(x)$, а $P = \{u_1,\ \ldots,\ u_n\}$ — совокупность точек разбиения отрезка $[0, 1]$, $0 < u_1 < \ldots < u_n < 1$.

Для оценки функции $F(x)$ будем использовать статистику, аналогичную статистике Надарая-Ватсона,

$$F_{nh_r}(x) = \frac{1}{nh_r} \sum_{i=1}^{n} K_r \left(\frac{x - u_i}{h_r} \right) W_i, \tag{2.1}$$

где $W_i = I(X_i < u_i)$ и $K_r(x)$ — финитная ядерная функция, т.е. является симметричной плотностью с компактным носителем.

Для оценок квантиля порядка λ распределения $F(x)$ будем использовать следующие статистики:

$$\hat{x}_{1,\lambda} = \frac{1}{nh_d} \sum_{i=1}^{n} \int\limits_{-\infty}^{\lambda} K_d \left(\frac{F_{nh_r}\left(i/n\right) - u}{h_d} \right) du; \tag{2.2}$$

$$\hat{x}_{2,\lambda} = \frac{\sum_{i=1}^{n} K_d \left(\frac{F_{nh_r}(i/n) - \lambda}{h_d} \right) \frac{i}{n}}{\sum_{i=1}^{n} K_d \left(\frac{F_{nh_r}(i/n) - \lambda}{h_d} \right)}; \tag{2.3}$$

$$\hat{x}_{3,\lambda} = \frac{\sum_{i=1}^{n} \int_{-\infty}^{\lambda} K_d \left(\frac{F_{nh_r}(i/n) - u}{h_d} \right) \frac{i}{n} du}{\sum_{i=1}^{n} \int_{-\infty}^{\lambda} K_d \left(\frac{F_{nh_r}(i/n) - u}{h_d} \right) du}; \tag{2.4}$$

$$\hat{x}_{4,\lambda} = F_{nh_r}^{-1}(\lambda). \tag{2.5}$$

Функция $K_d(x)$ является финитным ядром; величины $h_r = h_r(n)$ и $h_d = h_d(n)$ — сглаживающие параметры, они неслучайны и зависят от объема выборки n. Параметры h_r и h_d носят также название ширины окна просмотра данных.

Замечание 2.1. Оценка (**??**) была заимствована у H. Dette из [**?**] с некоторыми изменениями. Именно, в качестве оценки функции распределения в работе H. Dette была использована локальная оценка функции регрессии.

Основная идея самой оценки состоит в следующем. Суммирование в оценке (**??**) является приближением интеграла

$$\frac{1}{h_d} \int\limits_{0}^{1} \int\limits_{-\infty}^{\lambda} K_d \left(\frac{F_{nh_r}\left(x\right) - u}{h_d} \right) du dx.$$

Далее, если заменить $F_{nh_r}\left(x\right)$ на оцениваемую функцию распределения, то получим следующее выражение:

$$\frac{1}{h_d} \int\limits_{0}^{1} \int\limits_{-\infty}^{\lambda} K_d \left(\frac{F\left(x\right) - u}{h_d} \right) du dx = \int\limits_{0}^{1} I\left\{F(x) \le \lambda\right\} dx + o(1) = x_\lambda + o(1).$$

Сформулируем исходные предположения.

- Условия (H) на сглаживающие параметры h_r и h_d:

$\mathbf{H}_1:\ h_r \xrightarrow[n\to\infty]{} 0, \qquad h_d \xrightarrow[n\to\infty]{} 0;$

$\mathbf{H}_2:\ h_r/h_d \xrightarrow[n\to\infty]{} \infty;$

$\mathbf{H}_3:\ nh_r^5 = O(1);$

$\mathbf{H}_4:\ 1\Big/\left(nh_rh_d^{8/3}\right) = o(1).$

Условия (H) являются непротиворечивыми, поскольку последовательности $h_r = n^{-1/5}$ и $h_d = n^{-1/4}$ им удовлетворяют.

- Условия (K) на ядерные функции $K_r(x)$ и $K_d(x)$:

  $\mathbf{K}_1:\ \int_{-1}^{1} K_{r,d}(x)dx = 1.$

  $\mathbf{K}_2:\ K_{r,d}(x) \geq 0$, причем $K_{r,d}(x) = 0,\ x \notin [-1,1];$

  $\mathbf{K}_3:\ K_{r,d}(x) = K_{r,d}(-x),\ x \in \mathbb{R};$

  $\mathbf{K}_4:\ K_r(x)$ и $K_d(x)$ являются непрерывно дифференцируемой и четырежды непрерывно дифференцируемой на отрезке $[-1,1]$ функциями.

  $\mathbf{K}_5:$ Являются конечными следующие величины:

$$\nu_r^2 = \int K_r(z)z^2dz, \qquad \hat{\nu}_r^2 = \int K_r(z)z^4dz,$$
$$\nu_d^2 = \int K_d(z)z^2dz, \qquad \hat{\nu}_d^2 = \int K_d(z)z^4dz$$

- Условия (F) на функцию распределения $F(x)$:

  $\mathbf{F}_1:\ F(x)$ имеет четвертую непрерывную производную на [0,1];

  $\mathbf{F}_2:\ F'(x) = f(x) > 0, \forall x \in [0,1]$, т.е. $F(x)$ возрастает на $[0,1]$.

- Условия (P) на совокупность точек разбиения $P = \{u_1, \ldots, u_n\}$ отрезка $[0,1]$, $0 < u_1 < \ldots < u_n < 1$:

  $\mathbf{P}:$ Обозначим $u_0 = 0,\ u_{n+1} = 1$, тогда предположим, что

$$\max_{k=\overline{0;n}} \max\left\{\left|u_k - \frac{k}{n}\right|, \left|u_{k+1} - \frac{k}{n}\right|\right\} = O\left(\frac{1}{n}\right).$$

  Последнее условие фактически означает, что $u_k = \frac{k}{n} + O\left(\frac{1}{n}\right)$, где последовательность $nO\left(\frac{1}{n}\right)$ ограничена равномерно для всех $k = \overline{1,n}$.

Все приводимые утверждения в работе исходят из описанных в этом пункте условий (если не сделано иных оговорок), поэтому исходные предположения утверждений будем, как правило, опускать.

3. Вспомогательные утверждения

В этом пункте представлены общие результаты, необходимые для изучения асимптотики введенных выше оценок (**??**) — (**??**).

Сначала предложим без доказательства утверждение из книги [**?**], позволяющее оценить интеграл от функции по ее значениям в конечном числе точек.

Лемма 3.1 (Неравенство Koksma-Hlawka)**.** Если функция g имеет ограниченное изменение в кубе $[0;1]^s$ и $P = \{u_1;\ \ldots;\ u_n\}$ — множество точек из $[0;1]^s$, то

$$\left| \frac{1}{n} \sum_{k=1}^{n} g(u_k) - \int\limits_{[0,1]^s} g(u)du \right| \le D^*[P]V[g],$$

где

$$D^*[P] = \sup_{y_1,\ldots,y_s \in [0,1]} \left| \frac{A(y_1,\ldots,y_s)}{n} - y_1 \cdot \ldots \cdot y_s \right|,$$

$V[g]$ — вариация функции g в кубе $[0;1]^s$, а величина $A(y_1,\ldots,y_s)$ равна количеству точек из P, попавших в куб $[0,y_1) \times \ldots \times [0,y_s)$.

Замечание 3.1. При $s = 1$ величину $D^*[P]$, в соответствии с [**?**, теорема 2.6], можно определить как $\max_{k=\overline{0;n-1}} \max\{|u_k - \frac{k}{n}|, |u_{k+1} - \frac{k}{n}|\}$. Значит, в условиях (P), $D^*[P] = O\left(\frac{1}{n}\right)$.

Замечание 3.2. Положим в случае $s = 1$ функцию $g(u)$ равной нулю вне отрезка $[0,1]$. При достаточно малых h верно равенство $\int_0^1 g(u)du = \frac{1}{h}\int_0^1 g(\frac{u}{h})du$. Однако для одного и того же интеграла имеем две разные оценки:

$$\left| \frac{1}{n} \sum_{k=1}^{n} g(u_k) - \int\limits_0^1 g(u)du \right| \le V_0^1[g] O\left(\frac{1}{n}\right) = O\left(\frac{1}{n}\right),$$

$$\left| \frac{1}{n} \sum_{k=1}^{n} \frac{1}{h} g\left(\frac{u_k}{h}\right) - \frac{1}{h}\int\limits_0^1 g\left(\frac{u}{h}\right) du \right| = \left| \frac{1}{n} \sum_{k=1}^{n} \frac{1}{h} g\left(\frac{u_k}{h}\right) - \int\limits_0^1 g(u)du \right| \le$$
$$\le V_0^1[g_h] O\left(\frac{1}{n}\right) = O\left(\frac{1}{nh}\right),$$

поскольку $V_0^1[g_h] = V_0^1[g]/h$. Таким образом, последняя из двух оценок дает менее точный результат, что можно объяснить следующим: в сумме $\sum_{k=1}^{n} h^{-1} g\left(u_k h^{-1}\right)$ учитывается лишь часть наблюдений nh от всего объема n, который полностью учитывается в первой оценке.

Следующая лемма описывает асимптотику $E\left[F_{nh_r}(x)\right]$.

Лемма 3.2. Равномерно по $x \in (h_r; 1 - h_r)$ верно разложение

$$E\left[F_{nh_r}(x) - F(x)\right] = h_r^2 \frac{\nu_r^2}{2} F''(x) + O\left(\frac{1}{nh_r}\right),$$

где $F_{nh_r}(x)$ определена в (**??**), ν_r^2 см. стр.**??**.

Доказательство. Из определения $F_{nh_r}(x)$

$$E\left[F_{nh_r}(x) - F(x)\right] = \frac{1}{nh_r} \sum_{i=1}^{n} K_r\left(\frac{x - u_i}{h_r}\right) F(u_i) - F(x) =$$

применим неравенство Koksma-Hlawka

$$= \frac{1}{h_r}\int\limits_0^1 K_r\left(\frac{x-y}{h_r}\right)F(y)dy - F(x) + O\left(\frac{1}{nh_r}\right) =$$

$$= \int\limits_{(x-1)/h_r}^{x/h_r} K_r(t)F(x-h_rt)dt - F(x) + O\left(\frac{1}{nh_r}\right) =$$

поскольку $x \in (h_r; 1-h_r)$, заменим пределы интегрирования

$$= \int\limits_{-1}^{1} K_r(t)\left[F(x) - h_rtF'(x) + \frac{h_r^2t^2}{2}F''(x) - \frac{h_r^3t^3}{6}F'''(x) + \frac{h_r^4t^4}{24}F^{(4)}(x) + o(h_r^4)\right]dt -$$

$$- F(x) + O\left(\frac{1}{nh_r}\right) = h_r^2\frac{\nu_r^2}{2}F''(x) + h_r^4\frac{\hat{\nu}_r^2}{24}F^{(4)}(x) + o(h_r^4) + O\left(\frac{1}{nh_r}\right) =$$

$$= h_r^2F''(x)\frac{\nu_r^2}{2} + O\left(\frac{1}{nh_r}\right)$$

□

Следующая лемма описывает асимптотику центральных моментов величины $F_{nh_r}(x)$ порядка больше 1.

Лемма 3.3.

$$\mu_2 = O\left(\frac{1}{nh_r}\right),\ \mu_3 = O\left(\frac{1}{n^2h_r^2}\right),\ \mu_4 = O\left(\frac{1}{n^2h_r^2}\right),$$

равномерно по $x \in \mathbb{R}$, где

$$\mu_2 = E\left[F_{nh_r} - E\left(F_{nh_r}\right)\right]^2,$$
$$\mu_3 = E\left[F_{nh_r} - E\left(F_{nh_r}\right)\right]^3,$$
$$\mu_4 = E\left[F_{nh_r} - E\left(F_{nh_r}\right)\right]^4,$$

и F_{nh_r} определена в (**??**).

Доказательство. Из определения $F_{nh_r}(x)$ запишем

$$F_{nh_r}(x) = \sum_{j=1}^{n} D_j,\ \text{где } D_j = \frac{1}{nh_r}K_r\left(\frac{x-u_j}{h_r}\right)I\left(X_j < u_j\right),$$

причем все D_j независимы. Пусть $\varphi(t)$ — характеристическая функция случайной величины $F_{nh_r}(x)$, а $\varphi_j(t)$ — характеристические функции случайных величин D_j соответственно $(j = \overline{1;n})$.

Введем функцию

$$\psi(t) = \ln\varphi(t).$$

В силу независимости D_j, имеем

$$\psi(t) = \ln\prod_{j=1}^{n}\varphi_j(t) = \sum_{j=1}^{n}\ln\varphi_j(t) = \sum_{j=1}^{n}\psi_j(t),$$

где $\psi_j(t) = \ln\varphi_j(t)$.

Все рассматриваемые случайные величины принимают лишь конечное число значений, а значит имеют место следующие равенства (см. [?, с.209, 210]):

$$\varphi_j(t) = 1 + \sum_{\nu=1}^{\infty} \alpha_\nu \left(it\right)^\nu,$$

$$\psi_j(t) = \sum_{\nu=1}^{m} \frac{\chi_{\nu,j}}{\nu!} \left(it\right)^\nu + o\left(t^m\right), \forall m \in \mathbb{N}, \tag{3.1}$$

где α_ν — некоторые постоянные, а $\chi_{\nu,j}$ — семиинварианты распределения статистики D_j.

При сложении независимых случайных величин их семиинварианты складываются (см. [?, с.201]), поэтому

$$\chi_\nu = \sum_{j=1}^{n} \chi_{\nu,j},$$

χ_ν — семиинвариант распределения случайной величины $F_{nh_r}(x)$.

Выражения для центральных моментов тогда можно получить, следуя [?, с. 210], через семиинварианты, а именно:

$$\mu_2 = \chi_2, \tag{3.2}$$

$$\mu_3 = \chi_3, \tag{3.3}$$

$$\mu_4 = \chi_4 + 3\chi_2^2. \tag{3.4}$$

Непосредственные вычисления приводят к следующему:

$$\varphi_j(t) = (1 - F(u_j)) + F(u_j) e^{it\frac{1}{nh_r}K_r\left(\frac{x-u_j}{h_r}\right)} = \left[e^x = \sum_{s=0}^{\infty} \frac{x^s}{s!}\right] = 1 + a\sum_{s=1}^{\infty} \frac{\left(it\right)^s}{s!} b^s,$$

где

$$a = F(u_j), \quad b = \frac{1}{nh_r} K_r\left(\frac{x-u_j}{h_r}\right). \tag{3.5}$$

Далее, учитывая разложение Тейлора $\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + o\left(x^4\right)$ для $|x| < 1$, имеем:

$$\begin{aligned}
\psi_j(t) = \ln\varphi_j(t) = \ln\left\{1 + a\sum_{s=1}^{\infty} \frac{\left(it\right)^s}{s!} b^s\right\} = \\
= a\left\{itb + \frac{\left(it\right)^2}{2!}b^2 + \frac{\left(it\right)^3}{3!}b^3 + \frac{\left(it\right)^4}{4!}b^4\right\} - \frac{a^2}{2}\left\{itb + \frac{\left(it\right)^2}{2!}b^2 + \frac{\left(it\right)^3}{3!}b^3\right\}^2 + \\
+ \frac{a^3}{3}\left\{itb + \frac{\left(it\right)^2}{2!}b^2\right\}^3 - \frac{a^4}{4}\left(it\right)^4 b^4 + o\left(t^4\right) = itba + \left(it\right)^2 b^2 \left\{\frac{a}{2} - \frac{a^2}{2}\right\} + \\
+ \left(it\right)^3 b^3 \left\{\frac{a}{3!} - \frac{a^2}{2} + \frac{a^3}{3}\right\} + \left(it\right)^4 b^4 \left\{\frac{a}{4!} - \frac{a^2}{2}\frac{1}{4} - \frac{a^2}{2}\frac{2}{3!} + \frac{a^3}{3}\frac{3}{2} - \frac{a^4}{4}\right\} + o\left(t^4\right) = \\
= \frac{ba}{1!}it + \frac{b^2 a(1-a)}{2!}\left(it\right)^2 + \frac{b^3 a(a-1)(a-1/2)}{3!}\left(it\right)^3 + \\
+ \frac{b^4 a(1-a)(6a^2-6a+1)}{4!}\left(it\right)^4 + o\left(t^4\right).
\end{aligned}$$

Приравнивая коэффициенты последнего разложения к коэффициентам разложения (**??**), получим:

$$\chi_{2,j} = b^2 a(1-a), \qquad \chi_2 = \sum_{j=1}^{n} \chi_{2,j};$$

$$\chi_{3,j} = b^3 a(a-1)(a-1/2), \qquad \chi_3 = \sum_{j=1}^{n} \chi_{3,j};$$

$$\chi_{4,j} = b^4 a(1-a)(6a^2-6a+1), \qquad \chi_4 = \sum_{j=1}^{n} \chi_{4,j},$$

где a и b определяются из (**??**).

Теперь перейдем к вопросу об асимптотике величин χ_k.

$$\chi_2 = \sum_{j=1}^{n} \frac{1}{n^2 h_r^2} F(u_j)(1-F(u_j)) K_r^2\left(\frac{x-u_j}{h_r}\right) =$$
$$= \frac{1}{nh_r^2} \int\limits_0^1 F(y)(1-F(y)) K_r^2\left(\frac{x-y}{h_r}\right) dy + O\left(\frac{1}{n^2 h_r^2}\right).$$

Сделаем замену $z = \frac{x-y}{h_r}$ и оставим главный член

$$\chi_2 = \frac{1}{nh_r} F(x)(1-F(x)) \|K_r\|^2 + o\left(\frac{1}{nh_r}\right).$$

$$\chi_3 = \frac{1}{n^3 h_r^3} \sum_{j=1}^{n} G_1(u_j) K_r^3\left(\frac{x-u_j}{h_r}\right) =$$
$$= \frac{1}{n^2 h_r^3} \int\limits_0^1 G_1(y) K_r^3\left(\frac{x-y}{h_r}\right) dy + O\left(\frac{1}{n^3 h_r^3}\right) =$$
$$= \frac{1}{n^2 h_r^2} G_1(x) \int K_r^3(z) dz + o\left(\frac{1}{n^2 h_r^2}\right),$$

где $G_1(x) = F(x)(F(x)-1)(F(x)-1/2)$.

$$\chi_4 = \frac{1}{n^4 h_r^4} \sum_{j=1}^{n} G_2(u_j) K_r^4\left(\frac{x-u_j}{h_r}\right) = \frac{1}{n^3 h_r^3} G_2(x) \int K_r^4(z) dz + o\left(\frac{1}{n^3 h_r^3}\right),$$

где $G_2(x) = F(x)(1-F(x))(6F^2(x)-6F(x)+1)$.

В итоге для центральных моментов величины $F_{nh_r}(x)$ равномерно по x имеем следующее асимптотическое разложение (см. (**??**)-(**??**)):

$$\mu_2 = \frac{1}{nh_r} F(x)(1-F(x)) \|K_r\|^2 + o\left(\frac{1}{nh_r}\right),$$
$$\mu_3 = \frac{1}{n^2 h_r^2} G_1(x) \int K_r^3(z) dz + o\left(\frac{1}{n^2 h_r^2}\right),$$
$$\mu_4 = \frac{1}{n^2 h_r^2} F^2(x)(1-F(x))^2 \|K_r\|^4 + o\left(\frac{1}{n^2 h_r^2}\right),$$

чем завершается доказательство леммы. □

Из этой леммы легко получить асимптотику моментов случайной величины $\Delta_F = F_{nh_r}(x) - F(x)$, которая также будет необходима в дальнейшем.

Следствие 3.1.

$$E\left[\Delta_F^2\right] = O\left(\frac{1}{nh_r}\right),\ E\left[|\Delta_F|^3\right] = O\left(\frac{1}{(nh_r)^{3/2}}\right),\ E\left[\Delta_F^4\right] = O\left(\frac{1}{n^2 h_r^2}\right),$$

Доказательство. Заменим Δ_F суммой

$$\Delta_F = \Delta_{F,1} + \Delta_{F,2},$$

где $\Delta_{F,1} = F_{nh_r}(x) - E\left[F_{nh_r}(x)\right]$ и $\Delta_{F,2} = E\left[F_{nh_r}(x)\right] - F(x) = O\left(h_r^2\right)$. Используя лемму **??**, где

$$\mu_2 = E\left[\Delta_{F,1}^2\right],\ \mu_3 = E\left[\Delta_{F,1}^3\right],\ \mu_4 = E\left[\Delta_{F,1}^4\right],$$

распишем последовательно вычисления:

$$E\left[\Delta_F^2\right] = E\left[\Delta_{F,1} + \Delta_{F,1}\right]^2 = E\left[\Delta_{F,1}^2\right] + E\left[\Delta_{F,2}^2\right] = \mu_2 + \Delta_{F,2}^2 = \\ = O\left(\frac{1}{nh_r}\right) + O\left(h_r^4\right) = O\left(\frac{1}{nh_r}\right);$$

$$E\left[\Delta_F^4\right] = E\left[\Delta_{F,1} + \Delta_{F,1}\right]^4 = E\left[\Delta_{F,1}^4\right] + 4E\left[\Delta_{F,1}^3\Delta_{F,2}\right] + 6E\left[\Delta_{F,1}^2\Delta_{F,2}^2\right] + \\ + 4E\left[\Delta_{F,1}\Delta_{F,2}^3\right] + E\left[\Delta_{F,2}^4\right] = \mu_4 - 4\Delta_{F,2}\mu_3 + 6\Delta_{F,2}^2\mu_2 + \Delta_{F,2}^4 = \\ = O\left(\frac{1}{n^2h_r^2}\right) + O\left(\frac{h_r^2}{n^2h_r^2}\right) + O\left(\frac{h_r^4}{nh_r}\right) + O\left(h_r^8\right) = O\left(\frac{1}{n^2h_r^2}\right);$$

третий момент оценим четвертым (см. [**?**, с.198])

$$E\left[\left|\Delta_F\right|^3\right] \le \left\{E\left[\Delta_F^4\right]\right\}^{3/4} = O\left(\frac{1}{\left(nh_r\right)^{3/2}}\right)$$

Все полученные асимптотики равномерны по x, а значит утверждение леммы верно. $\square$

4. Основные результаты

4.1. Асимптотика оценки $\hat{x}_{1,\lambda}$

Представим статистику $\hat{x}_{1,\lambda}$ в следующем виде:

$$\hat{x}_{1,\lambda} = \frac{1}{nh_d} \sum_{i=1}^{n} \int\limits_{-\infty}^{\lambda} K_d \left(\frac{F_{nh_r}\left(i/n\right) - u}{h_d} \right) du = x_{\lambda,n} + \Delta,$$

где

$$x_{\lambda,n} = \frac{1}{nh_d} \sum_{i=1}^{n} \int\limits_{-\infty}^{\lambda} K_d \left(\frac{F\left(i/n\right) - u}{h_d} \right) du,$$

$$\Delta = \frac{1}{nh_d} \sum_{i=1}^{n} \int\limits_{-\infty}^{\lambda} \left\{ K_d \left(\frac{F_{nh_r}\left(i/n\right) - u}{h_d} \right) - K_d \left(\frac{F\left(i/n\right) - u}{h_d} \right) \right\} du.$$

Чтобы применить к $x_{\lambda,n}$ неравенство Koksma-Hlawka (см. лемму **??**) оценим вариацию функции

$$\tilde{f}(x) = \frac{1}{h_d} K_d \left(\frac{F(x) - \lambda}{h_d} \right)$$

на отрезке $[0;1]$.

Лемма 4.1.

$$V\left[\tilde{f}\right] = \sup \sum_{j=1}^{l} \left| \tilde{f}(x_j) - \tilde{f}(x_{j-1}) \right| = O\left(\frac{1}{h_d}\right).$$

Доказательство. Пусть $0 \le x_0 \le x_1 \le \ldots \le x_l \le 1$, тогда

$$\sum_{j=1}^{l} \left| \tilde{f}(x_j) - \tilde{f}(x_{j-1}) \right| = \frac{1}{h_d} \sum_{j=1}^{l} \left| K_d \left(\frac{F(x_j) - \lambda}{h_d} \right) - K_d \left(\frac{F(x_{j-1}) - \lambda}{h_d} \right) \right| =$$

$$= \frac{1}{h_d} \left\{ \sum_{j=1}^{l_1} (\ldots) + \sum_{j=l_2+1}^{l} (\ldots) + K_d \left(\frac{F(x_{l_1+1}) - \lambda}{h_d} \right) + \right.$$

$$\left. + K_d \left(\frac{F(x_{l_2-1}) - \lambda}{h_d} \right) \right\} + \frac{1}{h_d} \sum_{j=l_1+2}^{l_2-1} (\ldots),$$

где l_1 и l_2 таковы, что

$$F(x_{l_1}) \le \lambda - h_d; \qquad F(x_{l_1+1}) > \lambda - h_d,$$
$$F(x_{l_2}) \ge \lambda + h_d; \qquad F(x_{l_2-1}) < \lambda + h_d.$$

Поскольку $K_d(x) = 0$ для $|x| \ge 1$, то сумма $\sum_{j=1}^{l_1} (\ldots) + \sum_{j=l_2+1}^{l} (\ldots)$ занулится и

$$K_d \left(\frac{F(x_{l_1+1}) - \lambda}{h_d} \right) = K_d(-1) + K_d'(\xi) \left(\frac{F(x_{l_1+1}) - \lambda}{h_d} + 1 \right) \xrightarrow[n \to \infty]{} 0$$

при уменьшении ранга разбиения (за счет выбора l_1); аналогично можно показать, что $K_d \left(\frac{F(x_{l_2-1}) - \lambda}{h_d} \right) \xrightarrow[n \to \infty]{} 0$ (за счет выбора l_2).

В оставшемся ненулевом слагаемом все точки $\frac{F(x_j) - \lambda}{h_d}$ принадлежат отрезку $[-1;1]$, и значит, имеет место разложение Тейлора:

$$\frac{1}{h_d}\sum_{j=l_1+2}^{l_2-1}\left|K_d\left(\frac{F(x_j)-\lambda}{h_d}\right)-K_d\left(\frac{F(x_{j-1})-\lambda}{h_d}\right)\right| =$$

$$= \frac{1}{h_d}\sum_{j=l_1+2}^{l_2-1}K'_d(\xi_j)\left(\frac{F(x_j)-F(x_{j-1})}{h_d}\right) \le$$

$$\le \frac{1}{h_d^2}M\left\{F(x_{l_2-1})-F(x_{l_1+1})\right\} \le \frac{2}{h_d^2}Mh_d = \frac{2M}{h_d},$$

где $\xi_j \in [-1;1]$ и $\left|K'_d(\xi_j)\right| \le M$. Видно, что M не зависит от объема выборки n. Из этой оценки следует результат леммы. □

Асимптотическое поведение $x_{\lambda,n}$ представлено в следующей лемме.

Лемма 4.2.

$$x_{\lambda,n} = F^{-1}(\lambda) + a_{2,d}h_d^2 + o\left(h_d^2\right),$$

где

$$a_{2,d} = \frac{\nu_d^2}{2}\left(F^{-1}\right)''(\lambda),$$

а ν_d^2 определена на стр. **??**.

Доказательство. Из предыдущей леммы следует, что к $x_{\lambda,n}$ можно применить неравенство Koksma-Hlawka (см. лемму **??**):

$$x_{\lambda,n} = \frac{1}{h_d}\int\limits_0^1 dx\int\limits_{-\infty}^{\lambda} K_d\left(\frac{F(x)-u}{h_d}\right)du + O\left(\frac{1}{nh_d}\right) =$$

$$= \int\limits_0^1 dx\int\limits_{\frac{F(x)-\lambda}{h_d}}^{\infty} K_d(z)dz + O\left(\frac{1}{nh_d}\right) = \int\limits_0^1 dx\int\limits_{\frac{F(x)-\lambda}{h_d}}^{1} K_d(z)dz + O\left(\frac{1}{nh_d}\right).$$

Поскольку $\frac{F(x)-\lambda}{h_d} \le -1$ при $x \le \left(F^{-1}\right)(\lambda - h_d) \le 1$, то

$$x_{\lambda,n} = \int\limits_0^{F^{-1}(\lambda-h_d)} dx\int\limits_{-1}^{1} K_d(z)dz + \int\limits_{F^{-1}(\lambda-h_d)}^{1} dx\int\limits_{\frac{F(x)-\lambda}{h_d}}^{1} K_d(z)dz + O\left(\frac{1}{nh_d}\right).$$

Первый интеграл, очевидно, равен $\left(F^{-1}\right)(\lambda-h_d)$, а во втором сделаем замену $w = \frac{F(x)-\lambda}{h_d}$ и учтем, что $\lambda < F(1) = 1$ и $F(x) \in C^2$. Тогда

$$x_{\lambda,n} = F^{-1}(\lambda-h_d) + h_d\int\limits_{-1}^{\frac{F(1)-\lambda}{h_d}} dw\int\limits_w^1 K_d(z)\left(F^{-1}\right)'(\lambda+h_dw)dz+$$

$$+ O\left(\frac{1}{nh_d}\right) = F^{-1}(\lambda-h_d) + h_d\int\limits_{-1}^{1} dw\int\limits_w^1 K_d(z)\times\left\{\left(F^{-1}\right)'(\lambda)+\right.$$

$$\left. + \left(F^{-1}\right)''(\lambda)wh_d + o\left(h_d\right)\right\}dz + O\left(\frac{1}{nh_d}\right)$$

Так как

$$\int\limits_{-1}^{1} dw \int\limits_{w}^{1} K_d(z)dz = 1, \qquad \int\limits_{-1}^{1} dw \int\limits_{w}^{1} K_d(z)wdz = frac12\nu_d^2 - \frac{1}{2},$$

то

$$\begin{aligned} x_{\lambda,n} &= F^{-1}(\lambda - h_d) + \\ &+ h_d \left\{ \left(F^{-1}\right)'(\lambda) + \left(F^{-1}\right)''(\lambda) h_d \int\limits_{-1}^{1}\int\limits_{w}^{1} wK_d(z)dz + o\left(h_d\right) \right\} + \\ &+ O\left(\frac{1}{nh_d}\right) = F^{-1}(\lambda) + h_d^2 \left(F^{-1}\right)''(\lambda) \left\{ \frac{1}{2} + frac12\nu_d^2 - \frac{1}{2} \right\} + o\left(h_d^2\right) = \\ &= F^{-1}(\lambda) + h_d^2 \left(F^{-1}\right)''(\lambda) frac12\nu_d^2 + o\left(h_d^2\right). \end{aligned}$$

□

Далее рассмотрим асимптотику величины Δ.
Представим ее в виде суммы $\Delta = \Delta_1 + \Delta_2 + \Delta_3 + \Delta_4$, где

$$\Delta_1 = \frac{1}{nh_d^2} \sum_{i=1}^{n} \int\limits_{-\infty}^{\lambda} K_d' \left(\frac{F\left(i/n\right) - u}{h_d} \right) \left(F_{nh_r}\left(i/n\right) - F\left(i/n\right) \right) du, \tag{4.1.1a}$$

$$\Delta_2 = \frac{1}{nh_d^3} \sum_{i=1}^{n} \int\limits_{-\infty}^{\lambda} K_d'' \left(\frac{F\left(i/n\right) - u}{h_d} \right) \left(F_{nh_r}\left(i/n\right) - F\left(i/n\right) \right)^2 du, \tag{4.1.1b}$$

$$\Delta_3 = \frac{1}{nh_d^4} \sum_{i=1}^{n} \int\limits_{-\infty}^{\lambda} K_d''' \left(\frac{F\left(i/n\right) - u}{h_d} \right) \left(F_{nh_r}\left(i/n\right) - F\left(i/n\right) \right)^3 du, \tag{4.1.1c}$$

$$\Delta_4 = \frac{1}{nh_d^5} \sum_{i=1}^{n} \int\limits_{-\infty}^{\lambda} K_d^{(4)} \left(\frac{\xi_i - u}{h_d} \right) \left(F_{nh_r}\left(i/n\right) - F\left(i/n\right) \right)^4 du, \tag{4.1.1d}$$

где $\left|\xi_i - F\left(i/n\right)\right| \le \left|F\left(i/n\right) - F_{nh_r}\left(i/n\right)\right|$.
Асимптотика главной части статистики Δ, величины Δ_1, дается следующей леммой.

Лемма 4.3.

$$\left(\Delta_1 - a_{2,r}h_r^2\right)\sqrt{nh_r} \xrightarrow{d} N\left(0; g_2^2\right),$$

где $\|K_r\|^2 = \int K_r^2(x)dx$, ν_r^2 определена на стр. **??** и

$$a_{2,r} = -\frac{\nu_r^2}{2} F''\left(F^{-1}(\lambda)\right)\left(F^{-1}\right)'(\lambda),$$
$$g_2^2 = \lambda(1-\lambda)\|K_r\|^2 \left(\left(F^{-1}\right)'(\lambda)\right)^2.$$

,

Доказательство. Очевидно, что

$$\Delta_1 = -\frac{1}{nh_d} \sum_{i=1}^{n} K_d \left(\frac{F\left(i/n\right) - \lambda}{h_d} \right) \left(F_{nh_r}\left(i/n\right) - F\left(i/n\right) \right).$$

Введем величины

$$\Delta_{1,1} = -\frac{1}{nh_d}\sum_{i=1}^{n} K_d\left(\frac{F\left(i/n\right)-\lambda}{h_d}\right)\left(F_{nh_r}\left(i/n\right) - E\left[F_{nh_r}\left(i/n\right)\right]\right);$$
$$\Delta_{1,2} = -\frac{1}{nh_d}\sum_{i=1}^{n} K_d\left(\frac{F\left(i/n\right)-\lambda}{h_d}\right)\left(E\left[F_{nh_r}\left(i/n\right)\right] - F\left(i/n\right)\right),$$

Тогда $\Delta_1 = \Delta_{1,1} + \Delta_{1,2}$, причем величина $\Delta_{1,2}$ не случайна.

Из леммы **??** используем выражение $E\left[F_{nh_r}(x) - F(x)\right] = h_r^2\frac{\nu_r^2}{2}F''(x) + o\left(h_r^2\right)$ и получим

$$E\left[\Delta_1\right] = E\left[\Delta_{1,2}\right] = -\frac{\nu_r^2 h_r^2}{2nh_d}\sum_{i=1}^{n} K_d\left(\frac{F\left(i/n\right)-\lambda}{h_d}\right)F''(x)(1+o(1)) = -\frac{\nu_r^2 h_r^2}{2h_d}\times$$
$$\times\int\limits_0^1 K_d\left(\frac{F(x)-\lambda}{h_d}\right)F''(x)dx(1+o(1)) + O\left(\frac{h_r^2}{nh_d}\right) = -\frac{\nu_r^2 h_r^2}{2h_d}\int\limits_{-1}^1 K_d(z)x_z' F''(x)dz\times$$
$$\times\left(1+o(1)\right) + nh_r^2 nh_d = -\frac{\nu_r^2}{2}h_r^2\left(F^{-1}\right)'(\lambda)F''\left(F^{-1}(\lambda)\right) + o(h_r^2). \quad (4.1.2)$$

Замечание 4.1. На самом деле здесь мы использовали выражение для $E\left[F_{nh_r}(x) - F(x)\right]$, учитывая равномерность по $x \in (0;1)$, что не является обоснованным. Однако рассуждения остаются в силе, если провести следующие выкладки:

$$\sum_{i=1}^{n} K_d\left(\frac{F(i/n)-\lambda}{h_d}\right)\left(E\left[F_{nh_r}(x) - F(x)\right]\right) = \sum_{i=l_1}^{l_2} K_d\left(\frac{F(i/n)-\lambda}{h_d}\right)\times$$
$$\times\left(F''(i/n)\nu_r^2\left(2nh_d\right)^{-1}h_r^2 + o\left(h_r^2\right)\right) + \sum_{i=1}^{l_1-1} K_d\left(\frac{F(i/n)-\lambda}{h_d}\right)\left(E\left[F_{nh_r}(x)\ -\right.\right.$$
$$\left.\left.-\ F(x)\right]\right) + \sum_{i=l_2+1}^{n} K_d\left(\frac{F(i/n)-\lambda}{h_d}\right)\left(E\left[F_{nh_r}(x) - F(x)\right]\right) =$$

где l_1 и l_2 ограничивают аргумент i/n в пределах $(h_r, 1-h_r)$ и тогда при больших n последние две суммы занулятся

$$= \sum_{i=1}^{n} K_d\left(\frac{F(i/n)-\lambda}{h_d}\right)F''(i/n)\nu_r^2\left(2nh_d\right)^{-1}h_r^2\left(1+o\left(1\right)\right) -$$
$$-\sum_{i=1}^{l_1-1} K_d\left(\frac{F(i/n)-\lambda}{h_d}\right)F''(i/n)\nu_r^2\left(2nh_d\right)^{-1}h_r^2\left(1+o\left(1\right)\right) -$$
$$-\sum_{i=l_2+1}^{n} K_d\left(\frac{F(i/n)-\lambda}{h_d}\right)F''(i/n)\nu_r^2\left(2nh_d\right)^{-1}h_r^2\left(1+o\left(1\right)\right) =$$
$$= \sum_{i=1}^{n} K_d\left(\frac{F(i/n)-\lambda}{h_d}\right)F''(i/n)\nu_r^2\left(2nh_d\right)^{-1}h_r^2\left(1+o\left(1\right)\right)$$

В дальнейшем не будем подробно останавливаться на этом вопросе и будем неявно предполагать проделанные выше рассуждения.

Далее вычислим дисперсию Δ_1.

$Var\left[\Delta_1\right] = Var\left[\Delta_{1,1}\right] =$

$$= \frac{1}{n^4 h_d^2 h_r^2} \sum_{j=1}^{n} F(u_j)(1-F(u_j)) \left\{ \sum_{i=1}^{n} K_d \left(\frac{F\left(i/n\right) - \lambda}{h_d} \right) K_r \left(\frac{i/n - u_j}{h_r} \right) \right\}^2 =$$

$$= \frac{1}{n^2 h_d^2 h_r^2} \sum_{j=1}^{n} F(u_j)(1-F(u_j)) \left\{ \int_0^1 K_d \left(\frac{F\left(x\right) - \lambda}{h_d} \right) K_r \left(\frac{x - u_j}{h_r} \right) dx + O\left(\frac{1}{n}\right) \right\}^2;$$

сделаем замену $z = \frac{F(x)-\lambda}{h_d}$ и применим неравенство Koksma-Hlawka:

$$Var\left[\Delta_1\right] =$$

$$= \frac{1}{n h_d^2 h_r^2} \int_0^1 F(y)(1-F(y)) \left\{ \int_{-1}^{1} K_d(z) K_r \left(\frac{x-y}{h_r} \right) x_z' dz + O\left(\frac{1}{n}\right) \right\} dy + O\left(\frac{1}{n^2 h_r^2}\right).$$

Выпишем отдельно

$$K_r \left(\frac{x-y}{h_r} \right) \quad = \quad K_r \left(\frac{F^{-1}(\lambda + h_d z) - y}{h_r} \right) \quad = \quad K_r \left(\frac{F^{-1}(\lambda) - y}{h_r} + o(1) \right).$$

Учитывая последние выкладки и делая замену $t = \frac{F^{-1}(\lambda)-y}{h_r}$, запишем окончательно

$$Var\left[\Delta_1\right] = \frac{1}{n h_r} \lambda(1-\lambda) \left[\left(F^{-1}\right)'(\lambda) \right]^2 \|K_r\|^2 + o\left(\frac{1}{n h_r}\right). \tag{4.1.3}$$

Чтобы доказать асимптотическую нормальность величины Δ_1 достаточно доказать асимптотическую нормальность $\Delta_{1,1}$, поскольку $\Delta_{1,2}$ не случайна и ее предельное поведение нам известно. Для этого представим $\Delta_{1,1}$в виде суммы $\Delta_{1,1} = \sum_{j=1}^{n} \xi_j$, где

$$\xi_j = -\frac{1}{n^2 h_d h_r} \Big(I\left(X_j < u_j\right) - F(u_j) \Big) \sum_{i=1}^{n} K_d \left(\frac{F\left(i/n\right) - \lambda}{h_d} \right) K_r \left(\frac{i/n - u_j}{h_r} \right).$$

Имеем

$$\sum_{j=1}^{n} E\Big[\xi_j - E\left[\xi_j\right]\Big]^4 = \sum_{j=1}^{n} E\left[\xi_j\right]^4 = \frac{1}{n^8 h_d^4 h_r^4} \sum_{j=1}^{n} E\Big[I\left(X_j < u_j\right) - F(u_j) \Big]^4 \times$$

$$\times \left\{ \sum_{i=1}^{n} K_d \left(\frac{F\left(i/n\right) - \lambda}{h_d} \right) K_r \left(\frac{i/n - u_j}{h_r} \right) \right\}^4 =$$

обозначим $G(u) = F(u) - 4F^2(u) + 6F^3(u) - 3F^4(u)$ и продолжим

$$= \frac{1}{n^8 h_d^4 h_r^4} \sum_{j=1}^{n} G(u_j) \left\{ \int_0^1 K_d \left(\frac{F\left(x\right) - \lambda}{h_d} \right) K_r \left(\frac{x - u_j}{h_r} \right) dx + O\left(\frac{1}{n}\right) \right\}^4 =$$

$$= \frac{1}{n^4 h_r^4} \sum_{j=1}^{n} G(u_j) \left\{ \int_0^1 K_d\left(y\right) K_r \left(\frac{x - u_j}{h_r} \right) x_y' dy + O\left(\frac{1}{n h_d}\right) \right\}^4 =$$

$$= \frac{1}{n^3h_r^3}\int\limits_{-1}^{1} G(x - zh_r)\left\{\int\limits_{-1}^{1} K_d(y)K_r(z)x'_y dy\right\}^4 dz + O\left(\frac{1}{n^4h_r^4}\right) = O\left(\frac{1}{n^3h_r^3}\right).$$

Далее, поскольку

$$\frac{\sum_{j=1}^n E\left[\xi_j - E\left[\xi_j\right]\right]^4}{\left(Var\left[\sum_{j=1}^n \xi_j\right]\right)^2} = \frac{\sum_{j=1}^n E\left[\xi_j - E\left[\xi_j\right]\right]^4}{\left(Var\left[\Delta_1\right]\right)^2} = O\left(\frac{1}{nh_r}\right) \xrightarrow[n\to\infty]{} 0,$$

то к последовательности $\left\{\xi_j\right\}$ можем применить центральную предельную теорему Ляпунова. Суммируя последний результат с (**??**) и (**??**), получим утверждение леммы. $\square$

Далее перейдем к рассмотрению остатка величины Δ_1, который для краткости обозначим

$$D = \Delta_2 + \Delta_3 + \Delta_4,$$

см. (**??**)-(**??**).

В следующей лемме показано, что величина D является бесконечно малой величиной высшего порядка по сравнению с главным членом, величиной Δ_1.

Лемма 4.4.

$$D = o_d\left(\frac{1}{\sqrt{nh_r}}\right).$$

Доказательство. Проанализируем последовательно все три величины: Δ_2, Δ_3 и Δ_4, принимая во внимание следствие (**??**).

$$E\left|\Delta_2\right| \le \frac{1}{2nh_d^2}\sum_{i=1}^n \left|K''_d\left(\frac{F\left(i/n\right) - \lambda}{h_d}\right)\right| E\left[F_{nh_r}\left(i/n\right) - F\left(i/n\right)\right]^2 =$$

$$= O\left(\frac{1}{n^2h_rh_d^2}\right)\sum_{i=1}^n \left|K''_d\left(\frac{F\left(i/n\right) - \lambda}{h_d}\right)\right| =$$

$$= O\left(\frac{1}{nh_rh_d^2}\right)\int\limits_0^1 \left|K''_d\left(\frac{F\left(y\right) - \lambda}{h_d}\right)\right| dy + O\left(\frac{1}{n^2h_rh_d^2}\right) =$$

$$= O\left(\frac{1}{nh_rh_d^2}\right) = o\left(\frac{1}{\sqrt{nh_r}}\right);$$

$$E\left|\Delta_3\right| \le \frac{1}{6nh_d^3}\sum_{i=1}^n \left|K'''_d\left(\frac{F\left(i/n\right) - \lambda}{h_d}\right)\right| E\left[\left|F_{nh_r}\left(i/n\right) - F\left(i/n\right)\right|^3\right] =$$

$$= O\left(\frac{1}{n^{5/2}h_r^{3/2}h_d^3}\right)\sum_{i=1}^n \left|K'''_d\left(\frac{F\left(i/n\right) - \lambda}{h_d}\right)\right| = O\left(\frac{1}{n^{3/2}h_r^{3/2}h_d^2}\right) = o\left(\frac{1}{\sqrt{nh_r}}\right);$$

$$E\left|\Delta_4\right| \le \frac{1}{24nh_d^4}\sum_{i=1}^n \left|K^{(4)}_d\left(\frac{F\left(i/n\right) - \lambda}{h_d}\right)\right| E\left[F_{nh_r}\left(i/n\right) - F\left(i/n\right)^4\right] =$$

$$= O\left(\frac{1}{n^2h_r^2h_d^4}\right) = o\left(\frac{1}{\sqrt{nh_r}}\right);$$

Здесь мы также использовали тот факт, что $K_d^{(4)}(x)$ ограничена.

Таким образом, из неравенства Чебышева и оценки

$$|D| \le |\Delta_2| + |\Delta_3| + |\Delta_4|$$

получаем результат леммы. □

Суммируя леммы (**??**), (**??**) и (**??**), получаем следующую теорему:

Теорема 4.1. Пусть выполнены исходные предположения (H), (K), (F) и (P), тогда верно следующее предельное соотношение:

$$\left(\hat{x}_{1,\lambda} - F^{-1}(\lambda) - b_2\left(h_r, h_d\right)\right)\sqrt{nh_r} \xrightarrow{d} N\left(0; g_2^2\right),$$

где

$$b_2\left(h_r, h_d\right) = a_{2,d}h_d^2 + a_{2,r}h_r^2,$$
$$g_2^2 = \lambda(1-\lambda)\|K_r\|^2\left(\left(F^{-1}\right)'(\lambda)\right)^2$$

и

$$a_{2,r} = -\frac{\nu_r^2}{2}F''\left(F^{-1}(\lambda)\right)\left(F^{-1}\right)'(\lambda),$$
$$a_{2,d} = \frac{\nu_d^2}{2}\left(F^{-1}\right)''(\lambda),$$

ν_r^2, ν_d^2 из исходных предположений (стр. **??**) и $\|K_r\|^2 = \int K_r^2(x)dx$.

Замечание 4.2. Из сформулированной теоремы также следует состоятельность оценки $\hat{x}_{1,\lambda}$.

4.2. Асимптотика оценки $\hat{x}_{2,\lambda}$

4.2.1. Существование оценки $\hat{x}_{2,\lambda}$

Величина $\hat{x}_{2,\lambda}$, см. (**??**), может быть не определена для некоторых элементарных исходов, а именно в случае равенства нулю величины

$$\beta = \sum_{i=1}^{n} K_d\left(\frac{F_{nh_r}(i/n) - \lambda}{h_d}\right), \tag{4.2.1}$$

которая стоит в знаменателе $\hat{x}_{2,\lambda}$. В ряде случаев это не имеет существенного значения, поскольку исходная статистика может быть заменена эквивалентной (т.е. имеющая такое же распределение вероятностей), которая в свою очередь в ноль не обращается. В данном случае ситуация иная, но и она допускает схожее решение.

Оформим это замечание в виде леммы. Обозначим

$$A = \bigcap_{i=1}^{n} \left\{ K_d\left(\frac{F_{nh_r}(i/n) - \lambda}{h_d}\right) = 0 \right\}. \tag{4.2.2}$$

Лемма 4.5. Вероятность того, что случайная величина β, определенная в (**??**), равна нулю, положительна, т.е.

$$P\left(\beta = 0\right) > 0.$$

Доказательство. Поскольку величина β неотрицательна, то событие $\{\beta = 0\}$ равносильно событию A. Далее, в силу того, что $K_d(x) = 0$ при $|x| \geq 1$ и $\inf\limits_{x\in[-1+\delta;1-\delta]} K_d(x) > 0$ при $\delta \in (0;1)$, верно равенство

$$A = \bigcap_{i=1}^{n} \left\{ \left|F_{nh_r}(i/n) - \lambda\right| \geq h_d \right\}.$$

Рассмотрим событие $B = \bigcap_{j=1}^{n} \left\{X_j \geq u_j\right\}$, где величины X_j и u_j определены в исходной модели. Вероятность $P(B)$ равна $\prod_{j=1}^{n} \left(1 - F(u_j)\right)$ и отлична от нуля при $u_j < 1$ (Заметим, что u_j неслучайны и находятся в нашем распоряжении и, значит, $P(B)$ действительно отлична от нуля). Докажем, что $B \subseteq A$, откуда заключим $P(A) \geq P(B) > 0$.

Пусть произошло случайное событие, благоприятствующее событию B, т.е. $\omega \in B$. Тогда $F_{nh_r}(x) = 0, \forall x \in \mathbb{R}$ и при достаточно больших n (и малых h_d) $\left|F_{nh_r}(i/n) - \lambda\right| \geq h_d$, что доказывает включение $B \subseteq A$ и всю лемму. □

Для решения вопроса о существовании оценки $\hat{x}_{2,\lambda}$ нам понадобится

Лемма 4.6. Вероятность того, что случайная величина β примет нулевое значение, стремится к нулю, т.е.

$$P\left(\beta = 0\right) \xrightarrow[n\to\infty]{} 0.$$

Доказательство. Выберем индекс j так, чтобы x_λ лежал внутри отрезка $\left[j/n; (j+1)/n\right]$, где $F(x_\lambda) = \lambda$. Тогда

$$P(A) \leq P\left(K_d\left(\frac{F_{nh_r}(j/n) - \lambda}{h_d}\right) = 0\right) = P\left(\left|F_{nh_r}(j/n) - \lambda\right| \geq h_d\right) \leq$$
$$\leq P\left(\left|F_{nh_r}(x_\lambda) - F(x_\lambda)\right| \geq h_d/2 \cup \left|F_{nh_r}(j/n) - F_{nh_r}(x_\lambda)\right| \geq h_d/2\right) \leq P_1 + P_2,$$

где

$$P_1 = P\left(\left|F_{nh_r}(x_\lambda) - F(x_\lambda)\right| \geq h_d/2\right),$$

$$P_2 = P\left(\left|F_{nh_r}(j/n) - F_{nh_r}(x_\lambda)\right| \geq h_d/2\right).$$

Вероятность $P_1 \xrightarrow[n\to\infty]{} 0$, поскольку $\frac{F_{nh_r}(x_\lambda)-F(x_\lambda)}{h_d} \xrightarrow{p} 0$.

Рассмотрим вероятность P_2.

Поскольку функция $F_{nh_r}(x)$ непрерывно дифференцируема, то

$$F_{nh_r}(j/n) - F_{nh_r}(x_\lambda) = F'_{nh_r}(\xi)\left(j/n - x_\lambda\right) = F'_{nh_r} O\left(\frac{1}{n}\right).$$

Докажем, что $F'_{nh_r}(x) = O_p(1)$, откуда будет следовать сходимость по вероятности к нулю величины $F_{nh_r}(j/n) - F_{nh_r}(x_\lambda)$, и сходимость к нулю вероятности P_2.

$$E\left[F'_{nh_r}(x)\right] = \frac{1}{nh_r^2}\sum_{l=1}^{n} K'_r\left(\frac{x-u_l}{h_r}\right)F(u_l) = \frac{1}{h_r^2}\int\limits_0^1 K'_r\left(\frac{x-y}{h_r}\right)F(y)dy +$$

$$+ O\left(\frac{1}{nh_r^2}\right) = \frac{1}{h_r}\int\limits_0^1 K_r\left(\frac{x-y}{h_r}\right)F'(y)dy + O\left(\frac{1}{nh_r^2}\right) =$$

$$= \int\limits_0^1 K_r(z)F'(x-zh_r)dz + O\left(\frac{1}{nh_r^2}\right).$$

$$D\left[F'_{nh_r}(x)\right] = \frac{1}{nh_r^4}\int\limits_0^1 \left(K'_r\left(\frac{x-y}{h_r}\right)\right)^2 F(y)(1-F(y))dy +$$

$$+ O\left(\frac{1}{n^2h_r^4}\right) = \frac{1}{nh_r^3}\int\limits_0^1 \left(K'_r(z)\right)^2 F(x-zh_r)(1-F(x-zh_r))dz + O\left(\frac{1}{n^2h_r^4}\right).$$

В силу закона больших в форме Чебышева заключаем, что $F'_{nh_r}(x) = O_p(1)$. $\square$

Теперь перейдем к вопросу о том, какой случайной величиной можно заменить β, чтобы не нарушилось (асимптотическое) распределение статистики $\hat{x}_{2,\lambda}$. Ответ на него дает следующая

Лемма 4.7. Пусть дана случайная величина ξ такая, что $P(\xi \in B) \xrightarrow[n\to\infty]{} 0$, где множество B из борелевской $\sigma-$алгебры на прямой; обозначим

$$\tilde{\xi} = \begin{cases} \xi & ,\xi \notin B \\ c & ,\xi \in B, \end{cases}$$

где константа c произвольна.

Тогда для любой неслучайной последовательности $\gamma_n \xrightarrow[n\to\infty]{} \infty$

$$\left(\xi - \tilde{\xi}\right) = o_p\left(1/\gamma_n\right).$$

Замечание 4.3. В нашем случае роль ξ играет β, в роли множества B можно взять множество $\{0\}$, состоящее из одного нуля, а в качестве c любую ненулевую константу, например, единицу.

Доказательство. Пусть $\gamma_n \xrightarrow[n\to\infty]{} \infty$ и $\varepsilon > 0$. Тогда

$$P\left(\left|\xi - \tilde{\xi}\right| \gamma_n \geq \varepsilon\right) = P\left(0 \geq \varepsilon\right) + P\left(\left|\xi - c\right| \geq \varepsilon, \tilde{\xi} = c\right) =$$
$$= P\left(\left|\xi - c\right| \geq \varepsilon, \tilde{\xi} = c\right) \leq P\left(\tilde{\xi} = c\right) \xrightarrow[n\to\infty]{} 0.$$

□

Итак, исходная оценка имеет вид

$$\hat{x}_{2,\lambda} = \frac{\alpha_1}{\tilde{\alpha}_2}, \tag{4.2.3}$$

где

$$\alpha_1 = (nh_d)^{-1} \sum_{i=1}^{n} K_d\left(\frac{F_{nh_r}(i/n) - \lambda}{h_d}\right)\frac{i}{n} = S_1 + \Delta_1 \tag{4.2.4}$$

$$\tilde{\alpha}_2 = (nh_d)^{-1} \sum_{i=1}^{n} K_d\left(\frac{F_{nh_r}(i/n) - \lambda}{h_d}\right) = S_2 + \Delta_2, \tag{4.2.5}$$

$S_1, \Delta_1, S_2, \Delta_2$ соответственно из (**??**) — (**??**).

Далее под $\hat{x}_{2,\lambda}$ будем иметь ввиду статистику

$$\hat{x}_\lambda = \frac{\alpha_1}{\alpha_2}, \tag{4.2.6}$$

где α_1 из (**??**) и

$$\alpha_2 = \begin{cases} \tilde{\alpha}_2 & ,\tilde{\alpha}_2 \neq 0 \\ 1 & ,\tilde{\alpha}_2 = 0. \end{cases} = \tilde{\alpha}_2 + o_p\left(1/\gamma_n\right), \tag{4.2.7}$$

для любой неслучайной последовательности $\gamma_n \to \infty$; $\tilde{\alpha}_2$ из (**??**).

4.2.2. Предельное распределение оценки $\hat{x}_{2,\lambda}$

Для изучения асимптотики величины $\hat{x}_{2,\lambda}$ разложим ее:

$$\hat{x}_{2,\lambda} = \frac{\sum_{i=1}^{n} K_d\left(\frac{F_{nh_r}(i/n) - \lambda}{h_d}\right)\frac{i}{n}}{\sum_{i=1}^{n} K_d\left(\frac{F_{nh_r}(i/n) - \lambda}{h_d}\right)} = \frac{S_1 + \Delta_1}{S_2 + \Delta_2},$$

где

$$S_1 = (nh_d)^{-1} \sum_{i=1}^{n} \left[K_d\left(\frac{F(i/n) - \lambda}{h_d}\right)\frac{i}{n}\right], \tag{4.2.8}$$

$$\Delta_1 = (nh_d)^{-1} \sum_{i=1}^{n} \left[K_d\left(\frac{F_{nh_r}(i/n) - \lambda}{h_d}\right)\frac{i}{n} - K_d\left(\frac{F(i/n) - \lambda}{h_d}\right)\frac{i}{n}\right], \tag{4.2.9}$$

$$S_2 = (nh_d)^{-1} \sum_{i=1}^{n} K_d\left(\frac{F(i/n) - \lambda}{h_d}\right), \tag{4.2.10}$$

$$\Delta_2 = (nh_d)^{-1} \sum_{i=1}^{n} \left[K_d\left(\frac{F_{nh_r}(i/n) - \lambda}{h_d}\right) - K_d\left(\frac{F(i/n) - \lambda}{h_d}\right)\right]. \tag{4.2.11}$$

Лемма 4.8.
$$S_1 = F^{-1}(\lambda)\left(F^{-1}\right)'(\lambda) + a_{1,d}h_d^2 + o\left(h_d^2\right),$$
где
$$a_{1,d} = \frac{1}{2}\left[F^{-1}(\lambda)\left(F^{-1}\right)'''(\lambda) + 3\left(F^{-1}\right)'(\lambda)\left(F^{-1}\right)''(\lambda)\right]\nu_d^2. \tag{4.2.12}$$

Доказательство. Вариация функции
$$\tilde{f}(x) = \frac{1}{h_d}K_d\left(\frac{F(x)-\lambda}{h_d}\right)x$$
на отрезке $[0;1]$, аналогично случаю с оценкой $\hat{x}_{1,\lambda}$, имеет следующую асимптотику:
$$V\left[\tilde{f}\right] = \sup\sum_{j=1}^{l}\left|\tilde{f}(x_j) - \tilde{f}(x_{j-1})\right| = O\left(\frac{1}{h_d}\right).$$

Следовательно, можем применить неравенство Koksma-Hlawka:
$$S_1 = \int_0^1 \tilde{f}(x)dx + O\left(\frac{1}{nh_d}\right) = \frac{1}{h_d}\int_0^1 K_d\left(\frac{F(x)-\lambda}{h_d}\right)xdx + O\left(\frac{1}{nh_d}\right).$$
Поскольку $F'(x) > 0$, то замена $z = \big(F(x)-\lambda\big)/h_d$ приводит интеграл к следующему виду
$$S_1 = \int\limits_{\left(F(0)-\lambda\right)/h_d}^{\left(F(1)-\lambda\right)/h_d} K_d(z)F^{-1}(zh_d+\lambda)\left(F^{-1}\right)'(zh_d+\lambda)dz + O\left(\frac{1}{nh_d}\right).$$
Так как $\lambda \in (0;1)$, $F(0)=0$, $F(1)=1$ и $K_d(z)=0$ для z вне отрезка $[-1;1]$, то отрезок интегрирования можно заменить на $[-1;1]$. Поскольку $F'(x) > 0$ и $F(x) \in C^3[0;1]$, то $F^{-1}(x) \in C^3[0;1]$ и, следовательно, верно разложение Тейлора:
$$S_1 = \int_{-1}^{1} K_d(z)\left(F^{-1}(\lambda) + \left(F^{-1}\right)'(\lambda)zh_d + \left(F^{-1}\right)''(\lambda)\frac{(zh_d)^2}{2} + o(h_d^2)\right)\times$$
$$\times\left(\left(F^{-1}\right)'(\lambda) + \left(F^{-1}\right)''(\lambda)zh_d + \left(F^{-1}\right)'''(\lambda)\frac{(zh_d)^2}{2} + o(h_d^2)\right)dz + O\left(\frac{1}{nh_d}\right).$$
Учтем, что $\int K_d(z)dz = 1$ и $\int K_d(z)zdz = 0$:
$$S_1 = F^{-1}(\lambda)\left(F^{-1}\right)'(\lambda) + \left\{\frac{1}{2}F^{-1}(\lambda)\left(F^{-1}\right)'''(\lambda) + \right.$$
$$\left. + \left(F^{-1}\right)''(\lambda)\left(F^{-1}\right)'(\lambda)\frac{3}{2}\right\}h_d^2\nu_d^2 + o(h_d^2) + O\left(\frac{1}{nh_d}\right).$$
□

Аналогично проведем рассуждения для S_2.

Лемма 4.9. Величина S_2, определенная в (**??**), представима в следующем виде:
$$S_2 = \left(F^{-1}\right)'(\lambda) + \left(F^{-1}\right)'''(\lambda)\frac{\nu_d^2}{2}h_d^2 + o(h_d^2).$$

Доказательство. Имеем

$$S_2 = \frac{1}{nh_d}\sum_{i=1}^{n} K_d\left(\frac{F(i/n)-\lambda}{h_d}\right) = \frac{1}{h_d}\int\limits_0^1 K_d\left(\frac{F(x)-\lambda}{h_d}\right)dx + O\left(\frac{1}{nh_d}\right) =$$

$$= \int\limits_{-1}^{1} K_d(z)\left(F^{-1}\right)'(zh_d+\lambda)dz + O\left(\frac{1}{nh_d}\right) = \int\limits_{-1}^{1} K_d(z)\Big[\left(F^{-1}\right)'(\lambda) +$$

$$+\left(F^{-1}\right)''(\lambda)(zh_d) + \left(F^{-1}\right)'''(\lambda)\frac{(zh_d)^2}{2} + o\left((zh_d)^2\right)\Big]dz + O\left(\frac{1}{nh_d}\right) =$$

$$= \left(F^{-1}\right)'(\lambda) + \left(F^{-1}\right)'''(\lambda)\frac{\nu_d^2}{2}h_d^2 + o(h_d^2) + O\left(\frac{1}{nh_d}\right),$$

то есть

$$S_2 = \left(F^{-1}\right)'(\lambda) + \left(F^{-1}\right)'''(\lambda)\frac{\nu_d^2}{2}h_d^2 + o(h_d^2) + O\left(\frac{1}{nh_d}\right).$$

□

Далее изучим поведение величины Δ_1.

В силу того, что $K_d(x)\in C^2[-1;1]$, то для Δ_1 справедливо разложение Тейлора, аналогичное случаю с оценкой $\hat{x}_{1,\lambda}$: $\Delta_1 = \Delta_{1,1}+\Delta_{1,2}+\Delta_{1,3}+\Delta_{1,4}$, где

$$\Delta_{1,1} = \frac{1}{nh_d^2}\sum_{i=1}^{n}\left[K_d'\left(\frac{F(i/n)-\lambda}{h_d}\right)\Big(F_{nh_r}(i/n) - F(i/n)\Big)\frac{i}{n}\right],$$

$$\Delta_{1,2} = \frac{1}{nh_d^3}\sum_{i=1}^{n}\left[K_d''\left(\frac{F(i/n)-\lambda}{h_d}\right)\Big(F_{nh_r}(i/n) - F(i/n)\Big)^2\frac{i}{n}\right],$$

$$\Delta_{1,3} = \frac{1}{nh_d^4}\sum_{i=1}^{n}\left[K_d'''\left(\frac{F(i/n)-\lambda}{h_d}\right)\Big(F_{nh_r}(i/n) - F(i/n)\Big)^3\frac{i}{n}\right],$$

$$\Delta_{1,4} = \frac{1}{nh_d^5}\sum_{i=1}^{n}\left[K_d^{(4)}\left(\frac{\xi_i-\lambda}{h_d}\right)\Big(F_{nh_r}(i/n) - F(i/n)\Big)^4\frac{i}{n}\right],$$

и

$$\Big|\xi_i - F(i/n)\Big| \le \Big|F_{nh_r}(i/n) - F(i/n)\Big| \text{ для } i=\overline{1;n}.$$

Имеет место следующая

Лемма 4.10.

$$\left(\Delta_{1,1} - a_{1,r}h_r^2\right)\sqrt{nh_r^3} \xrightarrow{d} N\left(0;g_1^2\right),$$

где

$$a_{1,r} = \tilde{a}_{1,r}\frac{\nu_r^2}{2}\int\limits_{-1}^{1} K_d'(z)zdz,$$

$$g_1^2 = \lambda(1-\lambda)\left(F^{-1}\right)^2(\lambda)\left(\left(F^{-1}\right)'(\lambda)\right)^4\int\limits_{-1}^{1}\left(K_r'(t)\right)^2dt\left\{\int\limits_{-1}^{1}K_d'(z)zdz\right\}^2$$

и

$$\tilde{a}_{1,r} = F''\left(F^{-1}(\lambda)\right) F^{-1}(\lambda) \left(F^{-1}\right)''(\lambda) + F''\left(F^{-1}(\lambda)\right) \left(\left(F^{-1}\right)'(\lambda)\right)^2 + \\ + F'''\left(F^{-1}(\lambda)\right) \left(\left(F^{-1}\right)'(\lambda)\right)^2 F^{-1}(\lambda).$$

Доказательство. Используя лемму **??**, преобразуем величину $E\left[\Delta_{1,1}\right]$:

$$E\left[\Delta_{1,1}\right] = \frac{1}{nh_d^2}\sum_{i=1}^{n}\left[K_d'\left(\frac{F\left(i/n\right)-\lambda}{h_d}\right) E\left[F_{nh_r}\left(i/n\right) - F\left(i/n\right)\right]\frac{i}{n}\right] = \\ = \frac{1}{nh_d^2}\sum_{i=1}^{n}\left[K_d'\left(\frac{F\left(i/n\right)-\lambda}{h_d}\right)\left(h_r^2\frac{\nu_r^2}{2}F''(i/n) + O\left(\frac{1}{nh_r}\right)\right)\frac{i}{n}\right] = \\ = \frac{\nu_r^2}{2}\frac{h_r^2}{nh_d^2}\sum_{i=1}^{n}\left[K_d'\left(\frac{F\left(i/n\right)-\lambda}{h_d}\right)F''(i/n)\frac{i}{n}\right] + \\ + O\left(\frac{1}{nh_r}\right)\frac{1}{nh_d^2}\sum_{i=1}^{n}\left[K_d'\left(\frac{F\left(i/n\right)-\lambda}{h_d}\right)\frac{i}{n}\right] =$$

поскольку вариации $V\left[K_d'\right]$ и $V\left[F''(x)\right]$ ограничены, продолжим

$$= \frac{\nu_r^2}{2}\frac{h_r^2}{h_d^2}\int_0^1 K_d'\left(\frac{F\left(x\right)-\lambda}{h_d}\right)F''(x)xdx + \\ + O\left(\frac{1}{nh_rh_d}\right)\frac{1}{h_d}\int_0^1 K_d'\left(\frac{F\left(x\right)-\lambda}{h_d}\right)xdx + O\left(\frac{h_r^2}{nh_d^2}\right) + O\left(\frac{1}{n^2h_rh_d^2}\right) = \\ = \frac{\nu_r^2}{2}\frac{h_r^2}{h_d^2}\int_{-1}^1 K_d'\left(z\right)F''(x)x(z)x_z'(z)dz + O\left(\frac{1}{nh_rh_d}\right)$$

Поскольку $K_d'(z)$ нечетная функция, то под интегралом множитель при $K_d'(z)$ должен содержать нечетную степень z, иначе интеграл занулится. Учитывая этот факт, выделим главную часть множителя $F''(x)x(z)x_z'(z)$.

$$x(z) = F^{-1}(zh_d + \lambda) = F^{-1}(\lambda) + \left(F^{-1}\right)'(\lambda)(zh_d) + o(h_d),$$
$$x_z'(z) = \left(F^{-1}\right)'(zh_d + \lambda)h_d = \left(\left(F^{-1}\right)'(\lambda) + \left(F^{-1}\right)''(\lambda)(zh_d) + o(h_d)\right)h_d,$$

$$F''(x) = F''\left(F^{-1}(zh_d+\lambda)\right) = \\ = F''\left(F^{-1}(\lambda) + \left(F^{-1}\right)'(\lambda)(zh_d) + o(h_d)\right) = \\ = F''\left(F^{-1}(\lambda)\right) + F'''\left(F^{-1}(\lambda)\right)\left(F^{-1}\right)'(\lambda)zh_d + o(h_d),$$

и, в итоге,

$$F''(x)x(z)x_z'(z) = \left(a_0 + \tilde{a}_{1,r}zh_d + o(h_d)\right)h_d,$$

Тогда

$$E\left[\Delta_{1,1}\right] = \tilde{a}_{1,r}\frac{\nu_r^2}{2}h_r^2\int_{-1}^1 K_d'(z)zdz + O\left(\frac{1}{nh_dh_r}\right).$$

Проведем выкладки для дисперсии величины $\Delta_{1,1}$.

$$Var\left[\Delta_{1,1}\right] =$$

$$= \frac{1}{n^4 h_d^4 h_r^2} Var\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\left[K_d'\left(\frac{F\left(i/n\right)-\lambda}{h_d}\right)K_r\left(\frac{i/n-u_j}{h_r}\right)I\left(X_j<u_j\right)\frac{i}{n}\right]\right] =$$

поскольку случайные величины X_j независимы и $Var\left[I\left(X_j<u_j\right)\right] = F(u_j)(1-F(u_j))$, продолжим

$$= \frac{1}{n^4 h_d^4 h_r^2}\sum_{j=1}^{n} Var\left[I\left(X_j<u_j\right)\sum_{i=1}^{n}\left[K_d'\left(\frac{F\left(i/n\right)-\lambda}{h_d}\right)K_r\left(\frac{i/n-u_j}{h_r}\right)\frac{i}{n}\right]\right] =$$

$$= \frac{1}{n^4 h_d^4 h_r^2}\sum_{j=1}^{n} F(u_j)(1-F(u_j))\left(\sum_{i=1}^{n}K_d'\left(\frac{F\left(i/n\right)-\lambda}{h_d}\right)K_r\left(\frac{i/n-u_j}{h_r}\right)\frac{i}{n}\right)^2 =$$

т.к. вариации $V\left[K_d'\right]$ и $V\left[K_r\right]$ ограничены, то применим неравенство Koksma-Hlawka

$$= \frac{1}{n^2 h_d^4 h_r^2}\sum_{j=1}^{n} F(u_j)(1-F(u_j))\left(\int_0^1 K_d'\left(\frac{F\left(x\right)-\lambda}{h_d}\right)K_r\left(\frac{x-u_j}{h_r}\right)xdx +\right.$$

$$\left. + O\left(\frac{1}{n}\right)\right)^2 = \frac{1}{n^2 h_d^4 h_r^2}\sum_{j=1}^{n} F(u_j)(1-F(u_j))\left(\int_{-1}^{1} K_d'\left(z\right)K_r\left(\frac{x-u_j}{h_r}\right)xx_z'dz +\right.$$

$$\left. + O\left(\frac{1}{n}\right)\right)^2 =$$

из того, что вариации $V\left[K_d'\right], V\left[K_r\right]$ ограничены следует ограниченность вариации $V\left[K_d'\left(z\right)K_r\left(\frac{x-u_j}{h_r}\right.\right.$
и

$$= \frac{1}{nh_d^4 h_r^2}\int_0^1 F(y)(1-F(y))\times$$

$$\times\left(\int_{-1}^{1} K_d'\left(z\right)K_r\left(\frac{x-y}{h_r}\right)xx_z'dz + O\left(\frac{1}{n}\right)\right)^2 dy + O\left(\frac{1}{n^2 h_d^2 h_r^2}\right)$$

Рассмотрим отдельно величину $K_r\left(\frac{x-y}{h_r}\right)$:

$$K_r\left(\frac{x-y}{h_r}\right) = K_r\left(\frac{F^{-1}(zh_d+\lambda)-y}{h_r}\right) =$$

$$= K_r\left(\frac{F^{-1}(\lambda)-y+\left(F^{-1}\right)'(\lambda)zh_d}{h_r} + o\left(\frac{h_d}{h_r}\right)\right).$$

Тогда

$$Var\left[\Delta_{1,1}\right] = $$

$$= \frac{1}{nh_d^4 h_r} \int\limits_{\left(F^{-1}(\lambda)-1\right)/h_r}^{F^{-1}(\lambda)/h_r} F\left(F^{-1}(\lambda) - th_r\right)\left(1 - F\left(F^{-1}(\lambda) - th_r\right)\right) \times$$

$$\times \left[\int\limits_{-1}^{1} K_d'(z) x x_z' K_r \left(t + \frac{\left(F^{-1}\right)'(\lambda) z h_d}{h_r} + o\left(\frac{h_d}{h_r}\right)\right) dz + O\left(\frac{1}{n}\right)\right]^2 dt + O\left(\frac{1}{n^2 h_d^2 h_r^2}\right).$$

Применим разложение Тейлора для функций $x(z)$ и $x'(z)$ и учтем, что $\int K_d'(z)dz = 0$:

$$Var\left[\Delta_{1,1}\right] = \frac{1}{nh_r h_d^2} \int\limits_{-1}^{1} \lambda(1-\lambda) \times$$

$$\times \left\{ \int\limits_{-1}^{1} \left(K_d'(z) F^{-1}(\lambda) \left(F^{-1}\right)'(\lambda) K_r'(t) \left(F^{-1}\right)'(\lambda) \frac{z h_d}{h_r} + o\left(\frac{h_d}{h_r}\right)\right) dz + \right.$$

$$\left. + O\left(\frac{1}{n}\right) \right\}^2 dt + O\left(\frac{1}{n^2 h_d^2 h_r^2}\right) = g_1^2 \frac{1}{nh_r^3} + o\left(\frac{1}{nh_r^3}\right).$$

Дальнейшие рассуждения в точности повторяют рассуждения для оценки $\hat{x}_{1,\lambda}$ и приводят к применению центральной предельной теоремы Ляпунова. $\square$

Также верен аналог леммы (**??**).

Лемма 4.11.

$$\Delta_{1,2} + \Delta_{1,3} + \Delta_{1,4} = o_d\left(\frac{1}{\sqrt{nh_r^3}}\right),$$

где $\Delta_{1,2}$, $\Delta_{1,3}$ и $\Delta_{1,4}$ из (**??**).

Рассуждения для статистики Δ_2 в точности повторяют рассуждения для статистики Δ_1 за исключением некоторых тонкостей. Именно,

$$\Delta_2 = \Delta_{2,1} + \Delta_{2,2} + \Delta_{2,3} + \Delta_{2,4},$$

где

$$\Delta_{2,1} = \frac{1}{nh_d^2} \sum_{i=1}^{n} \left[K_d'\left(\frac{F\left(i/n\right) - \lambda}{h_d}\right) \left(F_{nh_r}\left(i/n\right) - F\left(i/n\right)\right) \right],$$

$$\Delta_{2,2} = \frac{1}{nh_d^3} \sum_{i=1}^{n} \left[K_d''\left(\frac{\xi - \lambda}{h_d}\right) \left(F_{nh_r}\left(i/n\right) - F\left(i/n\right)\right)^2 \frac{1}{2} \right],$$

$$\Delta_{2,3} = \frac{1}{nh_d^4} \sum_{i=1}^{n} \left[K_d'''\left(\frac{F\left(i/n\right) - \lambda}{h_d}\right) \left(F_{nh_r}\left(i/n\right) - F\left(i/n\right)\right)^3 \right],$$

$$\Delta_{2,4} = \frac{1}{nh_d^5} \sum_{i=1}^{n} \left[K_d^{(4)}\left(\frac{F\left(i/n\right) - \lambda}{h_d}\right) \left(F_{nh_r}\left(i/n\right) - F\left(i/n\right)\right)^4 \right]$$

и

$$\left|\xi_i - F\left(i/n\right)\right| \le \left|F_{nh_r}\left(i/n\right) - F\left(i/n\right)\right|, i = \overline{1;n},$$

Лемма 4.12.

$$\left(\Delta_{2,1} - a_{2,r}h_r^2\right)\sqrt{nh_r^3} \xrightarrow{d} N\left(0; g_2^2\right),$$

где

$$a_{2,r} = \frac{\nu_r^2}{2}\left\{F''\left(F^{-1}(\lambda)\right)\left(F^{-1}\right)''(\lambda) + F'''\left(F^{-1}(\lambda)\right)\left(\left(F^{-1}\right)'(\lambda)\right)^2\right\}\int_{-1}^{1} K_d'(z)zdz,$$

$$g_2^2 = \lambda(1-\lambda)\left(\left(F^{-1}\right)'(\lambda)\right)^4\int_{-1}^{1}\left(K_r'(t)\right)^2 dt\left\{\int_{-1}^{1} K_d'(z)zdz\right\}^2$$

Доказательство.

$$E\left[\Delta_{2,1}\right] = \frac{1}{h_d^2}\int_{-1}^{1} K_d'(z)\times$$

$$\times\left[h_r^2F''(x)\frac{\nu_r^2}{2} + h_r^4F^{(4)}(x)\frac{\hat{\nu}_r^2}{24} + O\left(\frac{1}{nh_r}\right)\right]x_z'dz + O\left(\frac{1}{nh_d^2}\right) =$$

$$= \frac{1}{h_d}\int_{-1}^{1} K_d'(z)\left(h_r^2\frac{\nu_r^2}{2}\tilde{a}_1zh_d + h_d o(h_r^2) + o(h_r^2h_d^2)\right)dz + O\left(\frac{1}{nh_d^2}\right) =$$

$$= a_{2,r}h_r^2 + o(h_r^2) + O\left(\frac{1}{nh_d^2}\right),$$

где

$$\tilde{a}_1\frac{\nu_r^2}{2}\int_{-1}^{1} K_d'(z)zdz = a_{2,r}.$$

Дисперсия величины $\Delta_{2,1}$ имеет следующий вид:

$$Var\left[\Delta_{2,1}\right] =$$

$$= \frac{1}{nh_d^4h_r}\int_{-1}^{1} F\left(F^{-1}(\lambda) - th_r\right)\left(1 - F\left(F^{-1}(\lambda) - th_r\right)\right)\times$$

$$\times\left[\int_{-1}^{1} K_d'(z)x_z'K_r\left(t + \frac{F^{-1}(\lambda)zh_d}{h_r} + o\left(\frac{h_d}{h_r}\right)\right)dz + O\left(\frac{1}{n}\right)\right]^2 dt + O\left(\frac{1}{n^2h_d^2h_r^2}\right),$$

и применив снова разложение Тейлора для функции $x'(z)$, получим:

$$Var\left[\Delta_{2,1}\right] = \frac{1}{nh_rh_d^2}\int_{-1}^{1}\lambda(1-\lambda)\times$$

$$\times\left[\int_{-1}^{1}\left(K_d'(z)\left(F^{-1}\right)'(\lambda)K_r'(t)\left(F^{-1}\right)'(\lambda)\frac{zh_d}{h_r}\right)dz\right]^2 dt+$$

$$+O\left(\frac{1}{n^2h_d^2h_r^2}\right)=g_2^2\frac{1}{nh_r^3}+o\left(\frac{1}{nh_r^3}\right),$$

Далее алгоритм повторяет рассуждения для Δ_1. $\square$

Чтобы найти окончательный предельный вид для оценки $\hat{x}_{2,\lambda}$ нам понадобится следующая лемма общего характера.

Лемма 4.13. Пусть $\eta=\frac{1}{1+\xi}$, где ξ - заданная случайная величина, причем $\xi=o_p(1)$ и $\xi\neq-1$. Тогда имеет место следующее разложение

$$\eta=1-\xi+O_p(\xi^2)$$

Доказательство. Введем случайную величину

$$\tilde{\xi}=\begin{cases}\xi & ,|\xi|<1\\ 0 & ,|\xi|\geq 1,\end{cases}$$

По лемме (**??**) $\left(\xi-\tilde{\xi}\right)=o_p\left(1/\gamma_n\right)$ для любой неслучайной последовательности $\gamma_n \xrightarrow[n\to\infty]{} \infty$. Тогда

$$\eta=\frac{1}{1+\xi}-\frac{1}{1+\tilde{\xi}}+\frac{1}{1+\tilde{\xi}}=\frac{\tilde{\xi}-\xi}{(1+\xi)(1+\tilde{\xi})}+\frac{1}{1+\tilde{\xi}}=O_p\left(\tilde{\xi}-\xi\right)+\frac{1}{1+\tilde{\xi}}=$$
$$=\frac{1}{1+\tilde{\xi}}+o_p\left(1/\gamma_n\right)=$$

поскольку $\tilde{\xi}\in(-1;1)$, то можно применить разложение Тейлора

$$=1-\tilde{\xi}+\frac{1}{(1+c)^3}\tilde{\xi}^2+o_p\left(1/\gamma_n\right),$$

где $|c|\leq|\tilde{\xi}|$ и

$$\frac{1}{(1+c)^3}\leq\frac{1}{(1-|\tilde{\xi}|)^3}=\frac{1}{(1+o_p(1))^3}=O_p(1).$$

В итоге получаем

$$\eta=1-\tilde{\xi}+O_p(\tilde{\xi}^2)+o_p(1/\gamma_n)=1-\xi+O_p(\xi^2).$$

$\square$

Обозначим

$$\mu_1=F^{-1}(\lambda)\left(F^{-1}\right)'(\lambda),$$
$$\mu_2=\left(F^{-1}\right)'(\lambda).$$

Из предыдущих результатов следует, что $\alpha_1-\mu_1=o_d(1)$ и $\alpha_2-\mu_2=o_d(1)$.

Преобразуем $\hat{x}_{2,\lambda}$ (α_1 и α_2 из (**??**) и (**??**)):

$$\frac{\alpha_1}{\alpha_2}=\frac{\alpha_1-\mu_1+\mu_1}{\alpha_2-\mu_2+\mu_2}=\frac{1}{\mu_2}\times\frac{\alpha_1-\mu_1}{1+(\alpha_2-\mu_2)/\mu_2}+\frac{\mu_1}{\mu_2}\times\frac{1}{1+(\alpha_2-\mu_2)/\mu_2}.$$

Случайная величина $(\alpha_2-\mu_2)/\mu_2$ удовлетворяет лемме **??**, поэтому

$$\frac{\alpha_1}{\alpha_2}=\frac{\alpha_1-\mu_1}{\mu_2}\left(1-(\alpha_2-\mu_2)/\mu_2+O_p\left((\alpha_2-\mu_2)^2\right)\right)+$$

$$+\frac{\mu_1}{\mu_2}\left(1-(\alpha_2-\mu_2)/\mu_2\right)+O_p\left((\alpha_2-\mu_2)^2\right)=$$

$$=\frac{\alpha_1-\mu_1}{\mu_2}-\frac{\mu_1}{\mu_2^2}(\alpha_2-\mu_2)+\frac{\mu_1}{\mu_2}+O_p\left((\alpha_1-\mu_1)(\alpha_2-\mu_2)^2\right)+O_p\left((\alpha_2-\mu_2)^2\right)=$$

поскольку $\forall\gamma_n \xrightarrow[n\to\infty]{} \infty$ верно $\alpha_2-\tilde{\alpha_2}=o_p(1/\gamma_n)$, то

$$=\frac{\alpha_1-\mu_1}{\mu_2}-\frac{\mu_1}{\mu_2^2}(\tilde{\alpha_2}-\mu_2)+\frac{\mu_1}{\mu_2}+O_p\left((\alpha_1-\mu_1)(\tilde{\alpha_2}-\mu_2)^2\right)+O_p\left((\tilde{\alpha_2}-\mu_2)^2\right)=$$

И для $\hat{x}_{2,\lambda}$ имеем следующее разложение:

$$\hat{x}_{2,\lambda}-\frac{\mu_1}{\mu_2}=\frac{\alpha_1-\mu_1}{\mu_2}-\frac{\mu_1}{\mu_2^2}(\tilde{\alpha_2}-\mu_2)+ \\ +O_p\left((\alpha_1-\mu_1)(\tilde{\alpha_2}-\mu_2)^2\right)+O_p\left((\tilde{\alpha_2}-\mu_2)^2\right). \quad (4.2.13)$$

Заметим, что выражение случайной величины через α_1 и $\tilde{\alpha_2}$ является существенным, поскольку именно распределения этих последних величин изучались нами выше.

Теорема 4.2.

$$\left(\hat{x}_{2,\lambda}-F^{-1}(\lambda)-b(h_r,h_d)\right)\sqrt{nh_r^3}\xrightarrow{d}N\left(0;g^2\right),$$

где

$$b(h_r,h_d)=\frac{1}{\mu_2}b_1(h_r,h_d)-\frac{\mu_1}{\mu_2^2}b_2(h_r,h_d),$$
$$g^2=\frac{g_1^2}{\mu_2^2}+\frac{g_2^2\mu_1^2}{\mu_2^4},$$

а также

$$b_1(h_r,h_d)=\mu_1+a_{1,d}h_d^2+a_{1,r}h_r^2,$$
$$b_2(h_r,h_d)=\mu_2+a_{2,d}h_d^2+a_{2,r}h_r^2$$

и

$$a_{1,d}=\frac{\nu_d^2}{2}\left\{F^{-1}(\lambda)\left(F^{-1}\right)'''(\lambda)+3\left(F^{-1}\right)'(\lambda)\left(F^{-1}\right)''(\lambda)\right\},$$

$$a_{1,r}=\tilde{a}_{1,r}\frac{\nu_r^2}{2}\int_{-1}^{1}K_d'(z)zdz,\qquad a_{2,d}=\left(F^{-1}\right)'''(\lambda)\frac{\nu_d^2}{2},$$

$$\tilde{a}_{1,r}=F''\left(F^{-1}(\lambda)\right)F^{-1}(\lambda)\left(F^{-1}\right)''(\lambda)+F''\left(F^{-1}(\lambda)\right)\left(\left(F^{-1}\right)'(\lambda)\right)^2+ \\ +F'''\left(F^{-1}(\lambda)\right)\left(\left(F^{-1}\right)'(\lambda)\right)^2F^{-1}(\lambda),$$

$$a_{2,r}=\frac{\nu_r^2}{2}\left\{F''\left(F^{-1}(\lambda)\right)\left(F^{-1}\right)''(\lambda)+ \\ +F'''\left(F^{-1}(\lambda)\right)\left(\left(F^{-1}\right)'(\lambda)\right)^2\right\}\int_{-1}^{1}K_d'(z)zdz,$$

$$g_1^2=\lambda(1-\lambda)F^{-1}(\lambda)^2\left(\left(F^{-1}\right)'(\lambda)\right)^4\int_{-1}^{1}\left(K_r'(t)\right)^2dt\left\{\int_{-1}^{1}K_d'(z)zdz\right\}^2$$

$$g_2^2 = \lambda(1-\lambda)\left(\left(F^{-1}\right)'(\lambda)\right)^4 \int\limits_{-1}^{1} \left(K_r'(t)\right)^2 dt \left\{\int\limits_{-1}^{1} K_d'(z) z dz\right\}^2$$

Замечание 4.4. Из этой теоремы следует состоятельность оценки $\hat{x}_{2,\lambda}$.

Доказательство. Напомним, что

$$\alpha_1 = S_1 + \Delta_1,$$
$$\tilde{\alpha}_2 = S_2 + \Delta_2.$$

Воспользуемся разложением (**??**), результатами лемм об асимптотическом распределении величин S_1, S_2, Δ_1 и Δ_2 и проведем простейшие преобразования:

$$\hat{x}_\lambda - \left(F^{-1}\right)(\lambda) - b(h_r, h_d) = \frac{\alpha_1 - b_1(h_r, h_d)}{\mu_2} - \frac{\mu_1}{\mu_2^2}\left(\tilde{\alpha}_2 - b_2(h_r, h_d)\right) + \\ + O_p\left(\left(\alpha_1 - \mu_1\right)\left(\tilde{\alpha}_2 - \mu_2\right)^2\right) + O_p\left(\left(\tilde{\alpha}_2 - \mu_2\right)^2\right) = \\ = \frac{1}{\sqrt{nh_r^3}}\left(\frac{1}{\mu_2} N\left(0; g_1^2\right) - \frac{\mu_1}{\mu_2^2} N\left(0; g_2^2\right)\right) + o_d\left(\frac{1}{\sqrt{nh_r^3}}\right).$$

Далее результат следует из того, что асимптотические случайные величины, стоящие в скобках, независимы. $\square$

4.3. Асимптотика оценки $\hat{x}_{3,\lambda}$

Для изучения асимптотики этой величины представим ее в следующем виде:

$$\hat{x}_{3,\lambda} = \frac{S+\Delta}{\hat{x}_{1,\lambda}},$$

где

$$S = (nh_d)^{-1} \sum_{i=1}^{n} \int\limits_{-\infty}^{\lambda} K_d\left(\frac{F(i/n)-u}{h_d}\right) \frac{i}{n} du, \tag{4.3.1}$$

$$\Delta = (nh_d)^{-1} \sum_{i=1}^{n} \int\limits_{-\infty}^{\lambda} \left\{ K_d\left(\frac{F_{nh_r}(i/n)-u}{h_d}\right) \frac{i}{n} - K_d\left(\frac{F(i/n)-u}{h_d}\right) \frac{i}{n} \right\} du \tag{4.3.2}$$

и величина $\hat{x}_{1,\lambda}$ определена в (**??**) и ее асимптотика известна.

Лемма 4.14.

$$S = \frac{1}{2}\left(F^{-1}(\lambda)\right)^2 + \frac{1}{2} h_d^2 \nu_d^2 \left(\left(F^{-1}\right)''(\lambda) F^{-1}(\lambda) + \left(\left(F^{-1}\right)'(\lambda)\right)^2 \right) + o\left(h_d^2\right)$$

Доказательство. Применим неравенство Koksma-Hlawka:

$$S = \frac{1}{h_d} \int\limits_{-\infty}^{\lambda} du \int\limits_{0}^{1} K_d\left(\frac{F(x)-u}{h_d}\right) x dx + O\left(\frac{1}{nh_d}\right) =$$

$$= \int\limits_0^1 dx \int\limits_{\left(F(x)-\lambda\right)/h_d}^{\infty} K_d(y) x dy + O\left(\frac{1}{nh_d}\right) = \int\limits_0^1 dx \int\limits_{\left(F(x)-\lambda\right)/h_d}^{1} K_d(y) x dy + O\left(\frac{1}{nh_d}\right) =$$

$$= \int\limits_0^{F^{-1}(\lambda-h_d)} dx \int\limits_{-1}^{1} K_d(y) x dy + \int\limits_{F^{-1}(\lambda-h_d)}^{1} dx \int\limits_{\left(F(x)-\lambda\right)/h_d}^{1} K_d(y) x dy + O\left(\frac{1}{nh_d}\right)$$

Первый интеграл берется непосредственно, а во втором сделаем замену $t = \frac{F(x)-\lambda}{h_d}$:

$$S = \frac{1}{2}\left\{F^{-1}(\lambda-h_d)\right\}^2 +$$

$$+ h_d \int\limits_{-1}^{\left(F(1)-\lambda\right)/h_d} dt \int\limits_t^1 K_d(y) \left(F^{-1}\right)'(\lambda+th_d) F^{-1}(\lambda+th_d) dy + O\left(\frac{1}{nh_d}\right) =$$

$$= \frac{1}{2}\left\{ F^{-1}(\lambda) - \left(F^{-1}\right)'(\lambda) h_d + \left(F^{-1}\right)''(\lambda) \frac{h_d^2}{2} + o\left(h_d^2\right) \right\}^2 +$$

$$+ h_d \left(F^{-1}\right)'(\lambda) F^{-1}(\lambda) \int\limits_{-1}^{1} dt \int\limits_t^1 K_d(y) dy +$$

$$+ h_d^2 \int\limits_{-1}^{1} dt \int\limits_t^1 K_d(y) t dy \left\{ \left(F^{-1}\right)''(\lambda) F^{-1}(\lambda) + \left[\left(F^{-1}\right)'(\lambda)\right]^2 \right\} + o\left(h_d^2\right)$$

Так как

$$\int_{-1}^{1} dt \int_{w}^{1} K_d(y)dy = 1,$$

$$\int_{-1}^{1} dt \int_{w}^{1} K_d(y)tdy = \frac{1}{2}\nu_d^2 - \frac{1}{2},$$

то

$$\begin{aligned} S = \frac{1}{2}\Big\{ \left(F^{-1}(\lambda)\right)^2 + \Big(\left(F^{-1}\right)'(\lambda)\Big)^2 h_d^2 - 2F^{-1}(\lambda)\left(F^{-1}\right)'(\lambda)h_d + \\ + F^{-1}(\lambda)\left(F^{-1}\right)''(\lambda)h_d^2\Big\} + \left(F^{-1}\right)'(\lambda)F^{-1}(\lambda)h_d + \\ + h_d^2\left\{\frac{1}{2}\nu_d^2 - \frac{1}{2}\right\}\left\{\left(F^{-1}\right)''(\lambda)F^{-1}(\lambda) + \left[\left(F^{-1}\right)'(\lambda)\right]^2\right\} + o\left(h_d^2\right) = \\ = \frac{1}{2}\left[F^{-1}(\lambda)\right]^2 + \frac{1}{2}h_d^2\nu_d^2\left\{\left(F^{-1}\right)''(\lambda)F^{-1}(\lambda) + \left[\left(F^{-1}\right)'(\lambda)\right]^2\right\} + o\left(h_d^2\right) \end{aligned}$$

□

Представим величину Δ в виде суммы $\Delta = \Delta_1 + \Delta_2 + \Delta_3 + \Delta_4$, где

$$\Delta_1 = \frac{1}{nh_d^2}\sum_{i=1}^{n}\int_{-\infty}^{\lambda} K_d'\left(\frac{F(i/n) - u}{h_d}\right)\frac{i}{n}\Big(F_{nh_r}(i/n) - F(i/n)\Big)du, \tag{4.3.3}$$

$$\Delta_2 = \frac{1}{nh_d^3}\sum_{i=1}^{n}\int_{-\infty}^{\lambda} K_d''\left(\frac{F(i/n) - u}{h_d}\right)\frac{i}{n}\Big(F_{nh_r}(i/n) - F(i/n)\Big)^2 du, \tag{4.3.4}$$

$$\Delta_3 = \frac{1}{nh_d^4}\sum_{i=1}^{n}\int_{-\infty}^{\lambda} K_d'''\left(\frac{F(i/n) - u}{h_d}\right)\frac{i}{n}\Big(F_{nh_r}(i/n) - F(i/n)\Big)^3 du, \tag{4.3.5}$$

$$\Delta_4 = \frac{1}{nh_d^5}\sum_{i=1}^{n}\int_{-\infty}^{\lambda} K_d^{(4)}\left(\frac{\xi_i - u}{h_d}\right)\frac{i}{n}\Big(F_{nh_r}(i/n) - F(i/n)\Big)^4 du, \tag{4.3.6}$$

$\Big|\xi_i - F(i/n)\Big| \le \Big|F(i/n) - F_{nh_r}(i/n)\Big|$.

Изучим асимптотику главной части статистики Δ, величину Δ_1.

Лемма 4.15.

$$\left\{\Delta_1 - a_{1,r}h_r^2\right\}\sqrt{nh_r} \xrightarrow{d} N\left(0; g_1^2\right),$$

где

$$a_{1,r} = -\frac{\nu_r^2}{2}F^{-1}(\lambda)\left(F^{-1}\right)'(\lambda)F''\left(F^{-1}(\lambda)\right),$$

$$g_1^2 = \lambda(1-\lambda)\left\{F^{-1}(\lambda)\left(F^{-1}\right)'(\lambda)\right\}^2\|K_r\|^2.$$

Доказательство. Очевидно, что

$$\Delta_1 = -\frac{1}{nh_d}\sum_{i=1}^{n} K_d\left(\frac{F(i/n) - \lambda}{h_d}\right)\frac{i}{n}\Big(F_{nh_r}(i/n) - F(i/n)\Big).$$

Введем величины

$$\Delta_{1,1} = -\frac{1}{nh_d}\sum_{i=1}^{n} K_d\left(\frac{F\left(i/n\right)-\lambda}{h_d}\right)\frac{i}{n}\left(F_{nh_r}\left(i/n\right) - E\left[F_{nh_r}\left(i/n\right)\right]\right);$$

$$\Delta_{1,2} = -\frac{1}{nh_d}\sum_{i=1}^{n} K_d\left(\frac{F\left(i/n\right)-\lambda}{h_d}\right)\frac{i}{n}\left(E\left[F_{nh_r}\left(i/n\right)\right] - F\left(i/n\right)\right),$$

Тогда $\Delta_1 = \Delta_{1,1} + \Delta_{1,2}$.

Учтем, что $E\left[F_{nh_r}(x) - F(x)\right] = h_r^2 \frac{\nu_r^2}{2} F''(x) + o\left(h_r^2\right)$, и получим

$$E\left[\Delta_1\right] = E\left[\Delta_{1,2}\right] = -\frac{\nu_r^2 h_r^2}{2nh_d}\sum_{i=1}^{n} K_d\left(\frac{F\left(i/n\right)-\lambda}{h_d}\right)\frac{i}{n}F''(x)(1+o(1)) =$$

$$= -\frac{\nu_r^2 h_r^2}{2h_d}\int\limits_0^1 K_d\left(\frac{F(x)-\lambda}{h_d}\right) xF''(x)dx(1+o(1)) + O\left(\frac{h_r^2}{nh_d}\right) =$$

$$= -\frac{\nu_r^2 h_r^2}{2h_d}\int\limits_{-1}^1 K_d(z)xx'_zF''(x)dz(1+o(1)) + O\left(\frac{h_r^2}{nh_d}\right) =$$

$$= -\frac{\nu_r^2}{2}h_r^2\left(F^{-1}\right)(\lambda)\left(F^{-1}\right)'(\lambda)F''\left[\left(F^{-1}\right)(\lambda)\right] + o(h_r^2).$$

Далее вычислим дисперсию Δ_1.

$$Var\left[\Delta_1\right] = Var\left[\Delta_{1,1}\right] =$$

$$= \frac{1}{n^4 h_d^2 h_r^2}\sum_{j=1}^{n} F(u_j)(1-F(u_j))\left\{\sum_{i=1}^{n} K_d\left(\frac{F\left(i/n\right)-\lambda}{h_d}\right)K_r\left(\frac{i/n-u_j}{h_r}\right)\frac{i}{n}\right\}^2 =$$

$$= \frac{1}{n^2 h_d^2 h_r^2}\sum_{j=1}^{n} F(u_j)(1-F(u_j))\left\{\int\limits_0^1 K_d\left(\frac{F\left(x\right)-\lambda}{h_d}\right)K_r\left(\frac{x-u_j}{h_r}\right)xdx + O\left(\frac{1}{n}\right)\right\}^2 =$$

Сделаем замену $z = \frac{F(x)-\lambda}{h_d}$ и применим неравенство К-Н

$$= \frac{1}{nh_d^2 h_r^2}\int\limits_0^1 F(y)(1-F(y))\left\{\int\limits_{-1}^1 K_d(z)K_r\left(\frac{x-y}{h_r}\right)xx'_zdz + O\left(\frac{1}{n}\right)\right\} + O\left(\frac{1}{n^2h_r^2}\right).$$

Выпишем отдельно

$$K_r\left(\frac{x-y}{h_r}\right) = K_r\left(\frac{F^{-1}(\lambda+h_dz)-y}{h_r}\right) = K_r\left(\frac{F^{-1}(\lambda)-y}{h_r}+o(1)\right).$$

Учитывая последние выкладки и делая замену $t = \frac{F^{-1}(\lambda)-y}{h_r}$, запишем окончательно

$$Var\left[\Delta_1\right] = \frac{1}{nh_r}\lambda(1-\lambda)\left\{F^{-1}(\lambda)\left(F^{-1}\right)'(\lambda)\right\}^2\|K_r\|^2 + o\left(\frac{1}{nh_r}\right).$$

□

Остаток величины Δ является бесконечно малой величиной высшего порядка по сравнению с главным членом, Δ_1. Обозначим $D = \Delta_2 + \Delta_3 + \Delta_4$(см. (**??**)-(**??**)).

Лемма 4.16.

$$D = o_d\left(\frac{1}{\sqrt{nh_r}}\right).$$

Доказательство аналогичного утверждения уже приводилось.

Лемма 4.17.

$$\left(S + \Delta - \frac{1}{2}\left(F^{-1}(\lambda)\right)^2 - b_1\left(h_r, h_d, 1\right)\right)\sqrt{nh_r} \xrightarrow{d} N\left(0; g_1^2\right),$$

где

$$b_1\left(h_r, h_d, 1\right) = a_{1,d}h_d^2 + a_{1,r}h_r^2,$$
$$g_1^2 = \lambda(1-\lambda)\|K_r\|^2\left(\left(F^{-1}\right)'(\lambda)F^{-1}(\lambda)\right)^2$$

и

$$a_{1,r} = -\frac{\nu_r^2}{2}F''\left(F^{-1}(\lambda)\right)\left(F^{-1}\right)'(\lambda)F^{-1}(\lambda)$$
$$a_{1,d} = \kappa_2\left(K_d\right)\left\{\left(F^{-1}\right)''(\lambda)F^{-1}(\lambda) + \left(\left(F^{-1}\right)'(\lambda)\right)^2\right\}$$

Далее представим статистику $\hat{x}_{3,\lambda}$ в виде $\frac{\beta}{\alpha}$, где

$$\beta = S_1 + \Delta_1,$$
$$\alpha = (nh_d)^{-1}\sum_{i=1}^{n}\int_{-\infty}^{\lambda} K_d\left(\frac{F_{nh_r}(i/n) - u}{h_d}\right)du.$$

Положив

$$\mu_1 = \frac{1}{2}\left(F^{-1}(\lambda)\right)^2, \tag{4.3.7}$$
$$\mu_2 = F^{-1}(\lambda), \tag{4.3.8}$$

из представления

$$\hat{x}_{3,\lambda} - \frac{\mu_1}{\mu_2} = \frac{\beta - \mu_1}{\mu_2} - \frac{\mu_1}{\mu_2^2}\left(\alpha - \mu_2\right) + $$
$$+ O_p\left(\left(\beta - \mu_1\right)\left(\alpha - \mu_2\right)^2\right) + O_p\left(\left(\alpha - \mu_2\right)^2\right)$$

получим следующий результат:

Теорема 4.3.

$$\left(\hat{x}_{3,\lambda} - \frac{1}{2}F^{-1}(\lambda) - b(h_r, h_d)\right)\sqrt{nh_r} \xrightarrow{d} N\left(0; g^2\right),$$

где

$$b(h_r, h_d) = \frac{1}{\mu_2}b_1(h_r, h_d) - \frac{\mu_1}{\mu_2^2}b_2(h_r, h_d),$$
$$g^2 = \frac{g_1^2}{\mu_2^2} + \frac{g_2^2\mu_1^2}{\mu_2^4},$$

$b_1(h_r, h_d)$, g_1^2 из леммы (**??**), $b_2(h_r, h_d)$, g_2^2 из теоремы (**??**) и μ_1, μ_2 из (**??**), (**??**).

Видно, что величина $\hat{x}_{3,\lambda}$ оценивает половину квантиля. Для непосредственной оценки квантиля умножим исходную статистику на 2, т.е.

$$\hat{x}_{3,\lambda} = 2\frac{\sum_{i=1}^{n}\int_{-\infty}^{\lambda} K_d\left(\frac{F_{nh_r}(i/n)-u}{h_d}\right)\frac{i}{n}du}{\sum_{i=1}^{n}\int_{-\infty}^{\lambda} K_d\left(\frac{F_{nh_r}(i/n)-u}{h_d}\right)du}.$$

и подытожим результат:

Теорема 4.4.

$$\left(\hat{x}_{3,\lambda} - F^{-1}(\lambda) - 2b(h_r,h_d)\right)\sqrt{nh_r} \xrightarrow{d} N\left(0;4g^2\right),$$

где $b(h_r,h_d)$ и g^2 из теоремы (**??**).

Замечание 4.5. Как и прежде, из сформулированной теоремы следует состоятельность оценки $\hat{x}_{3,\lambda}$.

5. Компьютерное моделирование эксперимента

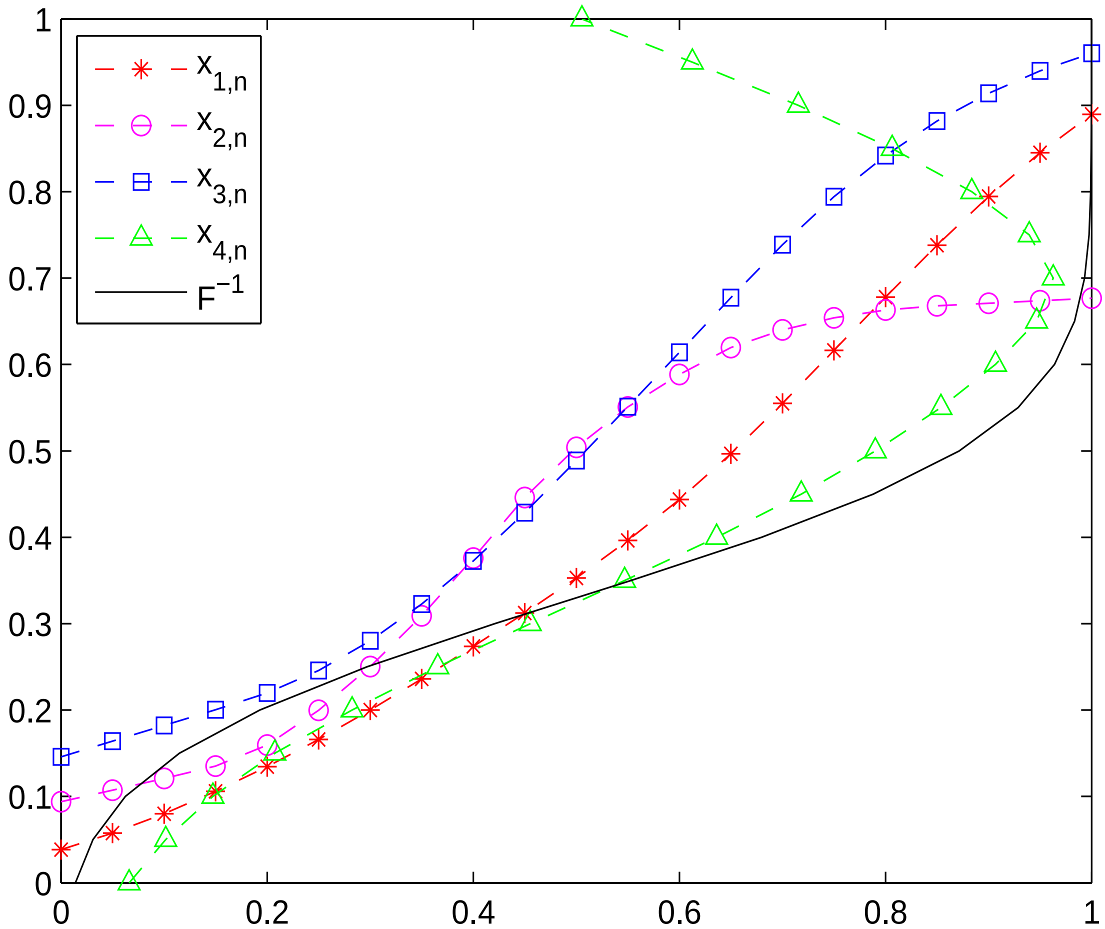
Исследование оценок (**??**) — (**??**) было проведено для следующих модельных данных.

Совокупность точек разбиения отрезка $[0, 1]$ была взята равномерной, т.е. $u_i = \frac{i}{n}$, где n — объем выборки. Распределения гипотетической дозы выбраны из двух классов: нормальное (со средними $a_1 = 0.33$, $a_2 = 0.66$, $a_3 = 0.5$ и стандартными отклонениями $b_1 = 0.15$ и $b_2 = 0.3$) и равномерное на отрезке $[0, 1]$. В обоих случаях величину X можно считать распределенной на отрезке $[0, 1]$ с достаточной степенью точности.

Все расчеты проведены для малого ($n = 50$) и большого ($n = 200$) объемов выборки.

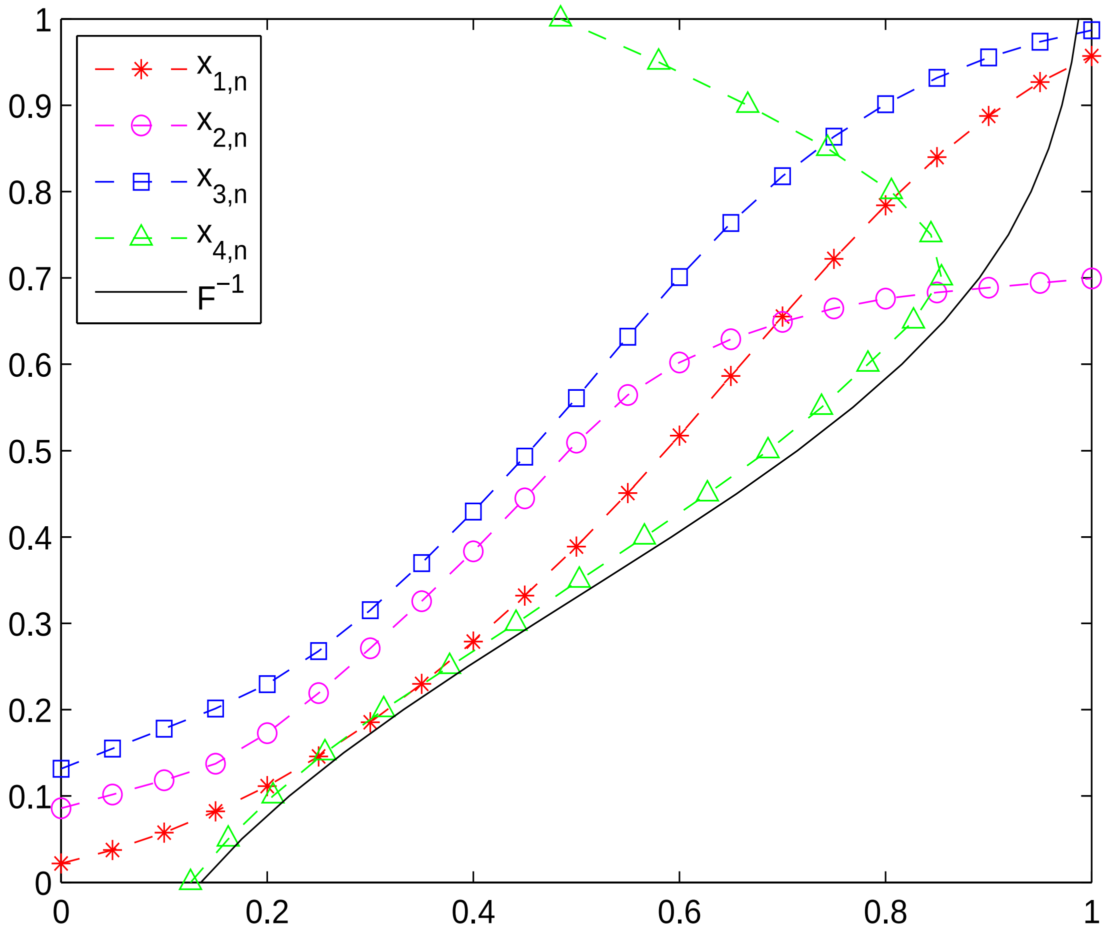
На Рис.**??** представлены результаты моделирования нормально распределенной гипотетической дозы X с математическим ожиданием $a_1 = 0.33$ и стандартным отклонением $b_1 = 0.15$. Выбранное среднее характеризует большую чувствительность биообъекта к препарату, что проявляется в небольших пороговых дозах, при которых наступает реакция организма. Малое значение отклонения характеризует небольшой разброс значений величины X относительно ее среднего. Объем выборки $n = 200$.

Рис. 1. Результаты моделирования нормально распределенной дозы X ($a = 0.33$, $b = 0.15$)



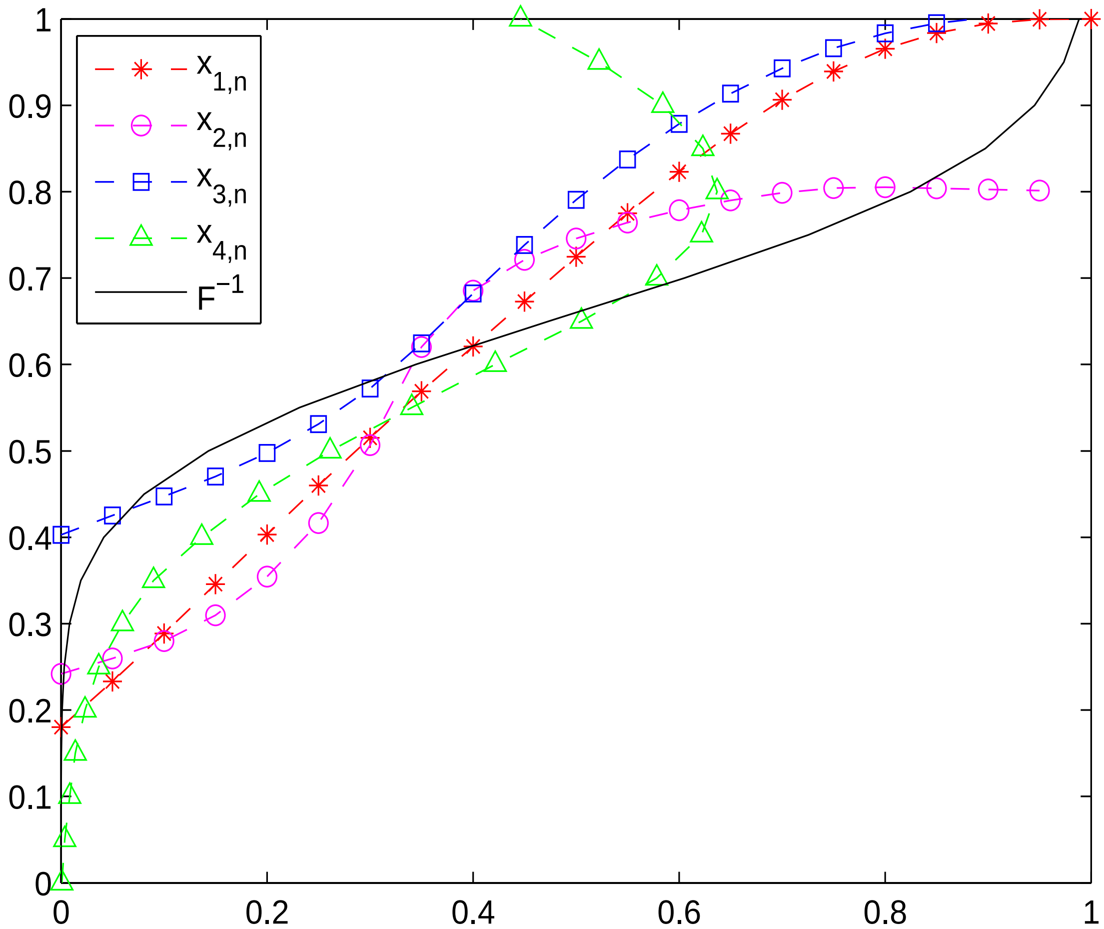
На Рис.?? представлены результаты моделирования нормально распределенной гипотетической дозы X с математическим ожиданием $a_2 = 0.33$ и стандартным отклонением $b_2 = 0.3$. Выбранное среднее характеризует большую чувствительность биообъекта к препарату, что проявляется в небольших пороговых дозах, при которых наступает реакция организма. Большое значение отклонения характеризует относительно большой разброс значений величины X относительно ее среднего. Объем выборки $n = 200$.

Рис. 2. Результаты моделирования нормально распределенной дозы X ($a = 0.33$, $b = 0.3$)



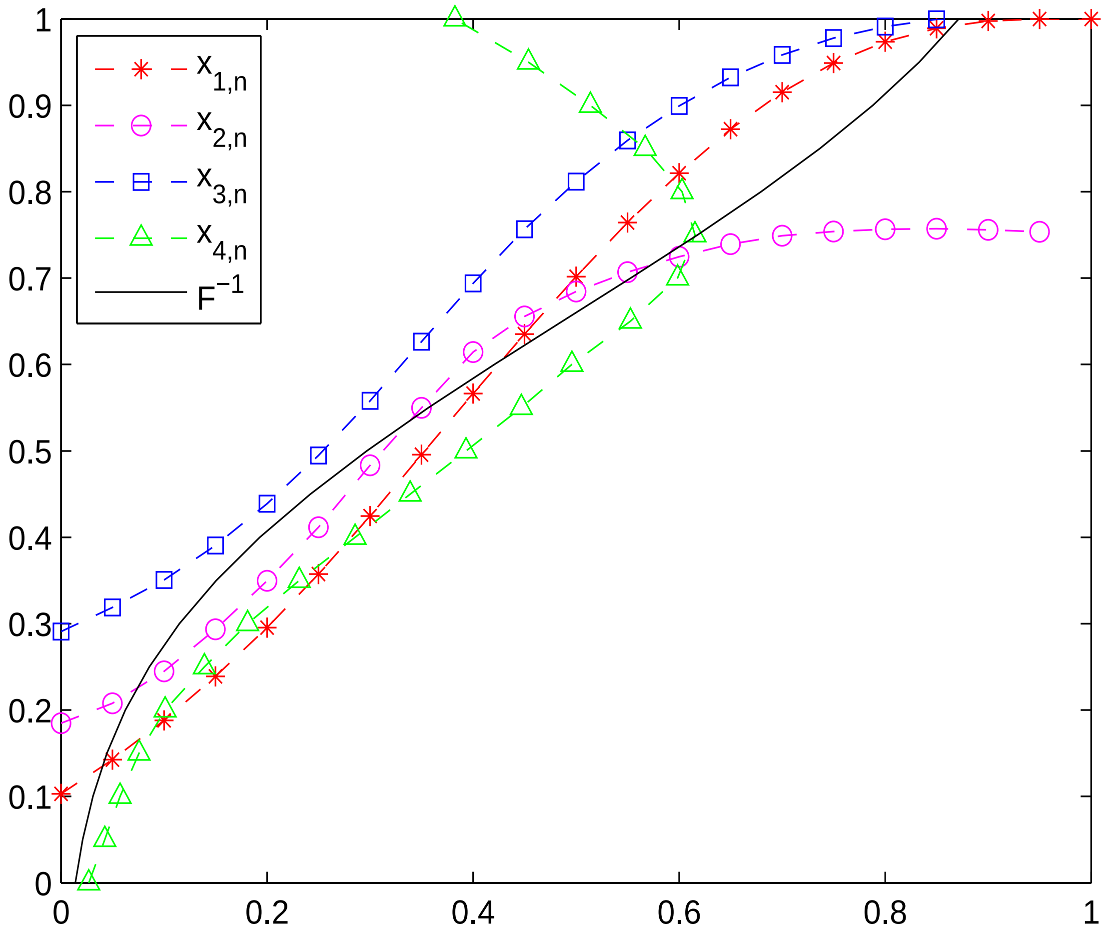
На Рис.?? представлены результаты моделирования нормально распределенной гипотетической дозы X с математическим ожиданием $a_2 = 0.66$ и стандартным отклонением $b_1 = 0.15$. Выбранное среднее характеризует малую чувствительность биообъекта к препарату, что проявляется в относительно больших пороговых дозах, при которых наступает реакция организма. Малое значение отклонения характеризует небольшой разброс значений величины X относительно ее среднего. Объем выборки $n = 200$.

Рис. 3. Результаты моделирования нормально распределенной дозы X ($a = 0.66$, $b = 0.15$)



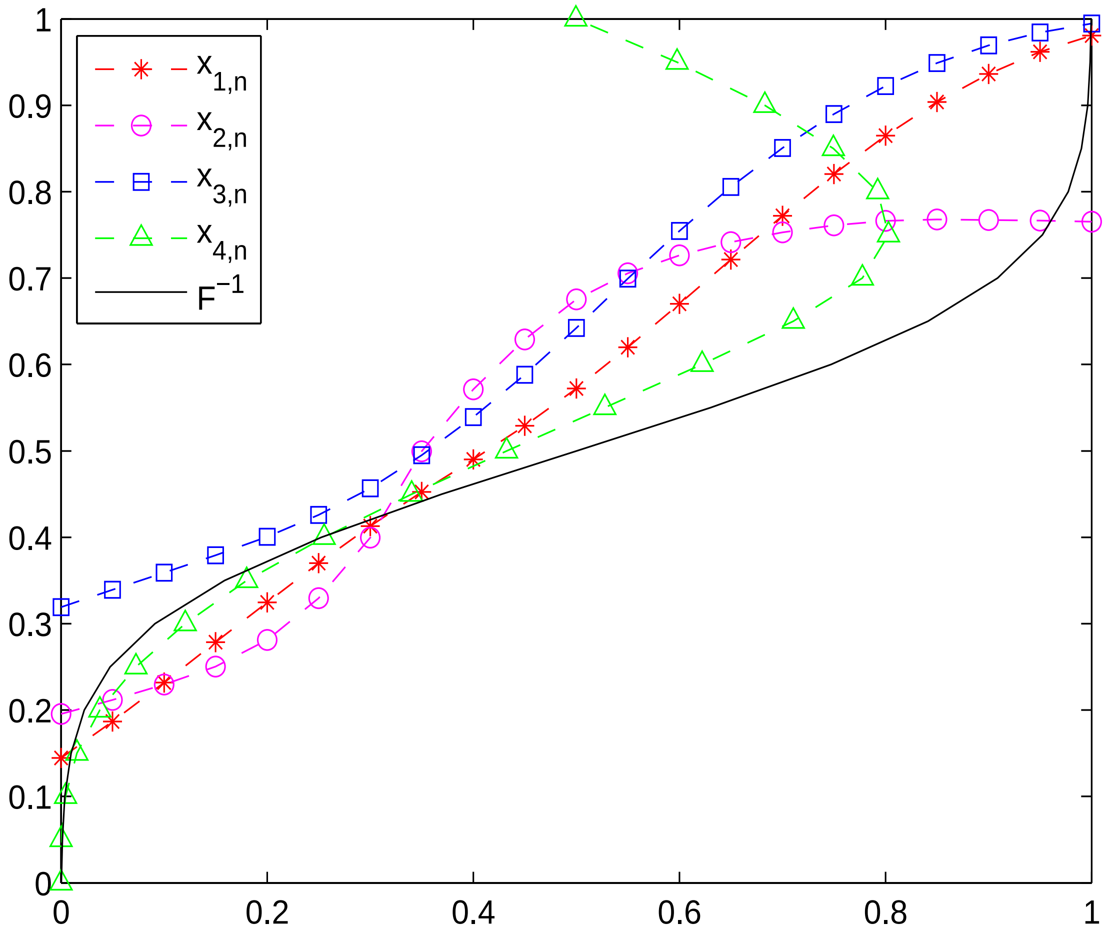
На Рис.?? представлены результаты моделирования нормально распределенной гипотетической дозы X с математическим ожиданием $a_2 = 0.66$ и стандартным отклонением $b_2 = 0.3$. Выбранное среднее характеризует малую чувствительность биообъекта к препарату, что проявляется в относительно больших пороговых дозах, при которых наступает реакция организма. Большое значение отклонения характеризует относительно большой разброс значений величины X относительно ее среднего. Объем выборки $n = 200$.

Рис. 4. Результаты моделирования нормально распределенной дозы X ($a = 0.66$, $b = 0.3$)



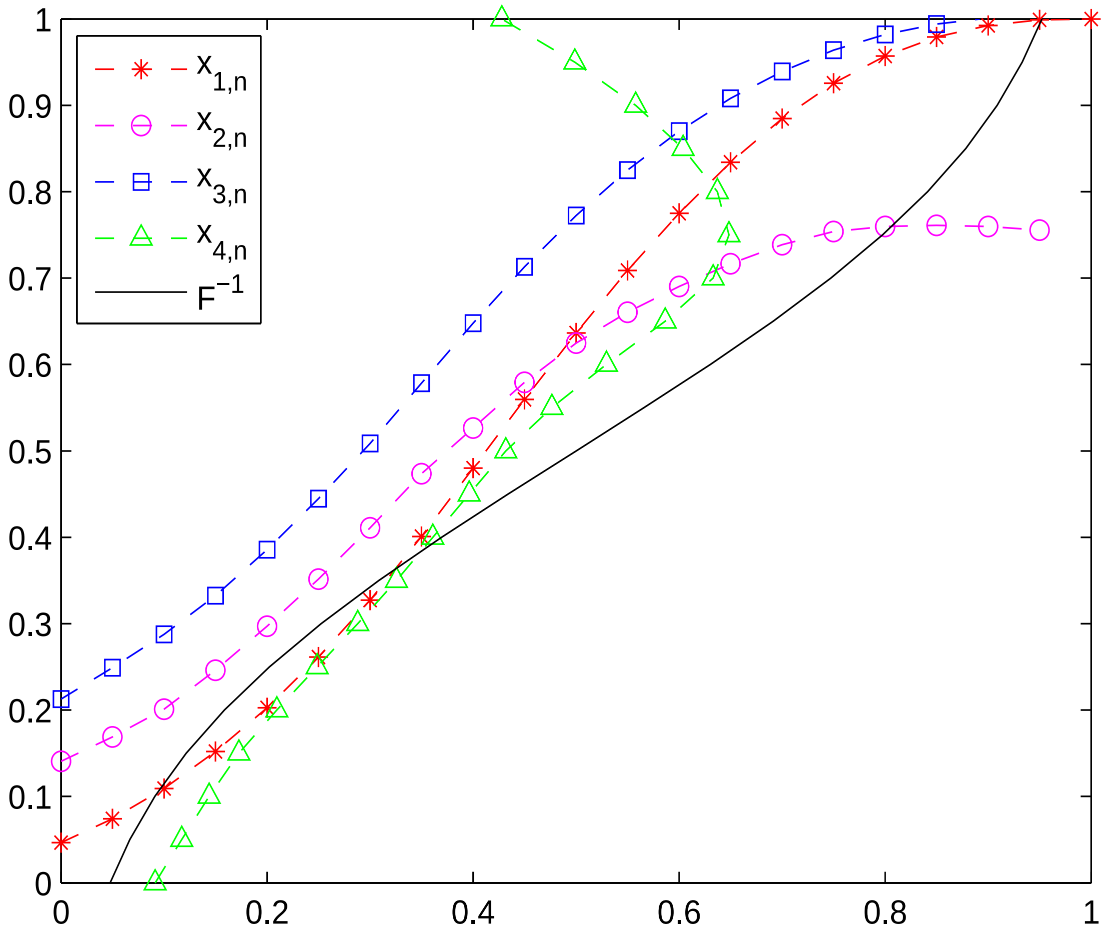
На Рис.**??** представлены результаты моделирования нормально распределенной гипотетической дозы X с математическим ожиданием $a_3 = 0.5$ и стандартным отклонением $b_1 = 0.15$. Выбранное среднее занимает промежуточное положение относительно средних $a_1 = 0.33$ и $a_2 = 0.66$. В данном случае пороговая доза будет принимать значения близкие к 0.5. Разброс значений относительно математического ожидания будет небольшим за счет малости параметра b_1. Объем выборки $n = 200$.

Рис. 5. Результаты моделирования нормально распределенной дозы X ($a = 0.5$, $b = 0.15$)



На Рис.?? представлены результаты моделирования нормально распределенной гипотетической дозы X с математическим ожиданием $a_3 = 0.5$ и стандартным отклонением $b_1 = 0.3$. Выбранное среднее занимает промежуточное положение относительно средних $a_1 = 0.33$ и $a_2 = 0.66$. В данном случае пороговая доза будет принимать значения близкие к 0.5. Разброс значений относительно математического ожидания будет значительным, поскольку стандартное отклонение b_1 велико. Объем выборки $n = 200$.

Рис. 6. Результаты моделирования нормально распределенной дозы X ($a = 0.5$, $b = 0.3$)



Из рисунков **??** — **??** видно, что в окрестности медианы распределения лучшую аппроксимацию дает естественная оценка $\hat{x}_{4,\lambda}$. Однако отрезок ее «эффективного» оценивания небольшой. Для значений пороговых доз близких к границе отрезка $[0,1]$ более удовлетворительную оценку дают величины $\hat{x}_{1,\lambda}$ и $\hat{x}_{3,\lambda}$.

На рисунках также видно положительное смещение оценок $\hat{x}_{1,\lambda}$, $\hat{x}_{2,\lambda}$, $\hat{x}_{3,\lambda}$ при значениях λ больших математического ожидания и отрицательное (в некоторых случаях положительное, но меньшее, чем для больших λ) — при меньших λ. Данное свойство не является следствием «малости» объема выборки и присуще асимптотическому поведению предложенных статистик.

Действительно, смещение оценок относительно оцениваемого квантиля $x_\lambda = F^{-1}(\lambda)$ определяется величинами $a_{1,r}$, $a_{2,r}$, $a_{1,d}$ и $a_{2,d}$, знаки которых положительны при $\lambda > E[X]$. При $\lambda < E[X]$ величины $a_{1,d}$ и $a_{2,d}$ становятся отрицательными, однако знак суммарного смещения остается (асимптотически) положительным за счет малости h_d по сравнению с h_r и положительности $a_{1,r}$ и $a_{2,r}$.

Остальные результаты моделирования вынесены в приложение.

Заключение

В работе были предложены оценки квантилей распределения пороговой дозы X для модели фиксированного плана зависимости «доза-эффект», доказана их состоятельность и асимптотическая нормальность.

Из предельных распределений видно, что статистики $\hat{x}_{1,\lambda}$ и $\hat{x}_{3,\lambda}$ сходятся с большей скоростью $\sqrt{nh_r}$, чем статистика $\hat{x}_{2,\lambda} \sim \sqrt{nh_r^3}$.

Было проведено компьютерное моделирование эксперимента для нормального и равномерного на $[0;1]$ распределений при конечных объемах выборки: $n_1 = 50$ и $n_2 = 200$. Из его результатов можно сделать вывод, что для средних значений квантиля x_λ лучшую аппроксимацию дает естественная оценка $\hat{x}_{4,\lambda}$; значения на границе отрезка $[0,1]$ лучше оценивают статистики $\hat{x}_{1,\lambda}$, $\hat{x}_{3,\lambda}$.

При анализе графиков было также выявлено (уже для конечных выборок $n_1 = 50$ и $n_2 = 200$) теоретическое асимптотическое смещение оценок относительно оцениваемого квантиля $x_\lambda = F^{-1}(\lambda)$, определяемое доказанными в работе теоремами.

Литература

1. Кочеганов В. М., Тихов М. С. Оценивание эффективных доз в зависимости «доза-эффект». — Обозрение Прикладной и Промышленной Математики, т.18, в.1, с.85-86.
2. Тихов М. С., Кочеганов В. М. Проверка гипотез согласия в зависимости «доза-эффект». — Обозрение Прикладной и Промышленной Математики, т.18, в.1, с.187-188.
3. Криштопенко С.В., Тихов М.С., Попова Е.Б. Доза-эффект. – М.: Медицина, 2008. – 288 с.
4. Криштопенко С.В., Тихов М.С., Попова Е.Б. Парадоксальная токсичность. —- Н.Новгород, изд-во НГМА, 2001 – 164 с.
5. Натансон И. П. Теория функций вещественного переменного — М.: Наука, 1974.
6. Колмогоров А.Н. Основные понятия теории вероятностей. М.: Наука, 1974. — 119 с.
7. Колмогоров А. Н., С. В. Фомин Элементы теории функций и функционального анализа — М.: Физматилит,2006.
8. Крамер Г. Математические методы статистики. — М.: Мир, 1975.
9. Гнеденко Б. В. Курс теории вероятностей. — М.: Издательство ЛКИ, 2007.
10. Липцер Р. Ш., Ширяев А. Н. Теория мартингалов. — М.: Наука, 1986.
11. Dette H., Neumeyer N., Pilz K.F. A note on nonparametric estimation of the effective dose in quantal bioassay. — December 19, 2003.
12. Niederreiter H. Random number generation and quasi-Monte Carlo methods — Society for industrial and applied mathematics, Philadelhia, Pennsilvania, 1992.
13. Finney, D.J. Statistical Methods in Biological Assay. Charles Griffin & Co., 1978.

Приложение

Рис. 7. Результаты моделирования для нормально распределенной дозы X, $n = 50$

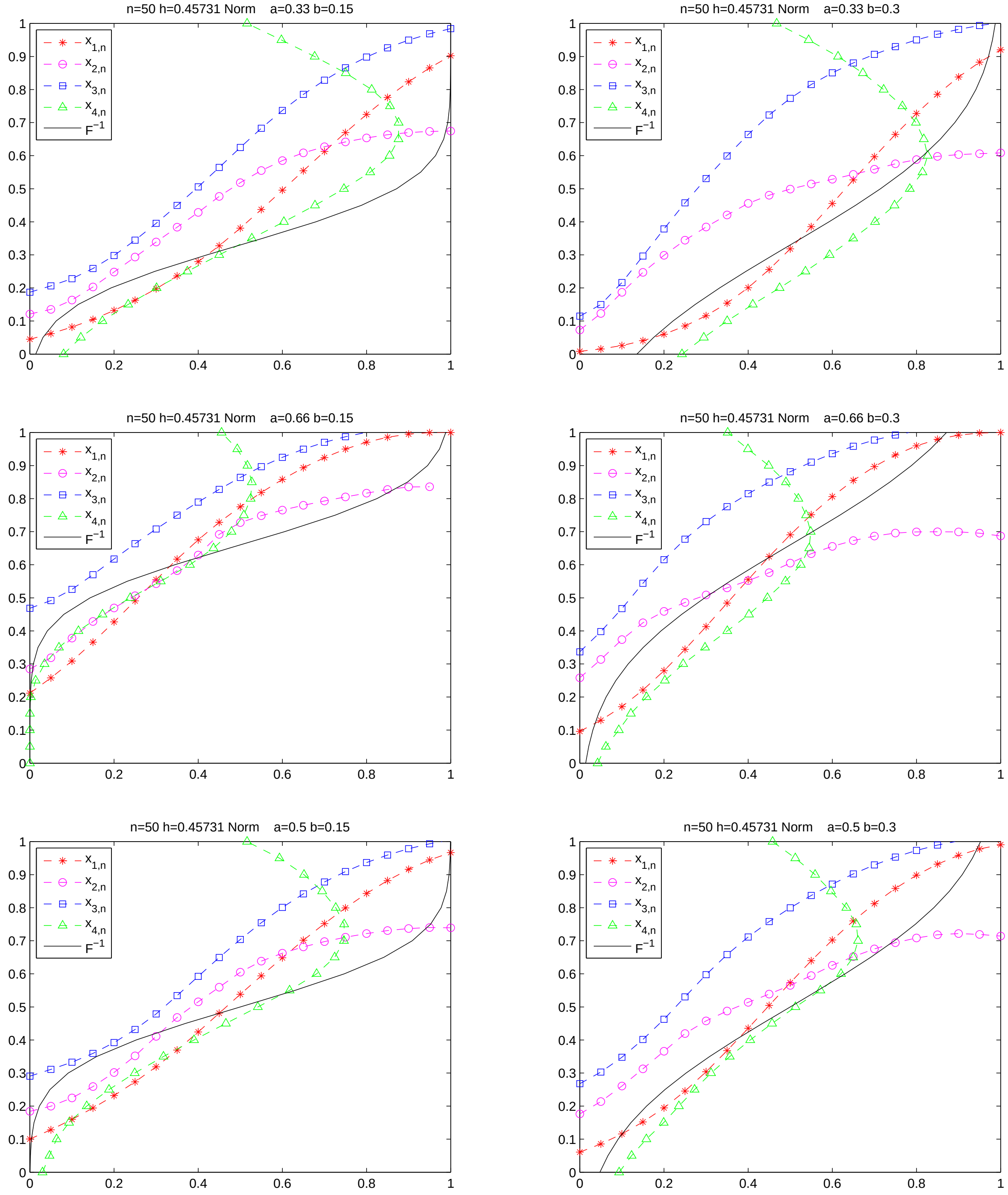


Рис. 8. Результаты моделирования для нормально распределенной дозы X, $n = 200$

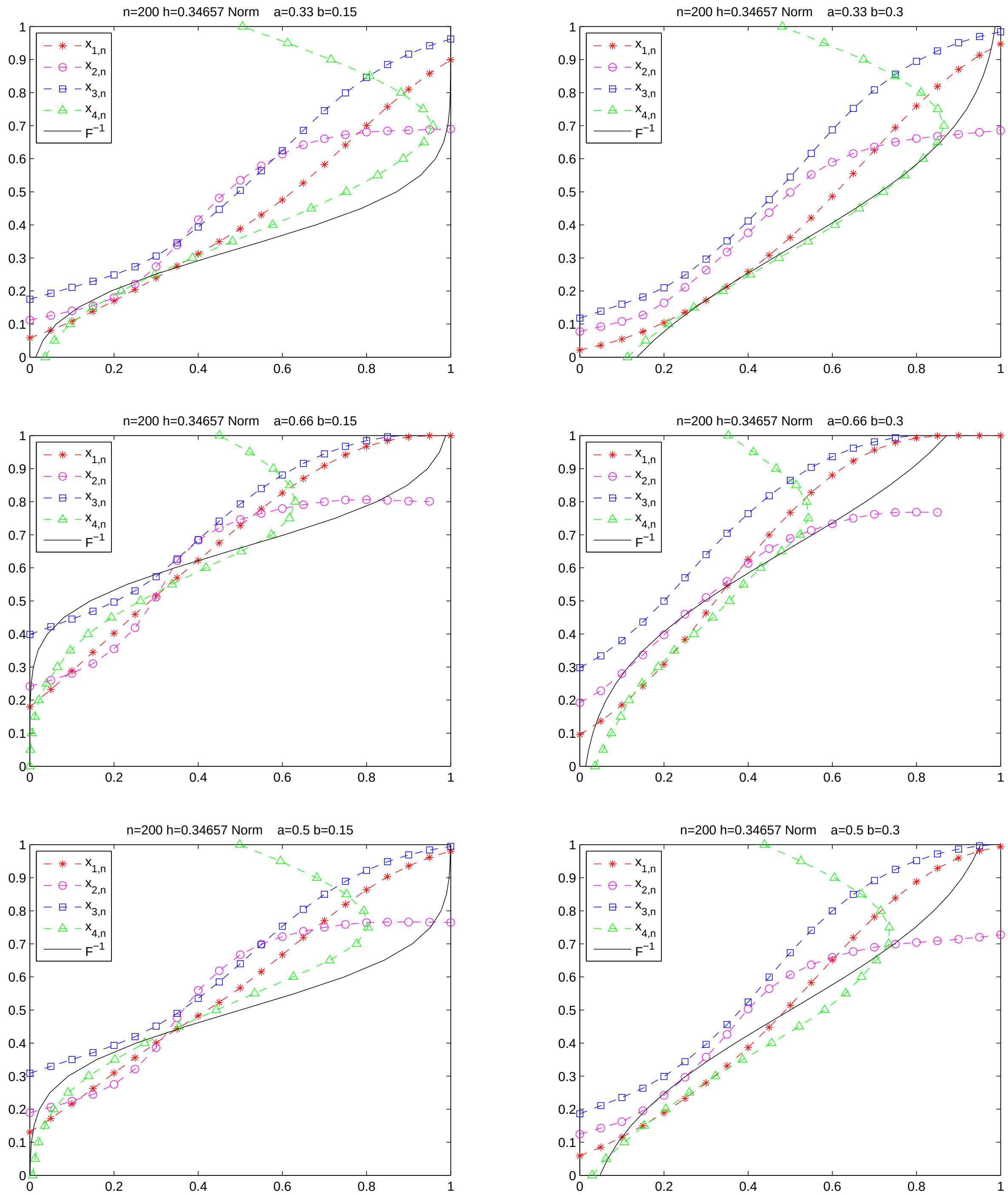


Рис. 9. Результаты моделирования для равномерно распределенной на $[0, 1]$ дозы X, $n = 50$

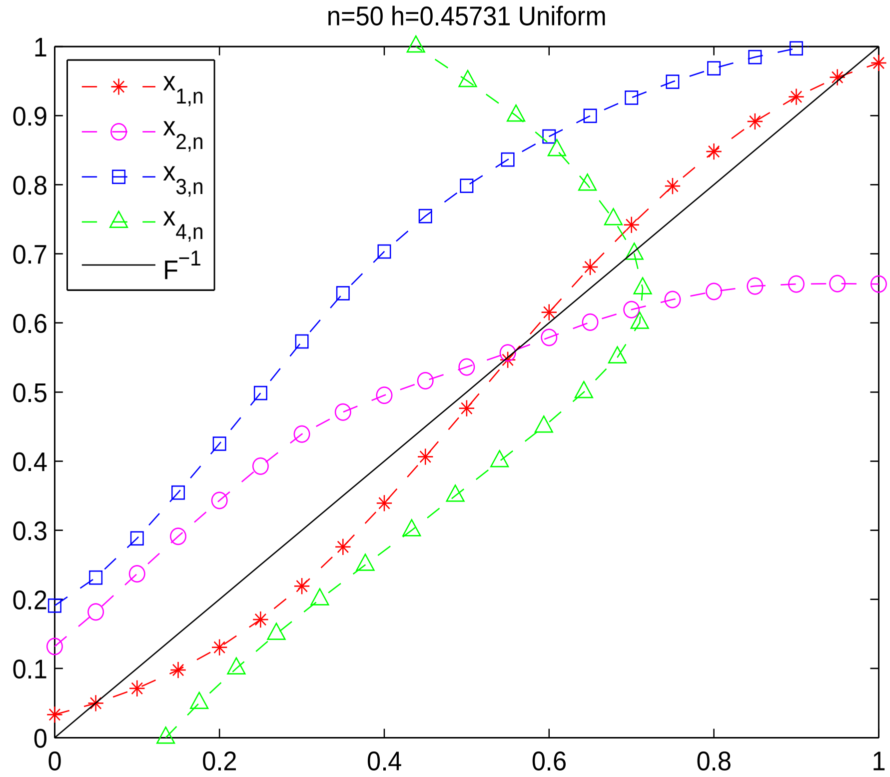


Рис. 10. Результаты моделирования для равномерно распределенной на $[0, 1]$ дозы X, $n = 200$