09.02.24. ~~III~~ СЕМЕСТР 6.

Модуль 1 (~10 нед)

1. Предельные теоремы теории вероятностей
2. Основные понятия
3. Точечные оценки
4. Интервальные оценки

} ДЗ1 РК1 ЛР1 ЛР2

Модуль 2 (~17 нед)

1. Проверка стат. гипотез } ДЗ2 РК2 ЛР3

Предельные теоремы теории вероятностей.

(1) Неравенства Чебышева.

Тк1 1е неравенство Чебышева

Пусть 1) X – сл. вел.
2) $X \geqslant 0$ (т.е. $P\{X<0\}=0$)
3) $\exists MX$

Тогда $\forall \varepsilon > 0 \quad P\{X \geqslant \varepsilon\} \leqslant \dfrac{MX}{\varepsilon}$ – 1е нер-во Чеб.

Док-во: (для случая непр. с. в. X)

Пусть f – функция плотности с. в. X.
Тогда по опр. $MX = \int\limits_{-\infty}^{+\infty} x f(x)\,dx = \left\{ \begin{array}{l} f(x)=0 \text{ при } x<0, \\ \text{т. к. } X \geqslant 0 \end{array} \right\} =$

$$= \int\limits_0^{+\infty} \underbrace{x}_{\geqslant 0} \underbrace{f(x)}_{\geqslant 0}\,dx = \underbrace{\int\limits_0^{\varepsilon} \underbrace{x}_{\geqslant 0} \underbrace{f(x)}_{\geqslant 0}\,dx}_{\geqslant 0} + \int\limits_{\varepsilon}^{+\infty} x f(x)\,dx \geqslant$$

$$\geqslant \int\limits_{\varepsilon}^{+\infty} x f(x)\,dx = \left\{ \begin{array}{l} x \in (\varepsilon; +\infty) \\ \Rightarrow x \geqslant \varepsilon \end{array} \right\} \geqslant \int\limits_{\varepsilon}^{+\infty} \varepsilon f(x)\,dx = \varepsilon \int\limits_{\varepsilon}^{+\infty} f(x)\,dx =$$

$$= \varepsilon \cdot P\{X \geqslant \varepsilon\}.$$

Т.е. $MX \geqslant \varepsilon \cdot P\{X \geqslant \varepsilon\}$

$\Rightarrow P\{X \geqslant \varepsilon\} \leqslant \dfrac{MX}{\varepsilon}$, ч.т.д.

Th2 $2^е$ неравенство Чебышева

Пусть 1) X — сл. вел;
2) $\exists DX$

Тогда
$$\forall \varepsilon > 0 \quad P\{|X - MX| \geq \varepsilon\} \leq \frac{DX}{\varepsilon^2}$$

Док-во:

$$DX = M[(X-MX)^2] \geq \left\{ \begin{array}{l} \text{Рассм. сл. вел.} \\ Y = (X-MX)^2 \geq 0 \\ \text{т.к. } \exists DX, \text{ то } \exists MY \end{array} \right\} \Rightarrow$$

применим к с.в. Y $1^е$ нер-во Чеб. для $\varepsilon_1 = \varepsilon^2$: $MY \geq \varepsilon_1 \cdot P\{Y \geq \varepsilon_1\}$

$$\geq \varepsilon^2 \cdot P\{(X-MX)^2 \geq \varepsilon^2\} = \varepsilon^2 P\{|X-MX| \geq \varepsilon\} \Rightarrow$$

$$\Rightarrow DX \geq \varepsilon^2 P\{|X-MX| \geq \varepsilon\}$$

$$\Rightarrow P\{|X-MX| \geq \varepsilon\} \leq \frac{DX}{\varepsilon^2}, \text{ чтд.}$$

Пример: Предельно допустимое давление в пневмокамере ракеты составляет 200 Па. После проверки большого числа ракет было получено среднее значение давления 150 Па. Оценить вер-ть того, что давление в пневмокамере случайно выбранной ракеты превысит предельно допустимое. Как изменится эта оценка, если дополнительно известно, что в результате проверки было также получено значение СКО давления, равное $\sigma = 5$ Па?

Решение:

1) Пусть X — сл. вел. прин. значение, равное давлению в пневмокамере очередной ракеты, Па

Тогда а) $X \geq 0$;
б) $MX = 150$ Па;
в) $DX = \sigma^2 = 25$ Па^2.

Нужно оценить ~~вер-ть того, что~~

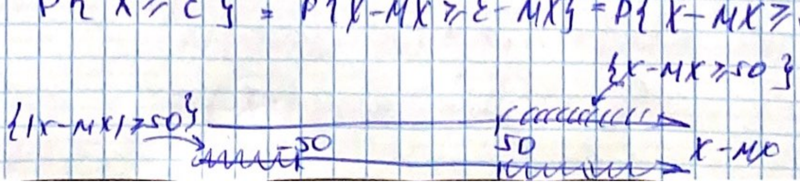
$$P\{X \geq \varepsilon\}, \text{ где } \varepsilon = 200 \text{ Па}$$

2) т.к. $X \geq 0$, то используем $1^е$ нер. Чеб:

$$P\{X \geq \varepsilon\} \leq \frac{MX}{\varepsilon} = \frac{150}{200} = \frac{3}{4}.$$

3) Используем $2^е$ нер. Чеб.

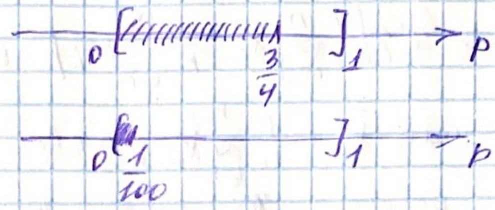
$$P\{X \geq \varepsilon\} = P\{X - MX \geq \varepsilon - MX\} = P\{X - MX \geq 50\} \leq$$



$$\overset{\circledast}{\le} P\{|X-MX|\ge 50\} \underset{\text{2е нер. Чеб.}}{\le} \frac{DX}{50^2} = \frac{25}{2500} = \frac{1}{100}$$

Замечание 1) 1е нер. Чеб. $P\{X\ge 200\}\le \frac{3}{4}$

2е нер. Чеб. $P\{X\ge 200\}\le \frac{1}{100}$



2) 2е нер. Чебышева дало сущ. более точную оценку, т.к. дополнительно использует инф-ю о величине дисперсии сл. вел. X.

3) Исп-е 1го нер. Чеб. для $\varepsilon \le MX$ и 2го нер. Чеб. для $\varepsilon \le \sqrt{DX}$ дает тривиальные оценки: $P \le 1$.

② Некоторые виды сходимости последовательностей с. в.

Пусть $X_1, X_2, \dots, X_n, \dots$ — посл-ть с.в., заданных на одном вероятностном пространстве.

Опр. Говорят, что последовательность $X_1, X_2, \dots$ с.в. сходится по вероятности к сл. в. Z, если

$$\forall \varepsilon > 0 \quad P\{|X_n - Z| \ge \varepsilon\} \xrightarrow[n\to\infty]{} 0 \qquad (*)$$

Обозн. $X_n \xrightarrow[n\to\infty]{P} Z$

Замечание Условие (*) из опр-я сходимости по вер-ти можно записать в экв. виде:

$$\forall \varepsilon > 0 \quad P\{|X_n - Z| < \varepsilon\} \xrightarrow[n\to\infty]{} 1.$$

Опр Говорят, что посл-ть $X_1, X_2, \dots$ сл. вел. слабо сходится к сл. вел. Z, если для любой точки x непрер-ти ф-ии распр. F_Z сл. вел. Z числовая посл-ть $F_{X_n}(x)$ значений ф-ий распр. с.в. X_n, $n \in \mathbb{N}$, сходится к значению $F_Z(x)$, т.е.

$$\forall x \in \mathbb{R},\ x \text{ — т. непр. } F_Z(x) \Rightarrow F_{X_n}(x) \xrightarrow[n\to\infty]{} F_Z(x).$$

Обозн. $F_{X_n}(x) \underset{n\to\infty}{\Longrightarrow} F_Z(x)$

③ Закон больших чисел.

Пусть 1) $X_1, X_2, \dots$ — посл-ть сл. вел., зад. на одном вер. пр-ве.

2) $\exists MX_i = m_i, \ i \in \mathbb{N}$

Опр. Говорят, что посл-ть $X_1, X_2, \dots$ сл. вел. удовлетворяет (слабому) закону больших чисел, если

$$\forall \varepsilon > 0 \quad P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_n - \frac{1}{n}\sum_{i=1}^{n} m_i\right| \geq \varepsilon\right\} \xrightarrow[n\to\infty]{} 0. \quad (**)$$

Замечания 1) Обозначим: $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \ n \in \mathbb{N}$

Тогда условие (**) можно записать в след. виде:

$$\forall \varepsilon > 0 \quad P\left\{\left|\bar{X}_n - \frac{1}{n}\sum_{i=1}^{n} m_i\right| \geq \varepsilon\right. \xrightarrow[n\to\infty]{} 0$$

Выполнение закона больших чисел для $\overbrace{X_1, X_2, \dots}^{\text{посл-ти}}$ означает, что при достаточно больших n значение с. в. $\bar{X}_n$ "почти не отклоняется" от детерм. величины $\frac{1}{n}\sum_{i=1}^{n} m_i$, т.е. для величины $\bar{X}_n$ эксперимент теряет случайный характер. (величина $\bar{X}_n$ вырождается в детерм. величину).

2) Обозначим: $Y_n = \frac{1}{n}\sum_{i=1}^{n} (X_i - m_i)$

Тогда условие (**) можно записать в виде

$$\forall \varepsilon > 0 \quad P\{|Y_n| \geq \varepsilon\} \xrightarrow[n\to\infty]{} 0$$

Т.о. выполнение ЗБЧ для посл-ти $X_1, X_2, \dots$ означает, что

$$Y_n \xrightarrow[n\to\infty]{P} 0.$$

Это означает, что при достаточно больших n сл. величина Y_n вырождается в детерм. величину $Z=0$.

Вопрос: какой д. б. последовательность $X_1, X_2, \dots$, чтобы она удовлетворяла ЗБЧ?

Th Чебышева (дост. условие того, что посл-ть с. в. удовл. ЗБЧ)

Пусть 1) $X_1, X_2, \dots$ — посл-ть попарно незав. с. в.

2) $\exists MX_i = m_i, \ i \in \mathbb{N}$
$\exists DX_i = \sigma_i^2$

3) Дисперсии ограничены в совок-ти, т.е.
$\exists C > 0 \ \forall i \in \mathbb{N} \ \sigma_i^2 \leq C$

Тогда посл-ть $X_1, X_2, \dots$ удовл. ЗБЧ.

Док-во:

1) Рассм. с.в. $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$, $i \in N$

Тогда

$$M\overline{X_n} = M\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} MX = \frac{1}{n}\sum_{i=1}^{n} m_i;$$

$$D\overline{X_n} = D\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2} D\left[\sum_{i=1}^{n} X_i\right] = \left\{\begin{matrix} X_i,\ i \in N \\ \text{поп. незав} \end{matrix}\right\} =$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} DX_i = \frac{1}{n^2}\sum_{i=1}^{n} \sigma_i^2$$

2) Зафикс. $\varepsilon > 0$ и применим к сл. вел. $\overline{X_n}$ $2^е$ нер. Чеб.

$$P\left\{\left|\overline{X_n} - \frac{1}{n}\sum_{i=1}^{n} m_i\right| \geq \varepsilon\right\} \leq \frac{D\overline{X_n}}{\varepsilon^2}$$

С учетом получ. выше значения $D\overline{X_n}$:

$$P\left\{\left|\overline{X_n} - \frac{1}{n}\sum_{i=1}^{n} m_i\right| \geq \varepsilon\right\} \leq \frac{1}{n^2\varepsilon^2}\sum_{i=1}^{n} \sigma_i^2.$$

Т.к. σ_i^2 ограничены в совок-ти, т.е

$$\forall i \in N \quad \sigma_i^2 \leq C, \text{ то } \sum_{i=1}^{n} \sigma_i^2 \leq \sum_{i=1}^{n} C = nC.$$

Т.о. $$P\left\{\left|\overline{X_n} - \frac{1}{n}\sum_{i=1}^{n} m_i\right| \geq \varepsilon\right\} \leq \frac{Cn}{n^2\varepsilon^2} = \frac{C}{n\varepsilon^2}$$

$$0 \leq P\left\{\left|\overline{X_n} - \frac{1}{n}\sum_{i=1}^{n} m_i\right| \geq \varepsilon\right\} \leq \frac{C}{n\varepsilon^2}$$

$n \to \infty$: $0 \to 0$

$$\frac{C}{n\varepsilon^2} \to 0$$

По Th о предельном переходе к нер-ву

$$P\left\{\left|\overline{X_n} - \frac{1}{n}\sum_{i=1}^{n} m_i\right| \geq \varepsilon\right\} \xrightarrow[n\to\infty]{} 0 \quad (***)$$

Т.к. относительно ε не сделано никаких доп. предположений, кроме $\varepsilon > 0$, то (***) вып-ся для любого $\varepsilon > 0$.

Т.о. посл-ть $X_1, X_2, \ldots$ удовл. ЗБЧ (см. замеч. 1), чтд.

Следствие 1 (th Чеб. для одинак. распр. с. в.)

Пусть 1) вып-но условие 1) th Чебышева
2) сл. вел. X_i $i \in \mathbb{N}$, одинаково распределены, причём
$$MX_i = m, \; DX_i = \sigma^2, \; i \in \mathbb{N}$$
Тогда пос-ть $X_1, X_2, \ldots$ удовл. ЗБЧ.

Док-во: Вып-е условие 2) рассм. следствие означает выполнение условие 2 th Чебышева.
При этом $DX_i = \sigma^2 \leq \sigma^2 = C$, что означает также вып-е усл. 3) th Чеб.
Т.о. пос-ть $X_1, X_2, \ldots$ удовл. ЗБЧ:
$$\forall \varepsilon > 0 \; P\{|\overline{X}_n - m| \geq \varepsilon\} \xrightarrow[n\to\infty]{} 0$$

16.02.24. **Лекция 2.**

Следствие 2. (th Бернулли, ЗБЧ в ф. Бернулли)

Пусть 1) проводится n испытаний по схеме Бернулли с вер-тью p успеха в отд. исп-ии
2) $r_n = \dfrac{\text{число успехов в серии из } n \text{ исп.}}{n}$ – наз. наблюдаемой частотой успеха

Тогда
$$r_n \xrightarrow[n\to\infty]{P} p$$
↑наз. теор. частотой успеха

Док-во: 1) Рассмотрим с. в. X_i:
$$X_i = \begin{cases} 1, & \text{если в } i\text{-м исп. серии – успех;} \\ 0, & \text{иначе,} \end{cases}$$
$i = \overline{1;n}$

Тогда набл. частота успеха
$$r_n = \frac{\sum_{i=1}^{n} X_i}{n}$$

2) Применим к пос-ти $X_1, X_2, \ldots$ th Чебышева:

а) $X_1, X_2, \ldots$ – нез. в совок-ти, т.к. отд. исп. в схеме Бернулли явл. незав.
$\Rightarrow X_1, X_2, \ldots$ – незав. попарно

б) $\left.\begin{array}{l} \exists \; MX_i = p \\ \exists \; DX_i = pq, \quad q = 1-p. \end{array}\right\} i \in \mathbb{N}$

в) Дисперсии ограничены в совок-ти, например, $C = pq$.

Тогда для посл-ти вып-ся ЗБЧ, т.е.

$$\forall \varepsilon > 0 \quad P\{|\underbrace{\frac{1}{n}\sum_{i=1}^{n} X_i}_{\equiv\, r_n} - \underbrace{\frac{1}{n}\sum_{i=1}^{n} m_i}_{\equiv\, p}| \geq \varepsilon\} \xrightarrow[n\to\infty]{} 0$$

$$\forall \varepsilon > 0 \quad P\{|r_n - p| \geq \varepsilon\} \xrightarrow[n\to\infty]{} 0,$$

т.е. $r_n \xrightarrow[n\to\infty]{P} p$, чтд.

Замечание. Тh Бернулли говорит о том, что при дост. больших n наблюдаемая частота успеха r_n почти не отличается от теоретич. частоты p.

4 Центральная предельная Th

Пусть (1) $X_1, X_2, \ldots$ – посл-ть незав. одинаково распр-ных с. в.

(2) $\exists MX_i = m$, $i \in N$
$\exists DX_i = \sigma^2$

Рассм. величину $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$, $n \in N$

Тогда $M[\overline{X_n}] = M[\frac{1}{n}\sum_{i=1}^{n} X_i] = \frac{1}{n}\sum_{i=1}^{n} \overbrace{MX_i}^{m} = m$

$$D[\overline{X_n}] = D[\frac{1}{n}\sum_{i=1}^{n} X_i] = \frac{1}{n^2} D[\sum_{i=1}^{n} X_i] = \{X_i, i \in N - \text{незав.}\} =$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{\sigma^2}{n}$$

Рассм. сл. вел.:

$$Y_n = \frac{\overbrace{\overline{X_n} - M[\overline{X_n}]}^{\text{центрирование}}}{\underbrace{\sqrt{D[\overline{X_n}]}}_{\text{нормиро-вание}}} = \frac{\overline{X_n} - m}{\sigma/\sqrt{n}}, \quad n \in N$$

Тогда

$$MY_n = M[\frac{\overline{X_n} - M\overline{X_n}}{\sqrt{D\overline{X_n}}}] = \frac{1}{\sqrt{D\overline{X_n}}} \cdot (M\overline{X_n} - M\overline{X_n}) = 0.$$

$$DY_n = D[\frac{\overline{X_n} - M\overline{X_n}}{\sqrt{D\overline{X_n}}}] = \frac{1}{D\overline{X_n}} D[\overline{X_n} - \overbrace{M\overline{X_n}}^{const}] = D\overline{X_n}/D\overline{X_n} = 1$$

Th ЦПТ в форме Ляпунова б/д

Пусть выполнены условия (1) – (2). $Z \sim N(0,1)$

Тогда посл-ть с.в. $Y_1, Y_2, \dots$ слабо сходится к с.в. Z, имеющей станд. норм. распр., т.е.

$$\forall x \in R \quad F_{Y_n}(x) \xrightarrow[n\to\infty]{} \underbrace{\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt}_{F_Z(x)}$$

Замечание: Th Чебышева (следствие 1 из нее) утверждает, что $\overline{X_n} \xrightarrow[n\to\infty]{P} m$.

ЦПТ уточняет характер этой сходимости:

$$\overline{X}_n \sim N(m, \sigma^2/n)$$

при дост. больших n.

$$\begin{cases} Y_n \underset{\text{прибл.}}{\sim} N(0,1), \quad n \gg 1 \\ \dfrac{\overline{X}_n - m}{\sigma/\sqrt{n}} = Y_n \Rightarrow \overline{X}_n = \dfrac{\sigma}{\sqrt{n}} \cdot Y_n + m \\ M\overline{X}_n = m, \; D\overline{X}_n = \dfrac{\sigma^2}{n} DY_n = \dfrac{\sigma^2}{n} \end{cases}$$

Пример:

ЭВМ производит суммирование $n = 10000$ чисел, каждое из которых округлено до 10^{-4}.

Считая, что ошибки округления независимы и распределены равномерно в диапазоне $(-0.5 \cdot 10^{-4}, +0.5 \cdot 10^{-4})$ найти диапазон, в котором с вер. 0.95 будет заключена ошибка результата суммирования.

Решение:

1) Пусть X_i – с.в. прин. значения, равные ошибке округления i-го числа, $i = \overline{1;n}$

Тогда $X_i \sim R(-0.5 \cdot 10^{-4}, 0.5 \cdot 10^{-4})$, $i = \overline{1;n}$.

$$m = MX_i = 0$$

$$\sigma^2 = DX_i = \frac{(10^{-4})^2}{12} = \frac{10^{-8}}{12} \approx 10^{-9}$$

При этом

$$S = \sum_{i=1}^{n} X_i \text{ – ошибка результата суммирования}$$

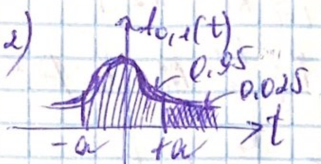
$$\Big\{ x_i = x_i^0 + X_i \; ; \; \underbrace{\sum_{i=1}^{n} x_i}_{\text{рез. сумм.}} = \sum_{i=1}^{n} x_i^0 + S$$

x_i^0 – точное знач. i-го слаг.; $\sum_{i=1}^{n} x_i^0$ – точное знач. суммы

Рассмотрим с.в. $Y_n = (\bar X_n - M\bar X_n)/\sqrt{D\bar X_n} = \frac{\bar X_n - 0\cdot m}{\sqrt{10^{-9}}/\sqrt{n}} = \frac{\bar X_n \cdot 10^{2}}{\sqrt{\frac{10^{-8}}{10}}} \approx$

$\approx \frac{\bar X_n \cdot 10^{2}}{\frac{1}{3}\cdot 10^{-4}} = 3\cdot 10^{6}\cdot \bar X_n \sim N(0,1)$

приближенно, с учетом ЦПТ и того, что $n = 10^4$, $n \gg 1$.

г) 

Нужно подобрать a так, чтобы

$$P\{|Y_n| < a\} = 0.95$$

Т.о. слева от a располагается 0.975 кг вер. массы.

$\Rightarrow$ $a = q_{0.975}$ — квантиль уровня $\alpha = 0.975$ стандартного норм. распределения

Из таблиц $\Rightarrow a = q_{0.975} = 1.96$

Т.о. с вероятностью 0.95: $-1.96 < Y_n < 1.96$

т.е. $|Y_n| < 1.96$

$|3\cdot 10^{+6}\cdot \bar X_n| < 1.96$

$|3\cdot 10^{+6}\cdot \frac{1}{n}\underbrace{\sum_{i=1}^{n} X_i}_{S}| < 1.96$

$|3\cdot 10^{2}\cdot S| < 1.96$

$|S| < \frac{1.96}{3\cdot 10^{2}} \approx 6.5\cdot 10^{-3}$

Ответ: $(-6.5\cdot 10^{-3}; +6.5\cdot 10^{-3})$

Следствие 1 (Интегр. теорема Муавра-Лапласа)

Пусть 1) проводится $n \gg 1$ испытаний по сх. Бернулли с вер. успеха p в одном исп.

2) k — общее число успехов в серии

Тогда

$$P\{k_1 \le k \le k_2\} \approx \Phi_0(x_2) - \Phi_0(x_1),$$

где $x_i = \frac{k_i - np}{\sqrt{npq}}$, $i = \overline{1,2}$, $q = 1-p$.

Док-во: 1) Рассмотрим с.в.

$X_i = \begin{cases} 1, & \text{если в } i\text{-ом испытании имел место усп.} \\ 0, & \text{иначе} \end{cases}$

$i = \overline{1;n}$

Тогда $MX_i = p$, $DX_i = pq$

2) $k = \sum_{i=1}^{n} X_i$

2) Рассм. с.в.

$$Y_n = \frac{\overline{X}_n - M\overline{X}_n}{\sqrt{D\overline{X}_n}} \underset{\text{прибл., т.к. } n \gg 1}{\sim} N(0,1)$$

$$P\{k_1 \le k \le k_2\} = P\{k_1 \le \sum_{i=1}^{n} X_i \le k_2\} =$$

$$= P\{\frac{k_1}{n} \le \overline{X}_n \le \frac{k_2}{n}\} = P\{\frac{k_1}{n} - p \le \overline{X}_n - p \le \frac{k_2}{n} - p\} =$$

$$= P\{\underbrace{\frac{\frac{k_1}{n} - p}{\sqrt{pq}/\sqrt{n}}}_{x_1} \le \underbrace{\frac{\overline{X}_n - p}{\sqrt{pq}/\sqrt{n}}}_{\sim Y_n} \le \underbrace{\frac{\frac{k_2}{n} - p}{\sqrt{pq}/\sqrt{n}}}_{x_2}\} \approx$$

$$\underset{Y_n \sim N(0,1)}{\approx} \Phi_0(x_2) - \Phi_0(x_1), \text{ чтд.}$$

~~Замечание~~

Пример.

В эксперименте Пирсона по подбрасыванию правильной монеты из $n = 24000$ подбрасываний герб выпал 12012 раз. Какова вер-ть того, что при повт. проведении эксперимента отклонение набл. частоты выпадения герба от теор. окажется таким же или больше.

Решение: