1. **Сформулировать и доказать первое неравенство Чебышева**.

① Неравенство Чебышева

Теорема №1: (1-ое нерав. Ч.)

Пусть 1) X - с/в;
2) $X \geq 0$, т.е. $P\{X<0\}=0$)
3) $\exists M(X)$

Тогда $\forall \varepsilon > 0$ $P\{X \geq \varepsilon\} \leq \frac{M(X)}{\varepsilon}$ (1-ое н.ч.)

Док-во: (для НСВ) $M(X) = \int\limits_{-\infty}^{+\infty} x f_X(x)\,dx = \{X \geq 0\} = \int\limits_{0}^{+\infty} x f_X(x)\,dx =$

$= \{$св-во аддитивности инт.$\} = \underbrace{\int\limits_{0}^{\varepsilon} x f_X(x)\,dx}_{\geq 0} + \int\limits_{\varepsilon}^{+\infty} x f_X(x)\,dx \geq \int\limits_{\varepsilon}^{+\infty} x f_X(x)\,dx$

$\left\{ \begin{array}{l} X \geq \varepsilon \Rightarrow \\ x f_X(x) \geq \varepsilon f_X(x) \end{array} \right\} \Rightarrow \int\limits_{\varepsilon}^{+\infty} x f_X(x)\,dx \geq \int\limits_{\varepsilon}^{+\infty} \varepsilon f_X(x)\,dx = \varepsilon \int\limits_{\varepsilon}^{+\infty} f_X(x)\,dx =$

$= \varepsilon P\{X \geq \varepsilon\}$

Т.о. $M(X) \geq \varepsilon \cdot P\{X \geq \varepsilon\} \Rightarrow P\{X \geq \varepsilon\} \leq \frac{M(X)}{\varepsilon}$

2. **Сформулировать и доказать второе неравенство Чебышева.**

Теорема №2: (2-ое нерав. Ч.)

Пусть 1) X - с/в;
2) $\exists M(X)$, $\exists D(X)$

Тогда $\forall \varepsilon > 0$ $P\{|X - M(X)| \geq \varepsilon\} \leq \frac{D(X)}{\varepsilon^2}$

Док-во:

1) $D(X) = M[(X - M(X))^2] = \left\{ \begin{array}{l} \text{Расм. с/в } Y = (X - M(X))^2 \geq 0 \\ \text{Применим к } Y \text{ 1-ое нер. Чеб.:} \\ P\{Y \geq \varepsilon_1\} \leq M(Y)/\varepsilon_1 \text{ для } \varepsilon_1 = \varepsilon^2 \\ M(Y) \geq \varepsilon_1 \cdot P\{Y \geq \varepsilon_1\} \end{array} \right\} =$

$= M(Y) \geq P\{Y \geq \varepsilon^2\} \cdot \varepsilon^2 = P\{(X - M(X))^2 \geq \varepsilon^2\} \cdot \varepsilon^2 = \cancel{P\{(X-M(X))^2}$

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

Т.о. $D(X) \geq \varepsilon^2 \cdot P\{|X - M(X)| \geq \varepsilon\} \Rightarrow P\{|X - M(X)| \geq \varepsilon\} \leq \frac{D(X)}{\varepsilon^2}$.

3. **Для последовательности случайных величин сформулировать определения сходимости по вероятности и слабой сходимости. Сформулировать закон больших чисел.**

② Сходимость последовательных с/в.

Пусть $X_1, X_2, \ldots$ – послед-ть с/в.

Опр.: говорят, что послед-ть с/в $X_1, X_2, \ldots$ сходится по вероятности к с/в Z, если:

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

Обозн. $a_n = P\{|X_n - Z| \geq \varepsilon\}$ – число. В опред. написано: числовая послед. a_n сходится к 0: $a_n \xrightarrow[n \to \infty]{} 0$.

Обозн.: $X_n \xrightarrow[n \to \infty]{P} Z$. $\boxed{2}$

Опр: говорят, что послед-ть с/в $X_1, X_2, \ldots$ слабо сходится к с/в Z, если $\forall x \in R$ такого, что F_Z непрерывна в x, вып.:

$$F_{X_n}(x) \xrightarrow[n \to \infty]{} F_Z(x).$$

(Здесь F_Z – ф. распр. с/в Z, F_{X_n}, $n \in N$ – ф. распр. с/в X_n)

В каждой точке x, в которой F_Z непрерывна, числовая послед-ть

$$\underset{\text{(число)}}{F_{X_n}(x)} \xrightarrow[n \to \infty]{} \underset{\text{(число)}}{F_Z(x)}$$

Обозн.: $F_{X_n} \underset{n \to \infty}{\Rightarrow} F_Z$.

4. **Сформулировать закон больших чисел. Доказать закон больших чисел в форме Чебышева.**

(3) Закон больших чисел

Пусть $X_1, X_2, \dots$ — посл-ть с/в, $\exists M(X_i) = m_i$, $i \in \mathbb{N}$

Опр: Говорят, что посл-ть $X_1, X_2, \dots$ удовл. закону больших чисел, если:

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

Замечание: 1) Обозн: $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, $n \in \mathbb{N}$, т.е.

$$\overline{X}_1 = X_1$$
$$\overline{X}_2 = \frac{1}{2}[X_1 + X_2]$$
$$\dots$$
$\overline{X}_n$ — с/в.

Теорема Чебышева (ЗБЧ в форме Чебышева — достаточное условие того, что посл. с/в удовл. ЗБЧ)

Пусть: 1) $X_1, X_2, \dots$ — посл. независ. с/в.

2) $\exists M(X_i) = m_i$, $\exists D(X_i) = \sigma_i^2$, $i \in \mathbb{N}$

3) Дисперсии с/в $X_1, X_2, \dots$ ограничены в совокупности:

$$(\exists C > 0)(\sigma_i^2 \leq C,\ i \in \mathbb{N})$$

$\boxed{3}$

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

Док-во: 1) Рассм. с/в $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Применим к с/в $\overline{X}_n$ 2НЧ:

$$P\{|\overline{X}_n - M[\overline{X}_n]| \geq \varepsilon\} \leq \frac{D(\overline{X}_n)}{\varepsilon^2}$$

$$M[\overline{X}_n] = \frac{1}{n}\sum_{i=1}^{n} m_i$$

$$D[\overline{X}_n] = D\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2} D\left[\sum_{i=1}^{n} X_i\right] = \{X_i \text{ независ.}\} =$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} D(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma_i^2 \leq \{\sigma_i^2 \leq C\} \leq \frac{1}{n^2}\sum_{i=1}^{n} C = \frac{n \cdot C}{n^2} = \frac{C}{n}.$$

$*$

Т.о. $\forall \varepsilon > 0 \quad 0 \leq P\{|\overline{X}_n - \frac{1}{n}\sum m_i| \geq \varepsilon\} \leq \frac{D(\overline{X}_n)}{\varepsilon^2} \leq \frac{C}{n\varepsilon^2}$

$**$

Устремим $n \to \infty$. Тогда по теореме о 2 милиционерах

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

5. **Сформулировать закон больших чисел в форме Чебышева. Доказать следствие этого закона для одинаково распределенных случайных величин и закон больших чисел в форме Бернулли.**

Ответ на первый вопрос смотреть в ответах на вопрос 4.

Следствие: (Th Чеб.)

Пусть: 1) $X_1, X_2, \dots$ — посл-ть независ. с/в
2) Все X_i, $i \in N$ одинаково распр.
3) $\exists M(X_i) = m$, $\exists D(X_i) = \sigma_i^2$, $i \in N$

Тогда: $\overline{X}_n \xrightarrow[n \to \infty]{P} m$

Док-во: 1) Условие 1 и 2 Th Чеб. вып-ны.
Условие 3 также вып-но, т.к. $\sigma_i^2 = \sigma^2 \le \sigma^2 = C$
Поэтому посл. $X_1, X_2, \dots$ удовл. ЗБЧ.

$\forall \varepsilon > 0: P\{|\overline{X}_n - \frac{1}{n}\sum_{i=1}^{n} m| \ge \varepsilon\} \xrightarrow[n \to \infty]{} 0 \Leftrightarrow \overline{X}_n \xrightarrow[n \to \infty]{P} m$

Следствие. Th Бернулли (ЗБЧ в форме Бернулли)

Пусть: 1) Проводится n испытаний по сх. Бернулли с вер. успеха p в отд. испытании.
2) $r_n = \dfrac{\{\text{число успехов в серии из } n \text{ исп.}\}}{n}$

наблюдаемая частота успеха в серии из n исп.

Тогда: $r_n \xrightarrow[n \to \infty]{P} p$

Док-во: 1) Рассм. с/в $X_i = \begin{cases} 1, & \text{успех} \\ 0 & \text{иначе} \end{cases}$

$X_i \sim B(1, p)$, $i \in N \Rightarrow M(X_i) = p$; $D(X_i) = pq$
(одинаково распр.)

X_i независ. (т.к. испытаний в сх. Бернулли нез.)

Т.о. $X_1, X_2 \dots$ удовл. следствию 1 из Th Чеб. $\Rightarrow \overline{X}_n \xrightarrow[n \to \infty]{P} M(X_i) = p$

2) $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i = r_n$

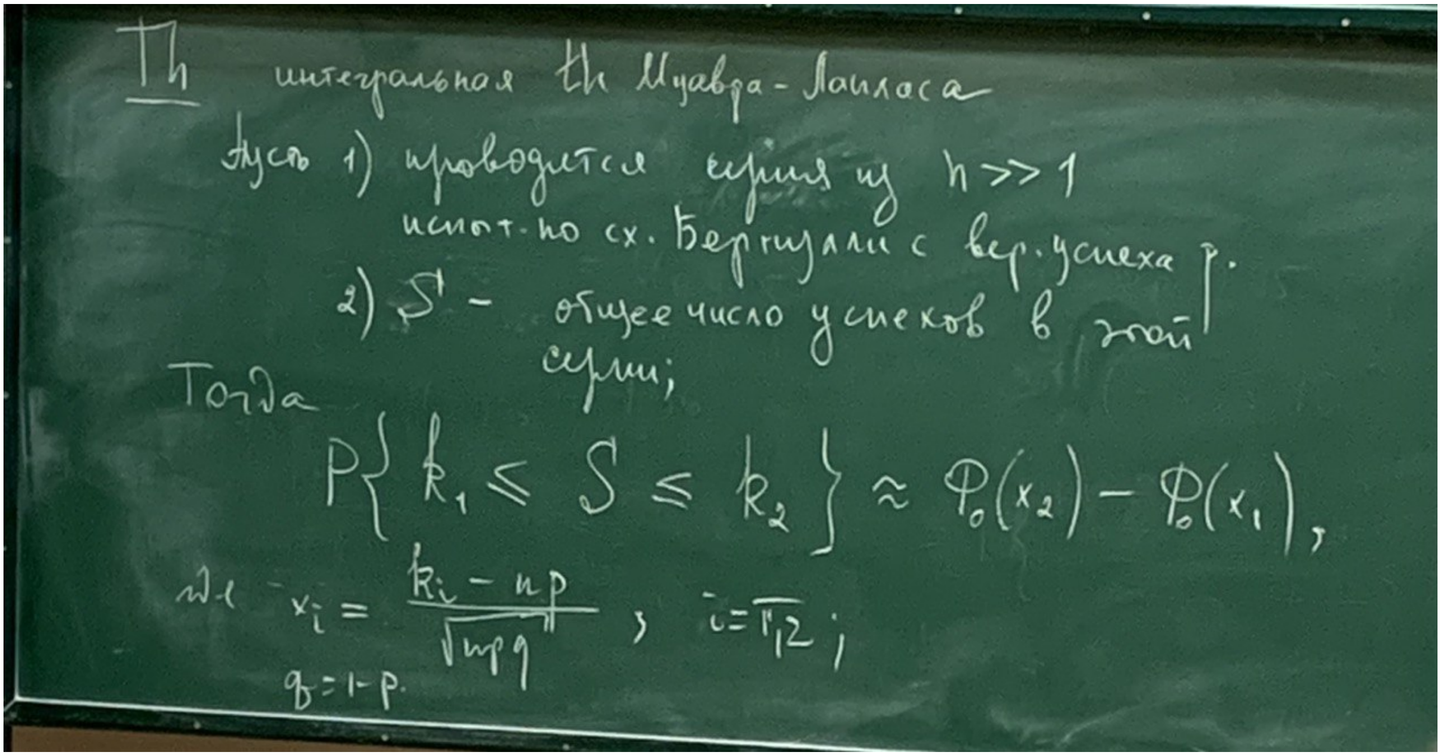
Т.о. $r_n \xrightarrow[n \to \infty]{P} p$.

Замечание: ЗБЧ в форме Бернулли означает, что при дост. большом объеме n

$r_n \approx p$.

5

6. **Сформулировать центральную предельную теорему. Доказать интегральную теорему Муавра-Лапласа.**



Th интегральная th Муавра-Лапласа

Пусть 1) проводится серия из $n >> 1$ испыт-ий по сх. Бернулли с вер. успеха p.

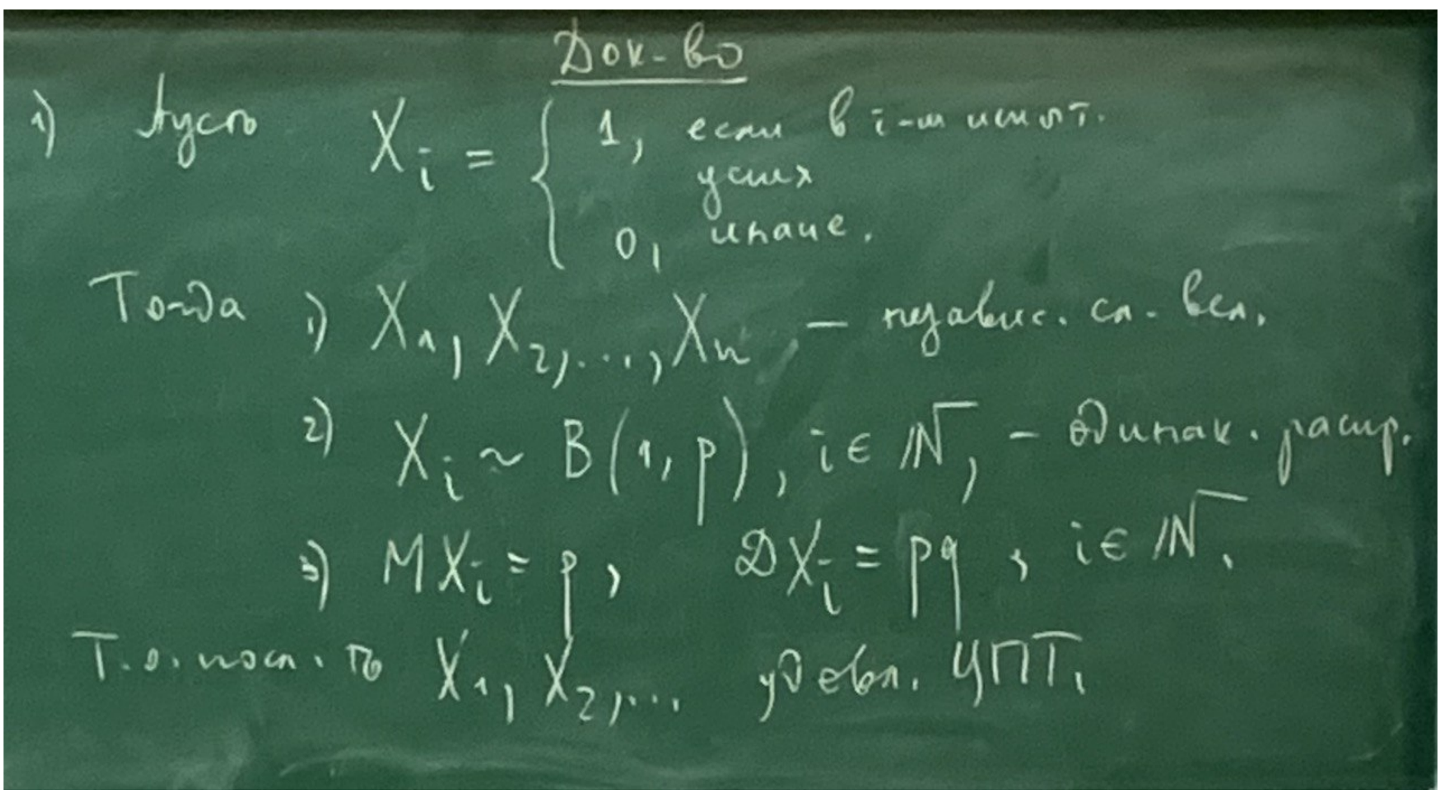
2) S — общее число успехов в этой серии;

Тогда

$$P\{k_1 \leqslant S \leqslant k_2\} \approx \Phi_0(x_2) - \Phi_0(x_1),$$

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

$q = 1 - p$.



Док-во

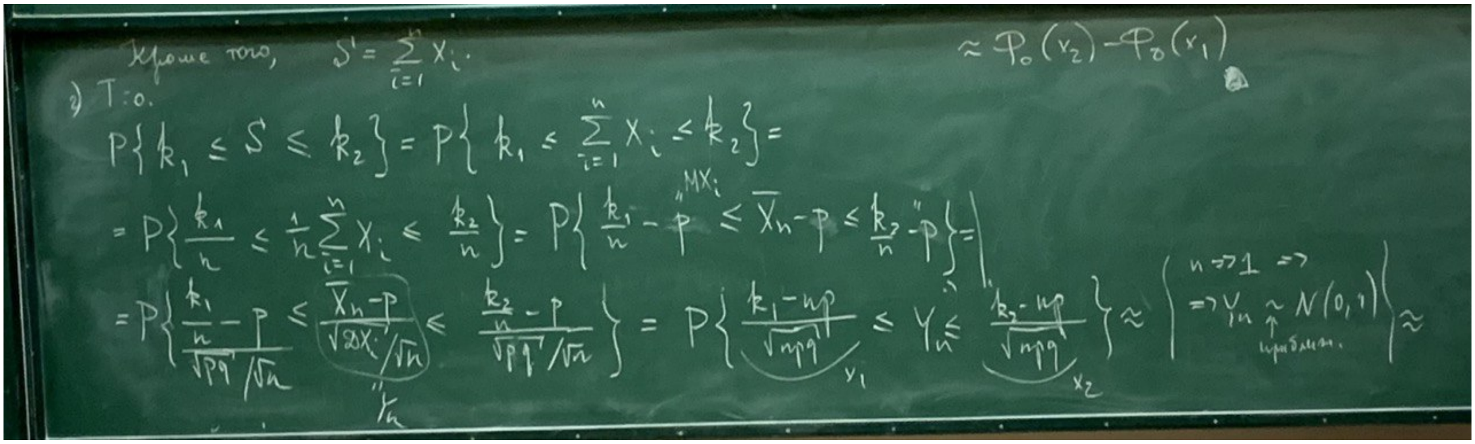
1) Пусть $X_i = \begin{cases} 1, & \text{если в } i\text{-м испыт. успех} \\ 0, & \text{иначе.} \end{cases}$

Тогда 1) $X_1, X_2, \ldots, X_n$ — независ. сл. вел.

2) $X_i \sim B(1, p)$, $i \in \mathbb{N}$, — одинак. распр.

3) $MX_i = p$, $DX_i = pq$, $i \in \mathbb{N}$.

Т.о. посл-ть $X_1, X_2, \ldots$ удовл. ЦПТ.



Кроме того, $S' = \sum_{i=1}^{n} X_i$. $\approx \Phi_0(x_2) - \Phi_0(x_1)$

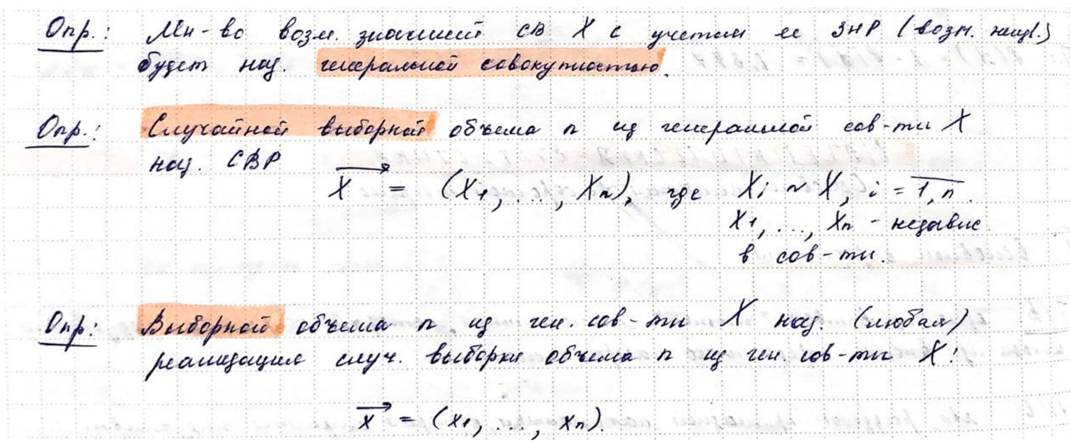
2) Т.о.

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

$$= P\left\{\frac{k_1}{n} \le \frac{1}{n}\sum_{i=1}^{n} X_i \le \frac{k_2}{n}\right\} = P\left\{\frac{k_1}{n} - \overset{MX_i}{p} \le \overline{X}_n - p \le \frac{k_2}{n} - p\right\} =$$

$$= P\left\{\frac{\frac{k_1}{n} - p}{\sqrt{pq}/\sqrt{n}} \le \underbrace{\frac{\overline{X}_n - p}{\sqrt{DX_i}/\sqrt{n}}}_{Y_n} \le \frac{\frac{k_2}{n} - p}{\sqrt{pq}/\sqrt{n}}\right\} = P\left\{\underbrace{\frac{k_1 - np}{\sqrt{npq}}}_{x_1} \le Y_n \le \underbrace{\frac{k_2 - np}{\sqrt{npq}}}_{x_2}\right\} \approx \left| \begin{array}{l} n \gg 1 \Rightarrow \\ \Rightarrow Y_n \underset{\text{приближ.}}{\sim} N(0,1) \end{array} \right| \approx$$

7. **Сформулировать определение случайной выборки и выборки, вариационного ряда. Записать и обосновать выражения для функций распределения случайной выборки и крайних членов вариационного ряда.** Про крайние значения



Опр.: Мн-во возм. значений СВ X с учетом ее ЗНР (возм. наук.) будет наз. генеральной совокупностью.

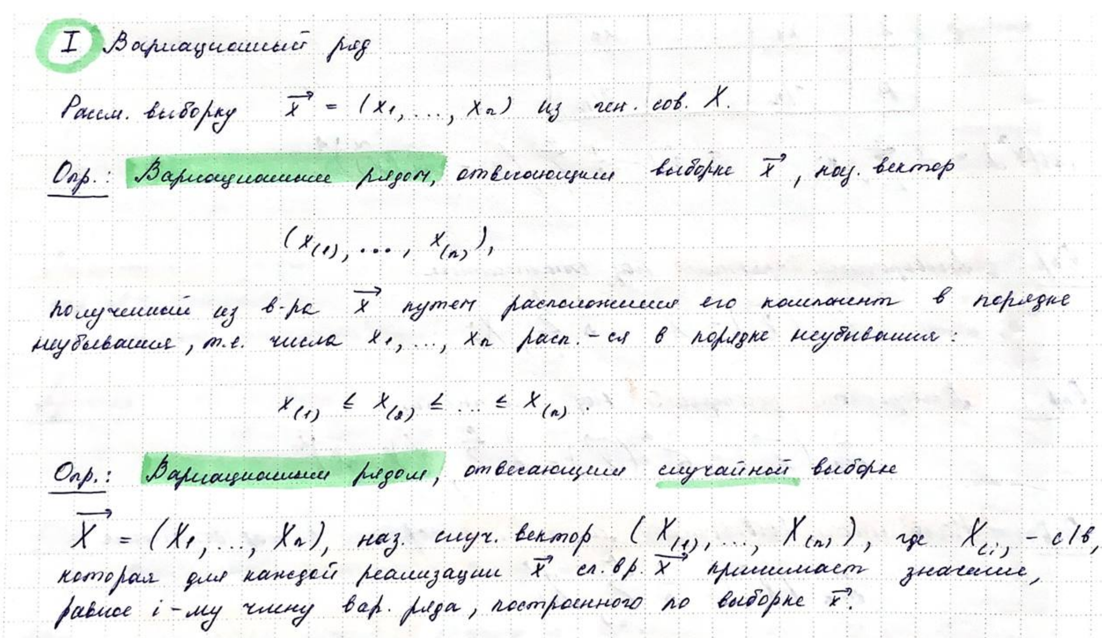
Опр.: Случайной выборкой объема n из генеральной сов-ти X наз. СВР

$$\vec{X} = (X_1, \dots, X_n), \text{ где } X_i \sim X,\ i = \overline{1,n},$$

$X_1, \dots, X_n$ — независ. в сов-ти.

Опр.: Выборкой объема n из ген. сов-ти X наз. (любая) реализация случ. выборки объема n из ген. сов-ти X:

$$\vec{x} = (x_1, \dots, x_n).$$



I Вариационный ряд

Рассм. выборку $\vec{x} = (x_1, \dots, x_n)$ из ген. сов. X.

Опр.: Вариационным рядом, отвечающим выборке $\vec{x}$, наз. вектор

$$(x_{(1)}, \dots, x_{(n)}),$$

полученный из в-ра $\vec{x}$ путем расположения его компонент в порядке неубывания, т.е. числа $x_1, \dots, x_n$ расп.-ся в порядке неубывания:

$$x_{(1)} \le x_{(2)} \le \dots \le x_{(n)}.$$

Опр.: Вариационным рядом, отвечающим случайной выборке $\vec{X} = (X_1, \dots, X_n)$, наз. случ. вектор $(X_{(1)}, \dots, X_{(n)})$, где $X_{(i)}$ — с/в, которая для каждой реализации $\vec{x}$ сл. в-ра $\vec{X}$ принимает значение, равное i-му члену вар. ряда, построенного по выборке $\vec{x}$.

2) Пусть F – ф-ия распр. св X. Тогда функция р. случайной выборки объема n из ген. св-ти X:

11. $F_{\vec{X}}(t_1, \dots, t_n) = P\{X_1 < t_1, \dots, X_n < t_n\} = \{X_i \text{ нез.}\} = P\{X_1 < t_1\} \cdot \ldots \cdot P\{X_n < t_n\} =$

$= F_{X_1}(t_1) \cdot \ldots \cdot F_{X_n}(t_n) = \{X_i \sim X \Rightarrow F_{X_i} = F\} = F(t_1) \cdot \ldots \cdot F(t_n)$

одинак. распр.

я ничего в лекциях не нашла, мб имеется в виду из вариационного ряда $x_{(i)}$

8. **Сформулировать определение начальных и центральных выборочных моментов порядка k, выборочного среднего и выборочной дисперсии. Являются ли эти статистики несмещенными оценками своих теоретических аналогов?**

Опр: Мн-во всех возм. значений сл. выборки $\vec{X}$ объема n из ген. сов. X наз. выборочным пространством и обозн. $\mathcal{X}_n$.

Опр.: Любую скалярную ф-ию $g(\vec{X})$ сл. выборки $\vec{X}$ будем наз. статистикой. При этом значение $g(\vec{x})$ этой ф-ии на выборке $\vec{x} \in \mathcal{X}_n$ будем наз. выборочным значением статистики g.

Замечание: Пусть $\vec{x}$ – выборка объема n из ген. сов. X. Это позволяет считать, что ЗНР сл. вел. X моделируется ЗНР ДСВ $\tilde{X}$:

$\tilde{X}$	x_1	...	x_n
p	$1/n$	...	$1/n$

$$M(\tilde{X}) = \frac{1}{n}\sum_{i=1}^{n} x_i; \qquad D(\tilde{X}) = \frac{1}{n}\sum_{i=1}^{n}(x_i - M(\tilde{X}))^2.$$

Опр.: Выборочным средним наз. статистика
$$\hat{m}_1(\vec{X}) = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Опр.: Выборочной дисперсией наз. статистика
$$\hat{m}_2(\vec{X}) = \hat{\sigma}^2(\vec{X}) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

Опр: Начальным выборочным моментом порядка k наз. статистика
$$\hat{m}_k(\vec{X}) = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

Опр: Центральным выборочным моментом порядка k наз. статистика
$$\hat{\mathring{m}}_k(\vec{X}) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k$$

Замечание: С учетом предыдущего замечания о том, что ЗНР СВ $\tilde{X}$ моделирует з-н распределения ген. сов. X, можно утверждать, что выборочные моменты явл. оценками своих теоретических (т.е. "истинных", таких, которые "на самом деле", и которые нам неизвестны) аналогов.

Про несмещенность пока не нашла

9. **Сформулировать определения эмпирической функции распределения и выборочной функции распределения. Сформулировать и доказать теорему о сходимости выборочной функции распределения.**

III Эмпирическая и выборочная ф-ии распределения

Пусть $\vec{x} = (x_1, \dots, x_n)$ - выборка из чел. сов. X. Обозначим $n(t, \vec{x})$ - число компонент вектора $\vec{x}$, которые меньше, чем t.

Опр.: Эмпирической ф-ией распр., построенной по выборке $\vec{x}$, наз. ф-ия

$$F_n : R \to R,$$

опр. правилом

$$F_n(t) = \frac{n(t, \vec{x})}{n}.$$

Замечание: 1) Ф-ия F_n может принимать значения из мн-ва

$$\left\{ 0 = \frac{0}{n}, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, \frac{n}{n} = 1 \right\}.$$

2) F_n явл. кусочно-постоянной ф-ией, которая имеет скачки (разрывы 1 рода) в точках

$$t = z_{(i)}$$

($z_{(i)}$ - все попарно различные значения компонент выборки $\vec{x}$).

3) Если все компоненты выборки $\vec{x}$ попарно различны, то

$$F_n(t) = \begin{cases} 0, & t \le x_{(1)} \\ \frac{1}{n}, & x_{(1)} < t \le x_{(2)} \\ \dots \\ \frac{n-1}{n}, & x_{(n-1)} < t \le x_{(n)} \\ 1, & x_{(n)} < t. \end{cases}$$

4) Эмп. ф. р. обладает всеми свойствами ф. р.

Пусть $\vec{X} = (X_1, \dots, X_n)$ - случ. выборка из чел. сов. X. Обозн. $n(t, \vec{X})$ - с/в, которая для каждой реализации $\vec{x}$ сл. выборки $\vec{X}$ принимает значение, равное $n(t, \vec{x})$.

Опр.: Выборочной ф. р., отвечающей сл. выборке $\vec{X}$, наз. отобр. $\hat{F} : R \to R$, определенное правилом:

$$\hat{F}(t) = \frac{n(t, \vec{X})}{n}.$$

Замечание: 1) при каждом фиксир. $t \in R$ $\hat{F}(t)$ явл. с/в, которая может принимать значения $0, 1/n, \dots, \frac{n-1}{n}, 1$.

2) Зафиксируем t_0. Обозн. $p = P\{X < t_0\} = const$. Тогда:

$P\{n(t_0, \vec{X}) = k\} = P\{$среди $X_1, \dots, X_n$ ровно k штук приняли значение $< t_0\} = C_n^k \cdot p^k \cdot q^{n-k}$.

Все X_i независ. в сов-ти и имеют одинаковое с X распр. $\Rightarrow$ $P\{$среди $\dots\} = P\{$в серии из n независ. исп. с/в X ровно k раз приняла значение $< t_0\}$. Если считать успехом событие $\{X$ приняла значение $< t_0\}$, то $P\{$в серии $\dots\} = P_n(k)$ - вер-ть реализации ровно k успехов в серии из n исп. по сх. Бернулли.

Из этого также следует, что $P\{\hat{F}(t_0) = \frac{k}{n}\} = P\{n(t_0, \vec{X}) = k\} = C_n^k p^k \cdot q^{n-k}$

Теорема: для любого фикс. $t_0 \in R$ посл-ть с/в

$$\hat F(t_0, \vec{X_1}),\ \hat F(t_0, \vec{X_2}),\ \dots,\ \hat F(t_0, \vec{X_n}),\ \dots$$

сходится по вероятности к значению в т. t_0 теоретической (т.е. "истинной") ф. распр. X, т.е.

$$\hat F(t_0, \vec{X_n}) \xrightarrow[n \to \infty]{P} F(t_0)$$

Выше мы обозн. сл. выборку $\vec{X} = (X_1, \dots, X_n)$ не включая объем выборки в обозначение в-ра.

$$\vec{X_1} = (X_1),\ \vec{X_2} = (X_1\ X_2),\ \dots,\ \vec{X_n} = (X_1\ X_2\ \dots\ X_n).$$

Док-во: При фикс. $t_0 \in R$ с/в $\hat F(t_0, \vec{X_n}) = \dfrac{n(t_0, \vec{X_n})}{n}$ равна отн. частоте успеха (т.е. события $\{X < t_0\}$) в серии из n исп. по сх. Б. В соотв. с ЗБЧ в форме Бернулли

$$\hat F(t_0, \vec{X_n}) \xrightarrow[n \to \infty]{P} p \text{ (теор. частота), но } p = P\{X < t_0\} = F(t_0)$$

10. **Понятия интервального статистического ряда, эмпирической плотности, гистограммы, полигона частот.**

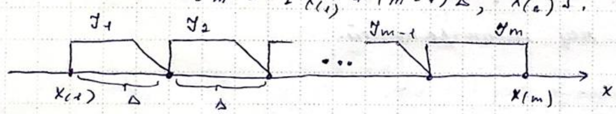
IV Интервальный стат. ряд

Выше было понятие стат. ряда. Однако если объем выборки достаточно велик ($n > 50$), то элементы выборки группируют в так называемый стат. ряд. Для этого отрезок $J = [x_{(1)}, x_{(n)}]$ разбивают на m равновеликих промежутков. Ширина каждого из них:

$$\Delta = \frac{|J|}{m} = \frac{x_{(n)} - x_{(1)}}{m}.$$

Далее полагают $J_i = [x_{(1)} + (i-1)\Delta;\ x_{(1)} + i \cdot \Delta)$, $i = \overline{1; m-1}$.

$$J_m = [x_{(1)} + (m-1)\Delta;\ x_{(n)}].$$



Опр: Интервальным стат. рядом, отвечающим выборке $\vec{x}$, наз. табл. вида

J_1	J_2	...	J_m
n_1	n_2	...	n_m

Здесь n_i - число эл-тов выборки $\vec{x}$, попавших в промежуток J_i.

Замечание: 1) $\sum\limits_{i=1}^{m} n_i = n$.

2) Для выбора m исп. формулу

$$m = \lfloor \log_2(n) \rfloor + 2 \quad \text{или} \quad m = \lfloor \log_2(n) \rfloor + 1$$

V Эмпирическая плотность

Пусть для данной выборки $\vec{x}$ построен инт. стат. ряд $(J_i;\ n_i)$, $i=\overline{1,m}$.

Опр.: Эмп. пл-ю распр. (соотв. выборке $\vec{x}$) наз. ф-ия:

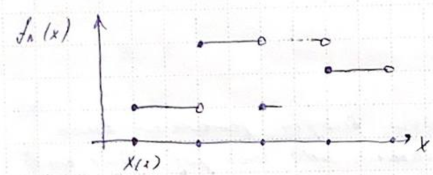
$$f_n(x)=\begin{cases}\dfrac{n_i}{n\cdot\Delta}, & \text{если } x\in J_i\\ 0 & \text{иначе}\end{cases}$$

Замечание: 1) Очевидно, что $\int\limits_{-\infty}^{+\infty} f_n(x)\,dx = \int\limits_{x_{(1)}}^{x_{(n)}} f_n(x)\,dx = \sum\limits_{i=1}^{m} \int\limits_{J_i} f_n(x)\,dx =$

$$= \sum_{i=1}^{m} \frac{n_i}{n\Delta}\cdot\Delta = \frac{1}{n}\cdot\sum_{i=1}^{m} n_i = \frac{n}{n} = 1.$$

Т.о. эмп. пл. удовл. условию нормировки. Легко показать, что она обладает всеми св-вами ф. пл-ти.

2) $f_n(x)$ яв. кусочно-постоянной ф.:



3) $f_n(x)$ яв. стат. аналогом ф-ии пл. Аналогично доказ. выше рез-ту для ф-ии распр.

$$\hat{F}_n(x) \xrightarrow[n\to\infty]{P} F(x)$$

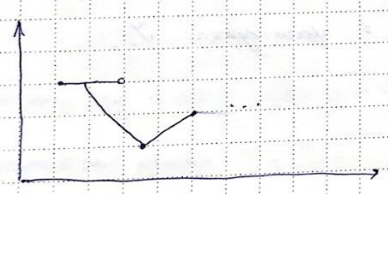
Можно пок., что $f_n(x)\approx f(x)$ при $n \gg 1$.

Опр.: График эмп. ф-ии наз. гистограммой.

VI Полигон частот

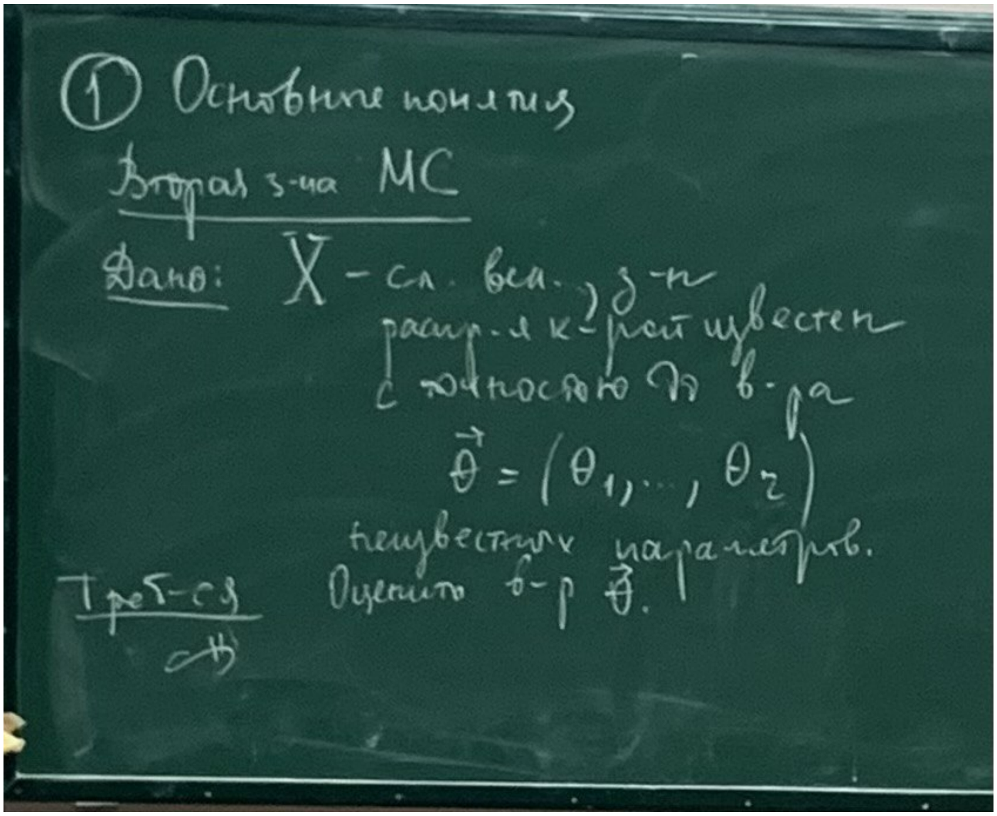
Пусть для нек. выборки $\vec{x}$ построена гистограмма.

Опр.: Полигоном частот наз. ломаная, звенья которой соед. середины верхних сторон соседних прямоугольников гистограммы.



11. **Постановка задачи идентификации неизвестных параметров закона распределения случайной величины. Определение точечной оценки. Определение несмещенной точечной оценки. Показать, что выборочная дисперсия является смещенной оценкой дисперсии. Записать формулу для исправленной выборочной дисперсии.**

Не уверена, что задача ид-ции именно такая, но похоже на правду.



Опр.: Точечной оценкой параметра θ наз. статистика $\hat{\theta}(\vec{X})$ (св), выборочное значение которой исп. в качестве значения параметра θ, т.е. принимается равенство $\theta = \hat{\theta}(\vec{x})$, $\vec{x}$ – выборка, числовой в-р, $\hat{\theta}$ – число, значение статистики $\hat{\theta}$ на выборке $\vec{x}$.

② Несмещенность точечной оценки

Пусть θ – неизв. пар-р з-на распр. СВ X. $\hat{\theta}(\vec{X})$ – точечная оценка для θ. Точечная оценка $\hat{\theta}$ параметра θ наз. несмещенной, если:

$$\exists\, M[\underbrace{\hat{\theta}}_{\text{СВ}}(\underbrace{\vec{X}}_{\text{сл. выборка}})] = \theta \leftarrow \text{теор. значение параметра}$$

Пример: X – ген. совокупность, $\exists\, DX = \sigma^2$.

$$\hat{\sigma}^2(\vec{X}) = \hat{m}_2^{\circ}(\vec{X}) = \underbrace{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}_{\text{выборочная дисперсия}}$$ – исп. в кач-ве точечной оценки для параметра σ^2.

Было доказано, что $M[\hat{\sigma}^2(\vec{X})] = \frac{n-1}{n}\sigma^2 \neq \sigma^2$, т.е. выборочная дисперсия яв. смещённой оценкой дисперсии.

Замечание: Рассм. статистику $S^2(\vec{X}) = \frac{n}{n-1}\hat{\sigma}^2(\vec{X}) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

Тогда $M[S^2(\vec{X})] = M[\frac{n}{n-1}\hat{\sigma}^2(\vec{X})] = \frac{n}{n-1}\cdot \underbrace{\frac{n-1}{n}\sigma^2}_{M[\hat{\sigma}^2]} = \sigma^2$, т.е.

$S^2(\vec{X})$ яв. несмещ. оценкой дисперсии.

Опр: Статистика $S^2(\vec{X})$ наз. исправленной выборочной дисперсией.

12. **Постановка задачи идентификации неизвестных параметров закона распределения случайной величины. Определение точечной оценки. Определение состоятельной оценки. Привести примеры состоятельной и несостоятельной оценок (с обоснованием).**

Ответ на первый и второй вопрос смотреть в ответах на вопрос 11.

3. Состоятельность точечной оценки

Пусть X – ген. совокупность; θ – параметр з-на распр. СВ X; $\hat{\theta}(\vec{X}_n)$ – точечная оценка параметра θ.

{ В настоящем пункте удобно вкл. n – объем выборки – в обозначение выборки: $\vec{X}_n = (X_1, \ldots, X_n)$ }

Опр: Точечная оценка $\hat{\theta}(\vec{X}_n)$ параметра θ наз. состоятельной, если:

$$\hat{\theta}(\vec{X}_n) \xrightarrow[n\to\infty]{P} \theta.$$

Замечание: 1) Условие из опр. сост. оценки экв. условию:

$$\forall \varepsilon > 0 \quad P\{|\hat{\theta}(\vec{X}_n) - \theta| \geq \varepsilon\} \xrightarrow[n\to\infty]{} 0$$

2) Если точечная оценка не яв. сост., то она никому не нужна.

Пример: Пусть X – ген. совокупность, $\theta = M(X)$, $\hat\theta_1(\vec X) = \bar X$ – точечная оценка. $\hat\theta_1$ яв. сост. для θ в предположении, что $\exists\, DX = \sigma^2 < \infty$.

1) Рассм. случ. выборку: $\vec X_n = (X_1, \dots, X_n)$.
При $n \to \infty$ получаем посл-ть $X_1, X_2, \dots$ независ. (т.к. это компоненты случ. выборки) СВ.

2) $X_i \sim X \Rightarrow \exists\, M(X_i) = \theta,\ \exists\, D(X_i) = \sigma^2$

3) Тогда согласно следствию ЗБЧ в форме Чебышева:

$$\underbrace{\frac{1}{n}\sum_{i=1}^{n} X_i}_{\bar X_n} \xrightarrow[n\to\infty]{P} \theta, \text{ т.е. } \hat\theta_1(\vec X_n) \xrightarrow[n\to\infty]{P} \theta$$

$\Updownarrow$ оценка $\hat\theta_1$ состоятельна.

29

Замечание: Сост-ть этой оценки можно док., не предполагая сущ. конечной дисперсии.

Пример: $X \sim N(\theta, \sigma^2)$ – ген. сов-ть, $\theta = M(X)$.

Покажем, что оценка $\hat\theta_4(\vec X_n) = X_1$ не яв. сост. для θ.

$$P\{|\hat\theta_4(\vec X_n) - \theta| \geq \varepsilon\} \xrightarrow[n\to\infty]{P} ?$$

$$P\{|X_1 - \theta| \geq \varepsilon\} = \{X_1 \sim X \sim N(\theta, \sigma^2)\} = 1 - P\{-\varepsilon \leq X - \theta \leq \varepsilon\} =$$

$$= 1 - P\{\theta - \varepsilon < X < \theta + \varepsilon\} = 1 - \left[\Phi_0\left(\frac{\theta + \varepsilon - \theta}{\sigma}\right) - \Phi_0\left(\frac{\theta - \varepsilon - \theta}{\sigma}\right)\right] =$$

$$= 1 + \left[\Phi_0\left(\frac{\varepsilon}{\sigma}\right) + \Phi_0\left(\frac{\varepsilon}{\sigma}\right)\right] = 1 - 2\cdot\Phi_0\left(\frac{\varepsilon}{\sigma}\right) = 0 \Leftrightarrow \Phi_0\left(\frac{\varepsilon}{\sigma}\right) = \frac{1}{2} \Leftrightarrow \frac{\varepsilon}{\sigma} = +\infty$$

Т.о. $\neg\left(\hat\theta_4(\vec X_n) \xrightarrow[n\to\infty]{P} \theta\right) \Rightarrow \hat\theta_4$ не яв. состоят. для θ

Замечание: 1) Можно док., что испр. выборочная дисперсия $S^2(\vec X_n)$ яв. состоят. оценкой дисперсии. Также можно док., что $\hat\sigma^2(\vec X_n)$ – тоже состоят. оценка дисперсии.

$$M[\hat\sigma^2(\vec X_n)] = \frac{n-1}{n}\sigma^2.$$

2) Можно док. более общий результат: выборочные моменты

$$\hat m_k(\vec X_n) = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

$$\hat{\mathring{m}}_k(\vec X_n) = \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar X)^k$$

также яв. состоят. оценками своих теор. аналогов при условии, что последние сущ-ют.

3) Можно пок., что при $k \geq 2$ выборочные моменты яв. смещенными оценками своих теор. аналогов.

13. **Постановка задачи идентификации неизвестных параметров закона распределения случайной величины. Определение точечной оценки. Определение эффективной оценки. Показать, что выборочное среднее является эффективной оценкой математического ожидания в классе линейных оценок.**

Ответ на первый и второй вопрос смотреть в ответах на вопрос 11.

(4) **Эффективность точечной оценки**

Пусть

1) X - СВ, з-н распр. которой зависит от неизв. параметра $\theta \in \mathbb{R}$.

2) $\hat{\theta}_1(\vec{X})$, $\hat{\theta}_2(\vec{X})$ - две несмещенные точечные оценки для θ.

3) $\exists D\hat{\theta}_1$, $\exists D\hat{\theta}_2$.

Опр: Оценка $\hat{\theta}_1$ наз. более эфф. оценкой параметра θ, чем оценка $\hat{\theta}_2$, если $D\hat{\theta}_1 < D\hat{\theta}_2$.

Опр: Оценка $\hat{\theta}$ наз. эфф. оценкой параметра θ, если:

1) $\hat{\theta}$ явл. несмещенной оценкой для θ;
2) $\hat{\theta}$ обладает наименьшей дисперсией среди всех несмещенных точечных оценок пар-ра θ.

Замечание: Иногда говорят не об эфф. "вообще" точечной оценке, а об оценке, эфф. в некотором классе.

33

Пусть $\textcircled{H}$ - нек. класс несмещённых оценок θ.

Опр.: Оценка $\hat{\theta} \in \textcircled{H}$ наз. эфф. в классе $\textcircled{H}$, если она имеет наим. дисперсию среди всех оценок этого класса, т.е.:

$$(\forall \tilde{\theta} \in \textcircled{H})(D\hat{\theta} \le D\tilde{\theta}).$$

Пример: Пусть X - СВ, $\exists MX = m$, $\exists DX = \sigma^2$.
Покажем, что оценка $\hat{m}_1(\vec{X}) = \bar{X}$ явл. эфф. оценкой для m в классе линейных оценок.

Решение: 1) Лин. оценка имеет вид:

$$\hat{m}(\vec{X}) = \sum_{i=1}^{n} \lambda_i X_i = \lambda_1 X_1 + \ldots + \lambda_n X_n, \text{ где } \lambda_i \in R. \quad (*)$$

Тогда мат. ожидание лин. оценки $(*)$:

а) $M[\hat{m}] = \lambda_1 MX_1 + \ldots + \lambda_n MX_n = \{X_i \sim X;\ MX = m\} = (\lambda_1 + \ldots + \lambda_n) \cdot m$

Т.к. оценка д.б. несмещённой, то $M[\hat{m}] = m \Rightarrow \sum_{i=1}^{n} \lambda_i = 1$.

б) Дисперсия оценки $(*)$:

$$D[\hat{m}] = \sum_{i=1}^{n} \lambda_i^2 DX_i = \{X_i \sim X;\ DX_i = DX = \sigma^2\} = \sigma^2 \cdot \sum_{i=1}^{n} \lambda_i^2$$

2) Попробуем подобрать константы λ_i в $(*)$ т. ч.:

$$\begin{cases} D[\hat{m}] \longrightarrow \min \\ \sum_{i=1}^{n} \lambda_i = 1 \end{cases}$$

Для этого нужно решить задачу условной оптимизации:

$$\varphi(\lambda_1, \ldots, \lambda_n) = \lambda_1^2 + \ldots + \lambda_n^2.$$

[Общий вид: $f(x_1, \ldots) \to \min$
$\varphi_1(x_1, \ldots, x_n) = 0$
$\ldots$
$\varphi_m(x_1, \ldots) = 0$]

Запишем функцию Лагранжа:

$$L(\lambda_1, \ldots, \lambda_n, \mu) = \lambda_1^2 + \ldots + \lambda_n^2 - \mu\left(\sum_{i=1}^{n} \lambda_i - 1\right) = f - \varphi$$

Необх. условие extremum'а:

$$\begin{cases} \dfrac{\partial L}{\partial \lambda_i} = 0 & \dfrac{\partial L}{\partial \lambda_i} = 2\lambda_i - \mu = 0 \\ \dfrac{\partial L}{\partial \mu} = 0 & \dfrac{\partial L}{\partial \mu} = -\left(\sum_{i=1}^{n} \lambda_i - 1\right) = 0 \end{cases}$$

Из первых уравнений $\Rightarrow \lambda_i = \dfrac{\mu}{2}$, $i = \overline{1,n}$
Подставим в последнее ур-е: $\sum_{i=1}^{n} \frac{\mu}{2} = 1 \Rightarrow \dfrac{\mu n}{2} = 1 \Rightarrow \mu = \dfrac{2}{n}$.
Т.о. $\lambda_i = \mu/2 = 1/n$

31

Можно показать, что найденные функция

$(\lambda_1, \ldots, \lambda_n, \mu) = \left(\frac{1}{n}, \ldots, \frac{1}{n}, \frac{2}{n}\right)$ соответствует усл. min целевой ф-ии.

Т.о., подставляя λ_i в (+), получаем искомую оценку для МО с миним. дисперсией в классе лин. оценок:

$$\hat{m}(\vec{X}) = \frac{1}{n} X_1 + \ldots + \frac{1}{n} X_n = \overline{X}.$$

Дисперсия этой оценки:

$$D[\hat{m}] = \sigma^2 \sum_{i=1}^{n} \lambda_i^2 = \sigma^2 \sum_{i=1}^{n} \frac{1}{n^2} = \frac{\sigma^2}{n}.$$

14. **Постановка задачи идентификации неизвестных параметров закона распределения случайной величины. Определение точечной оценки. Определение эффективной оценки. Сформулировать теорему о единственности эффективной оценки.**

    Ответ на первый и второй вопрос смотреть в ответах на вопрос 11.

    Ответ на третий вопрос в ответах на вопрос 13.

Теорема: (о единственности эфф. оценки)

Пусть 1) $\hat{\theta}_1(\vec{X})$ и $\hat{\theta}_2(\vec{X})$ – две эфф. ($\Rightarrow$ несмещенные) оценки пар-ра θ.

Тогда $\hat{\theta}_1(\vec{X}) = \hat{\theta}_2(\vec{X})$.

Без док-ва.

15. **Постановка задачи идентификации неизвестных параметров закона распределения случайной величины. Определение точечной оценки. Определение эффективной оценки. Определение количества информации по Фишеру. Сформулировать теорему о неравенстве Рао-Крамера.**

    Ответ на первый и второй вопрос смотреть в ответах на вопрос 11.

    Ответ на третий вопрос в ответах на вопрос 13.

(5) Неравенство Рао-Крамера

Пусть 1) X - СВ, з-н распр. которой зависит от в-ра $\vec{\theta} = (\theta_1, \ldots, \theta_r)$ параметров.

2) $\vec{X} = (X_1, \ldots, X_n)$ - сл. выборка из ген. совокупности X.

Опр.: Функцией правдоподобия, отвеч. сл. выборке $\vec{X}$, наз. ф-ию:

$$L(\vec{X}, \vec{\theta}) = p(X_1, \vec{\theta}) \cdot \ldots \cdot p(X_n, \vec{\theta}),$$

где $p(X_i, \vec{\theta}) = \begin{cases} f(X_i, \vec{\theta}), & \text{если } X \text{ - НСВ}, \\ P\{X = X_i\}, & \text{если } X \text{ - ДСВ} \end{cases}$

(здесь f - ф-ия пл-ти распр. вер. СВ X)

Пусть $r = 1$, т.е. $\vec{\theta} = (\theta_1) = (\theta)$

Опр.: Кол-во инф-ии по Фишеру, содерж. в сл. выборке $\vec{X}$, наз. число

$$I(\theta) = M\left[\left(\frac{\partial \ln L}{\partial \theta}\right)^2\right].$$

Теорема: (неравенство Рао-Крамера)

Пусть 1) рассм. модель явл. регулярной;

2) $\hat{\theta}(\vec{X})$ - несмещ. оценка θ.

35 Тогда: $D[\hat{\theta}] \geq \dfrac{1}{I(\theta)}$

б/д.

16. **Постановка задачи идентификации неизвестных параметров закона распределения случайной величины. Определение точечной оценки. Определение эффективной оценки. Показать, что выборочное среднее является эффективной оценкой математического ожидания нормальной случайной величины при известной дисперсии**

Ответ на первый и второй вопрос смотреть в ответах на вопрос 11.

Ответ на третий вопрос в ответах на вопрос 13.

На последний вопрос не нашла пока ответ

17. **Постановка задачи идентификации неизвестных параметров закона распределения случайной величины. Определение точечной оценки. Описать метод моментов построения точечной оценки. Привести пример.**

Ответ на первый и второй вопрос смотреть в ответах на вопрос 11.

I Метод моментов

Пусть 1) X - СВ, з-н распр. которой зависит от вектора $\vec{\theta} = (\theta_1, \theta_2, \dots, \theta_r)$ неизв. параметров

2) $\exists\, r$ первых моментов СВ X, т.е. $\exists\, M[X^k]$, $k = 1 \dots r$.

Тогда в методе моментов:

1) Вычисляются теоретические моменты 1-го, 2-го, .., r-ого порядка, которые зависят от неизвестных параметров

$$m_k(\theta_1, \dots, \theta_r) = \underbrace{M[X^k]}_{\text{теор. момент порядка } k}, \quad k = 1 \dots r$$

2) Теор. моменты приравниваются к их выборочным аналогам.

$$(*) \begin{cases} m_1(\theta_1, \dots, \theta_r) = \hat{m}_1(\vec{X}) \\ \dots \\ m_r(\theta_1, \dots, \theta_r) = \hat{m}_r(\vec{X}) \end{cases}, \quad \text{где} \quad \hat{m}_k(\vec{X}) = \frac{1}{r}\sum_{i=1}^{r} X_i^k$$

Система уравнений (в общем случае нелинейных) относительно независ. параметров $\theta_1 \dots \theta_r$.

40

3) Решаем полученную систему: Полученные зависимости исп. в качестве точечных оценок для неизв. парам.-ов

$$\begin{cases} \theta_1 = \hat{\theta}_1(\vec{x}) \\ \dots \\ \theta_r = \hat{\theta}_r(\vec{x}) \end{cases}$$

Замечание: иногда некоторое уравнение системы (*) удобнее записывать отн-но центральных, а не начальных моментов. В этом случае k-ое уравнение будет иметь вид:

$$\overset{\circ}{m}_k(\theta_1, \dots, \theta_r) = \hat{\overset{\circ}{m}}_k(\vec{x}), \text{ где } \overset{\circ}{m}_k(\theta_1, \dots, \theta_r) = M[(X - MX)^k],$$

$$\hat{\overset{\circ}{m}}_k(\vec{x}) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k.$$

Пример: Пусть СВ $X \sim R[a; b]$; a, b - неизвестные. С исп. метода моментов построить точечные оценки для a и b.

1) $X \sim R[a; b] \Rightarrow f_X(x) = \begin{cases} \frac{1}{b-a}, & x \in [a; b] \\ 0, & \text{иначе} \end{cases}$ (плотность)

2 параметра $\Rightarrow r = 2$

$$\begin{cases} m_1(a, b) = \hat{m}_1(\vec{x}) & \text{- отн. нач. момента } 1^{\text{го}} \text{ порядка} \\ \overset{\circ}{m}_2(a, b) = \hat{\overset{\circ}{m}}_2(\vec{x}) & \text{- отн. центрального момента } 2^{\text{го}} \text{ порядка} \end{cases}$$

2) Найдем теоретические моменты:

$$m_1(a, b) = MX = \frac{a+b}{2}; \quad \overset{\circ}{m}_2(a, b) = DX = \frac{(b-a)^2}{12}$$

Запишем выборочные моменты:

$$\hat{m}_1(a, b) = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}; \quad \hat{\overset{\circ}{m}}_2(a, b) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sigma^2(\vec{x})$$

$$\hookrightarrow = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = S^2(\vec{x})$$

3) Приходим к системе уравнений:

$$\begin{cases} \frac{a+b}{2} = \bar{x} \\ \frac{(b-a)^2}{12} = S^2 \end{cases} \Rightarrow \begin{cases} a+b = 2\bar{x} \\ (b-a) = \underset{\uparrow +, \text{ т.к. } b > a}{\pm} 2\sqrt{3}S \end{cases} \Rightarrow \begin{cases} 2b = 2\bar{x} + 2\sqrt{3}S \\ 2a = 2\bar{x} - 2\sqrt{3}S \end{cases} \Rightarrow \begin{cases} a = \bar{x} - S\sqrt{3} \\ b = \bar{x} + S\sqrt{3} \end{cases}$$

Ответ: $a = \bar{x} - S\sqrt{3}$; $b = \bar{x} + S\sqrt{3}$.

18. **Постановка задачи идентификации неизвестных параметров закона распределения случайной величины. Определение точечной оценки. Описать метод максимального правдоподобия построения точечной оценки. Привести пример.**

Ответ на первый и второй вопрос смотреть в ответах на вопрос 11.

II Метод максимального правдоподобия

Пусть 1) X – СВ, з-н распр. которой зависит от вектора параметров $\vec{\theta} = (\theta_1, \ldots, \theta_r)$. Ранее было введено понятие функции правдоподобия случайной выборки $\vec{x}$:

$$L(\vec{X}, \vec{\theta}) = p(X_1, \vec{\theta}) \cdot \ldots \cdot p(X_n, \vec{\theta}), \text{ где}$$

$$p(X_i; \vec{\theta}) = \begin{cases} f(X_i, \vec{\theta}), & \text{НСВ} \\ P(X_i; \vec{\theta}), & \text{ДСВ} \end{cases}$$

Можно показать, что чем ближе значение вектора $\vec{\theta}$ к теор. значению вектора $\vec{\theta}$, тем большее значение принимает ф-ия $L(\vec{x}, \vec{\theta})$.
В этом методе в качестве точечных оценок неизв. параметров исп. значения, доставляющие max значение ф-ии правдоподобия. Для реализации метода необх. решить задачу $L(\vec{x}, \vec{\theta}) \to \max\limits_{\vec{\theta}}$. (*) Тогда $\vec{\theta}(\vec{x}) = \arg\max\limits_{\vec{\theta}} L(\vec{x}, \vec{\theta})$.

Замечания: 1) Для решения задачи (*) можно исп. необх. условие экстремума диф-емой ф-ии нескольких переменных:

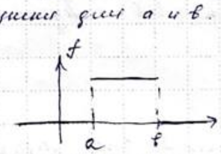
$$\begin{cases} \dfrac{\partial L(\vec{x}, \vec{\theta})}{\partial \theta_1} = 0 \\ \ldots \\ \dfrac{\partial L(\vec{x}, \vec{\theta})}{\partial \theta_r} = 0 \end{cases} \text{ – уравнения правдоподобия}$$

2) Функция L – произведение n сомножителей, работать с ней не всегда удобно, поэтому часто вместо задачи (*) решают эквив. задачу:

$$\ln(L) \to \max. \text{ Тогда: } \begin{cases} \dfrac{\partial \ln(L(\vec{x}, \vec{\theta}))}{\partial \theta_1} = 0 \\ \dfrac{\partial \ln(L(\vec{x}, \vec{\theta}))}{\partial \theta_r} = 0 \end{cases}$$

Пример: $X \sim R[a; b]$, где a и b – неизв. параметры. С исп. метода макс. правдоподобия построить точечные оценки для a и b.

Решение: 1) $f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a; b] \\ 0, & \text{иначе} \end{cases}$



$$L(\vec{x}, a, b) = \{X - \text{НСВ}\} = f(x_1, a, b) \cdot \ldots \cdot f(x_n, a, b) \overset{(***)}{=} \frac{1}{b-a} \cdot \ldots \cdot \frac{1}{b-a} = \frac{1}{(b-a)^n}$$

ошибка

41

2) $\ln L = \ln \frac{1}{(b-a)^n} = -n \ln (b-a) \longrightarrow \max\limits_{(a,b)}$

Уравнения правдоподобия:

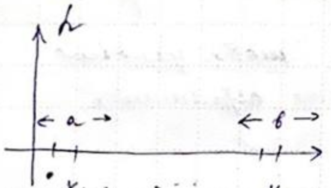
$$\begin{cases} \dfrac{\partial \ln L}{\partial a} = -n \cdot \dfrac{-1}{b-a} = \dfrac{n}{b-a} = 0 \\ \dfrac{\partial \ln L}{\partial b} = -\dfrac{n}{b-a} = 0 \end{cases} \Rightarrow n = 0 \ (?) \ \text{Противоречие.}$$

3) Противоречие возникло из-за того, что неверно записана формула (***)

Правильно: $L(\vec{x}, a, b) = \begin{cases} \dfrac{1}{(b-a)^n}, \text{ если } \begin{cases} X_{(1)} \geq a \\ X_{(n)} \leq b \end{cases} \\ 0 \text{ иначе} \end{cases}$

Вычисление $\dfrac{\partial L(\vec{x}, a, b)}{\partial a}$ и $\dfrac{\partial L(\vec{x}, a, b)}{\partial b}$ затруднительно, т.к. от a и b также зависят границы области, в которой $L > 0$.

Попробуем "в лоб" найти $\max\limits_{a,b} L(\vec{x}, a, b)$.



а) если $\begin{cases} a \leq X_{(1)} \\ b \geq X_{(n)} \end{cases}$, то $L > 0$, иначе $L = 0$

т.к. $L \to \max$, то эти условия д.б. выполнены.

б) при выполнении этих условий $L = \dfrac{1}{(b-a)^n}$.

Т.к. $L \to \max$, то $(b-a) \to \min$. Т.о. приходим к задаче:

$$\begin{cases} b - a \to \min \\ a \leq X_{(1)} \\ b \geq X_{(n)} \end{cases}$$

Ответ: $\hat{a}(\vec{x}) = X_{(1)}$; $\hat{b}(\vec{x}) = X_{(n)}$.

19. **Постановка задачи идентификации неизвестных параметров закона распределения случайной величины. Сформулировать определение γ-доверительного интервала. Сформулировать определение центральной статистики и изложить общий алгоритм построения γ-доверительного интервала для скалярного параметра.**

Ответ на первый вопрос смотреть в ответах на вопрос 11.

Опр: Интервальной оценкой параметра θ с уровнем (уровня) γ (γ-интерв. оценкой) наз. пару статистик $\underline{\theta}(\vec{x})$ и $\overline{\theta}(\vec{x})$, таких, что $P\{\theta \in (\underline{\theta}(\vec{x}); \overline{\theta}(\vec{x}))\} = \gamma$.

Опр.: γ-доверительным интервалом (ДИ уровня γ) для параметра θ наз. реализацию (выборочное значение) интервальной оценки уровня γ для этого параметра, т.е. интервал: $(\underline{\theta}(\vec{x}); \overline{\theta}(\vec{x}))$ с детерминированными границами.

Общий алгоритм построения инт. оценки с исп. центральной статистики:

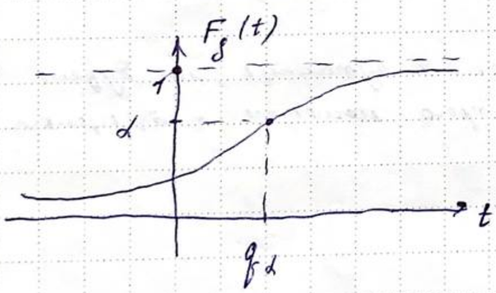
Пусть 1) $g(\vec{X}; \theta)$ - центральная статистика

2) $g(\vec{X}, \theta)$ как ф-ия параметра θ явл. монотонно возр.

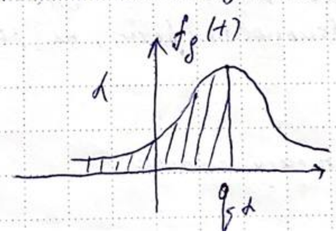
3) $F_g(t)$ явл. монотонно возр.

4) Выбраны некот. $\alpha_1, \alpha_2 \geq 0$ такие, что $\alpha_1 + \alpha_2 = 1 - \gamma$.

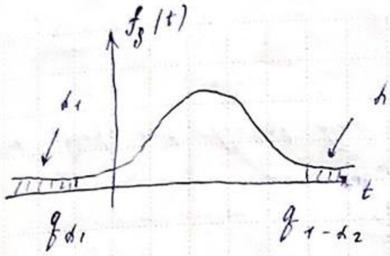
Замечание: из 3) $\Rightarrow$ ур-е $F_g(t) = \alpha$ для $\forall \alpha \in (0; 1)$ имеет единств. решение $t = q_\alpha$ (квантиль уровня α СВ g).



При этом СВ g:



$\Leftrightarrow P\{X < q_\alpha\} = \alpha$



$\alpha_1 + \alpha_2 + \gamma = 1$

Из свойств ЦСВ: $P\{q_{\alpha_1} < g(\vec{X}, \theta) < q_{1-\alpha_2}\} = \gamma$

Запишем событие, стоящее под вер-ю в экв. виде:

$\{q_{\alpha_1} < g(\vec{X}, \theta) < q_{1-\alpha_2}\} \Leftrightarrow$ {Выражаем из двойного неравенства θ, с учетом условия 2) знаки нер-в не меняются} $\Leftrightarrow$

$\Leftrightarrow \{g^{-1}(\vec{X}, q_{\alpha_1}) < \theta < g^{-1}(\vec{X}, q_{1-\alpha_2})\}$

Обозн.: $\underline{\theta}(\vec{X}) = g^{-1}(\vec{X}, q_{\alpha_1})$; $\overline{\theta}(\vec{X}) = g^{-1}(\vec{X}, q_{1-\alpha_2})$

Т.к. соб-я экв., то:

$P\{\underline{\theta}(\vec{X}) < \theta < \overline{\theta}(\vec{X})\} = \gamma \Rightarrow \underline{\theta}(\vec{X}), \overline{\theta}(\vec{X})$ – γ-дов. интервал для θ.

Замечание:

1) Построенная инт. оценка зависит от выбора α_1 и α_2. Обычно исп. значения $\alpha_1 = \alpha_2 = \frac{1-\gamma}{2}$. Однако это не всегда так.

2) Например, если взять $\alpha_1 = 1-\gamma$, $\alpha_2 = 0$, то $q_{1-\alpha_2} = +\infty \Rightarrow$
$\underline{\theta}(\vec{X}) = g^{-1}(\vec{X}, q_{1-\gamma})$; $\overline{\theta}(\vec{X}) = +\infty$

Т.о. $\underline{\theta}(\vec{X})$ – нижняя γ-дов. граница для θ.

3) Аналогично при $\alpha_1 = 0$, $\alpha_2 = 1-\gamma$ получим $\underline{\theta}(\vec{X}) = -\infty$
$\overline{\theta}(\vec{X}) = g^{-1}(\vec{X}, q_{1-(1-\gamma)}) =$
$= g^{-1}(\vec{X}, q_\gamma)$ – верхняя граница для θ.

20. **Постановка задачи идентификации неизвестных параметров закона распределения случайной величины. Сформулировать определение γ-доверительного интервала. Сформулировать определение центральной статистики. Изложить и обосновать метод построения доверительного интервала для математического ожидания нормальной случайной величины в случае известной дисперсии**

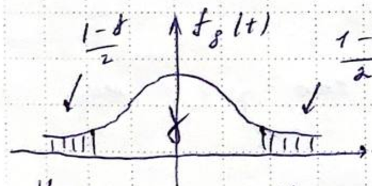
Ответ на первый вопрос смотреть в ответах на вопрос 11.

Ответ на второй и третий вопрос смотреть в ответах на вопрос 19.

Пример: Пусть $X \sim N(m, \sigma^2)$, где m – неизв, σ^2 – изв.
Построить γ-доверит. инт. оценку для m.

Решение: Рассм. статистику $g(\vec{X}, m) = \frac{m - \bar{X}}{\sigma} \cdot \sqrt{n} \sim N(0,1)$.

Исп. $\alpha_1 = \alpha_2 = \frac{1-\gamma}{2}$.



$u_{\frac{1-\gamma}{2}}$ $\quad$ $u_{1-\frac{1-\gamma}{2}}$

$\Vert$

$-u_{\frac{1+\gamma}{2}}$ т.к. $f_g(t)$ – четная

Т.о.:

$$P\left\{-u_{\frac{1+\gamma}{2}} < \underbrace{\frac{m - \bar{X}}{\sigma}\sqrt{n}}_{g(\vec{X}, m)} < u_{\frac{1+\gamma}{2}}\right\} = \gamma$$

$$P\left\{\bar{X} - \frac{\sigma \cdot u_{\frac{1+\gamma}{2}}}{\sqrt{n}} < m < \bar{X} + \frac{\sigma u_{\frac{1+\gamma}{2}}}{\sqrt{n}}\right\} = \gamma$$

21. **Постановка задачи идентификации неизвестных параметров закона распределения случайной величины. Сформулировать определение γ-доверительного интервала. Сформулировать определение центральной статистики. Изложить и обосновать метод построения доверительного интервала для математического ожидания нормальной случайной величины в случае неизвестной дисперсии**

Ответ на первый вопрос смотреть в ответах на вопрос 11.

Ответ на второй и третий вопрос смотреть в ответах на вопрос 19.

Пример: $X \sim N(m, \sigma^2)$, где m и σ^2 неизв. Построить инт. оценку для m.

Решение: Рассм. статистику $g(\vec{X}, m) = \dfrac{m - \bar{X}}{S(\vec{X})}\sqrt{n} \sim ?$

а) Представим g в виде $g(\vec{X}, m) = \dfrac{\frac{m-\bar{X}}{\sigma}\cdot\sqrt{n}}{\frac{S(\vec{X})}{\sigma}} = \dfrac{\frac{m-\bar{X}}{\sigma}\cdot\sqrt{n}\cdot\sqrt{n-1}}{\sqrt{\frac{(n-1)S^2(\vec{X})}{\sigma^2}}} =$

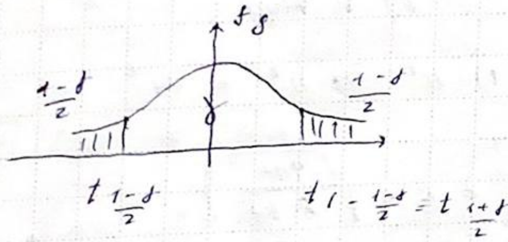
$$= \left\{\begin{array}{l} \text{Обозн. } \xi = \frac{m-\bar{X}}{\sigma}\cdot\sqrt{n} \\ \eta = \frac{(n-1)S^2(\vec{X})}{\sigma^2} \end{array}\right\} = \frac{\xi}{\eta}\cdot\sqrt{n-1}$$

б) $\xi \sim N(0,1)$ Можно показать, что $\eta \sim \chi^2(n-1)$ и что ξ и η нез.

Т.о. $g(\vec{X}, m) \sim \underbrace{St(n-1)}$

не зависит от $m \Rightarrow$ центр. статистика

Положим $\alpha_1 = \alpha_2 = \dfrac{1-\gamma}{2}$



$$P\left\{-t_{\frac{1+\gamma}{2}} < \frac{m-\bar{X}}{S(\vec{X})}\sqrt{n} < t_{\frac{1+\gamma}{2}}\right\} = \gamma$$

t_α — квантиль

22. **Постановка задачи идентификации неизвестных параметров закона распределения случайной величины. Сформулировать определение γ-доверительного интервала. Сформулировать определение центральной статистики. Изложить и обосновать метод построения доверительного интервала для дисперсии нормальной случайной величины.**

Ответ на первый вопрос смотреть в ответах на вопрос 11.

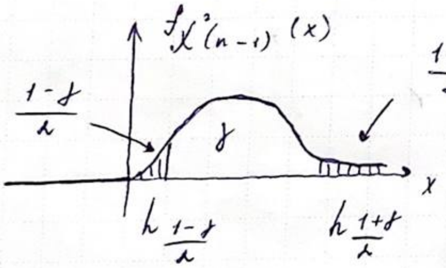
Ответ на второй и третий вопрос смотреть в ответах на вопрос 19.

Пример: Пусть СВ $X \sim N(m, \sigma^2)$, где σ^2 - неизв.; m - неизв/изв.
Построить интерв. оценку для σ^2.

Рассм. статистику: $g(\vec{X}, \sigma^2) = \dfrac{(n-1)\cdot S^2(\vec{X})}{\sigma^2}$

Можно показать, что $g \sim \chi^2(n-1)$.

Т.о., принимая $\alpha_1 = \alpha_2 = \dfrac{1-\gamma}{2}$, получаем:



$$P\left\{h_{\frac{1-\gamma}{2}} < g(\vec{X}, \sigma^2) < h_{\frac{1+\gamma}{2}}\right\} = \gamma$$

$$P\left\{\frac{1}{h_{\frac{1-\gamma}{2}}} > \frac{\sigma^2}{(n-1)\cdot S^2(\vec{X})} > \frac{1}{h_{\frac{1+\gamma}{2}}}\right\} = \gamma$$

$$P\left\{\frac{(n-1)S^2(\vec{X})}{h_{\frac{1+\gamma}{2}}} < \sigma^2 < \frac{(n-1)S^2(\vec{X})}{h_{\frac{1-\gamma}{2}}}\right\} = \gamma$$

Т.о.: $\underline{\sigma}^2(\vec{X}) = \dfrac{(n-1)S^2(\vec{X})}{h_{\frac{1+\gamma}{2}}}$

$\overline{\sigma}^2(\vec{X}) = \dfrac{(n-1)\cdot S^2(\vec{X})}{h_{\frac{1-\gamma}{2}}}$, где $h_\gamma = h_\gamma^{(n-1)}$ - квантиль уровня γ распр-я $\chi^2(n-1)$.