Мат. Стат. РК1 Теория

① Сформулировать и доказать первое неравенство Чебышева.

Теорема (1-ое нерав. Чебышева)

Пусть: 1) X – случ. величина;

2) $X \geq 0$ (т.е. $P\{X < 0\} = 0$);

3) $\exists\, MX$

Тогда:

$$\forall \varepsilon > 0: \; P\{X \geq \varepsilon\} \leq \frac{MX}{\varepsilon}$$

Доказ.

(для непрерывной случ. вел.)

$$MX = \int_{-\infty}^{+\infty} x f_X(x)\,dx = \{X \geq 0\} = \int_0^{+\infty} x f_X(x)\,dx =$$

$$= \left\{\begin{array}{l}\text{св-во}\\ \text{аддитивн.}\\ \text{интеграла}\end{array}\right\} = \int_0^{\varepsilon} \underbrace{x}_{\geq 0}\, \underbrace{f_X(x)}_{\geq 0}\,dx + \int_{\varepsilon}^{+\infty} x f_X(x)\,dx \geq$$

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

$$\geq \int_{\varepsilon}^{+\infty} \varepsilon f_X(x)\,dx = \varepsilon \int_{\varepsilon}^{+\infty} f_X(x)\,dx = \varepsilon \cdot P\{X \geq \varepsilon\}$$

Т.е. $MX \geq \varepsilon P\{X \geq \varepsilon\} \Rightarrow P\{X \geq \varepsilon\} \leq \dfrac{MX}{\varepsilon}$

ч.т.д.

② Сформулировать и доказать второе неравенство Чебышева.

<u>Теорема</u> (2-ое нерав. Чебышева)

Пусть: 1) X – случ. велич.;

2) $\exists MX$, $\exists DX$.

Тогда:

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

<u>Доказат.</u>

1) $DX = M[(X - MX)^2] =$

= { Рассм. случ. велич. $Y = (X - MX)^2 \geq 0$

Применим к Y 1-ое нерав. Чебышева

$P\{Y \geq \varepsilon_1\} \leq \frac{MY}{\varepsilon_1}$

<u>для $\varepsilon_1 = \varepsilon^2$</u>

$MY \geq \varepsilon_1 \cdot P\{Y \geq \varepsilon_1\}$ } $\geq$

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

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

Т.о. $DX \geq \varepsilon^2 \cdot P\{|X - MX| \geq \varepsilon\} \Rightarrow$

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

ч.т.д.

③ Для последовательности случайных величин сформулир. опред. сходимости по вероятности и слабой сходимости. Сформулировать закон больших чисел.

Пусть $X_1, X_2, \dots$ — последов. случ. вел.

<u>Опр.</u> Говорят, что последов. случ. величин $X_1, X_2, \dots$ сходится по вероятности к случ. величине Z, если:

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

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

<u>Опр.</u> Говорят, что последов. случ. величин $X_1, X_2, \dots$ слабо сходится к случ. величине Z, если $\forall x \in R$ такого, что F_Z непрерывна в x, выполняется:

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

(Здесь F_Z — функция распредел. случ. величины Z

$F_{X_n}, n \in N$ — функция распредел. случ. величины X_n)

Закон больших чисел (ЗБЧ)

Пусть $X_1, X_2, \ldots$ — послед. случ. велич.

$$\exists MX_i = m_i, \; i \in N$$

V

Опр. Говорят, что последов. $X_1, X_2, \ldots$ удовлетв. закону больших чисел, если:

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

---

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

ЗБЧ V

Теорема (ЗБЧ в форме Чебышева)

Пусть: 1) $X_1, X_2, \ldots$ — последов. независ. случ. вел.

2) $\exists MX_i = m_i, \; \exists DX_i = \sigma_i^2, \; i \in N$

3) Дисперсия случ. велич. $X_1, X_2, \ldots$ ограничена в совокупности, т.е.

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

Тогда последовательность $X_1, X_2, \ldots$ удовлетворяет ЗБЧ.

Док-во

1) Рассмотрим случ. вел. $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, n \in N$

Применим к случ. велич. $\overline{X}_n$ 2-ое нер-во Чебышева:

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

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

$$D[\overline{X}_n] = D[\frac{1}{n}\sum_{i=1}^{n} X_i] = \frac{1}{n^2} D[\sum_{i=1}^{n} X_i] =$$

$$= \left\{ \begin{matrix} X_i, i \in N - \\ \text{независим.} \end{matrix} \right\} = \frac{1}{n^2} \cdot \sum_{i=1}^{n} DX_i = \frac{1}{n^2}\sum_{i=1}^{n} \sigma_i^2 \leq$$

$$\leq \{\sigma_i^2 \leq C\} \leq \frac{1}{n^2}\sum_{i=1}^{n} C = \frac{nC}{n^2} = \frac{C}{n}$$

т.о.

$$\forall \varepsilon > 0 \quad 0 \leq P\{|\overline{X}_n - \frac{1}{n}\sum_{i=1}^{n} m_i| \geq \varepsilon\} \leq \frac{D\overline{X}_n}{\varepsilon^2} \leq \frac{C}{n\varepsilon^2}$$

Устремим $n \to \infty$.

Тогда по теореме о двух милиционерах

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

$\Rightarrow$ последов. $X_1, X_2, \ldots$ удовл. ЗБЧ.

ч.т.д.

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

<u>Теорема</u> (ЗБЧ в форме Чебышева)

Пусть 1) $X_1, X_2, \ldots$ последов. незав. случ. величин

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

3) Дисперсии случ. величин $X_1, X_2 \ldots$ ограничена в совокупности, т.е. $(\exists C > 0)(\sigma_i^2 \leq C,\ i \in N)$

Тогда последовательность $X_1, X_2 \ldots$ удовлетворяет ЗБЧ.

<u>Следствие 1</u> (т. Чебышева)

Пусть 1) $X_1, X_2, \ldots$ – последов. независ. случ. вел.

2) Все X_i, $i \in N$, одинаково распределены

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

Тогда:

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

Док-во

1) Условия 1) и 2) теоремы Чебышева выполн., условие 3) также выполн., т.к. $\sigma_i^2 = \sigma^2 \leq C$

Поэтому последоват. $X_1, X_2, \ldots$ удовл. ЗБЧ:

$$\forall \varepsilon > 0 \quad P\Big\{\Big|\overline{X}_n - \underbrace{\frac{1}{n}\sum_{i=1}^{n} m}_{m}\Big| \geq \varepsilon\Big\} \xrightarrow[n\to\infty]{} 0 \iff$$

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

Следствие 2 (ЗБЧ в форме Бернулли)

Пусть:

1) проводится n исп. по схеме исп. Бернулли, вер. успеха p в отдельном испытании

2) $r_n = \dfrac{\{\text{число успехов в серии из } n \text{ исп.}\}}{N}$

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

Док-во

1) Рассмотр. случ. величины:

$$X_i = \begin{cases} 1, & \text{если в } i\text{-ом исп. имел место успех} \\ 0, & \text{— // — неудача} \end{cases}$$

$X_i \sim B(1, p)$, $i \in \mathbb{N} \Rightarrow MX_i = p$, $DX_i = pq$

Все X_i одинаково распределены.

X_i независимы (т.к. исп. в схеме Бернулли независимы)

Т.о. $X_1, X_2, \ldots$ удовлетв. следствию 1 из теоремы Чебышева $\Rightarrow$

$$\Rightarrow \overline{X}_n \xrightarrow[n\to\infty]{P} MX_i = p$$

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

число успехов в серии из n испыт.

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

6) Сформулировать центральную предельную теорему. Доказать интегральную теор. Муавра-Лапласа.

Центральная предельная теорема.

Пусть: 1) $X_1, X_2, \ldots$ - последов. независ. случ. вел.

2) Все X_i, $i \in \mathbb{N}$, одинаково распред.

3) $\exists MX_i = m$, $\exists DX_i = \sigma^2$

Рассмотрим случ. велич. $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$

$$M[\overline{X}_n] = M\left[\frac{1}{n} \sum_{i=1}^{n} X_i\right] = \frac{1}{n} \sum_{i=1}^{n} MX_i = \frac{1}{n} \sum_{i=1}^{n} m = m$$

$$D[\overline{X}_n] = D\left[\frac{1}{n} \sum_{i=1}^{n} X_i\right] = \frac{1}{n^2} D\left[\sum_{i=1}^{n} X_i\right] = \{X_1, X_2, \ldots \text{-независ.}\} =$$

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

Рассмотрим случ. вел.

$$Y_n = \frac{\bar{X}_n - M[\bar{X}_n]}{\sqrt{D\bar{X}_n}} = \frac{\bar{X}_n - m}{\frac{\sigma}{\sqrt{n}}}, n \in N$$

Тогда $M[Y_n] = M\left[\frac{\bar{X}_n - M\bar{X}_n}{\sqrt{D\bar{X}_n}}\right] =$

$$= \frac{1}{\sqrt{D\bar{X}_n}} [M\bar{X}_n - M\bar{X}_n] = 0$$

$$D[Y_n] = D\left[\frac{\bar{X}_n - M\bar{X}_n}{\sqrt{D\bar{X}_n}}\right] = \frac{1}{D\bar{X}_n} D[\bar{X}_n - M\bar{X}_n] = 1$$

<u>Теорема</u> (ЦПТ)

Пусть 1) выполнены условия (1)-(3)

2) $Y_n = \frac{\bar{X}_n - m}{\frac{\sigma}{\sqrt{n}}} = \frac{\frac{1}{n}\sum\limits_{i=1}^{n} X_i - m}{\frac{\sigma}{\sqrt{n}}}, n \in N$

Тогда последов. $Y_1, Y_2, \dots$ слабо сходится к случайной величине $Z \sim N(0, 1)$,

т.е.

$$\forall x \in R \quad F_{Y_n}(x) \xrightarrow[n \to \infty]{} \Phi(x) = \underbrace{\frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}} dt}_{\text{функция распределен. стандартной случ. величины}}$$

Теорема (интегральная теорема Муавра-Лапласа)

Пусть: 1) проводится серия из $n >> 1$ исп. по схеме Бернулли с вер. успеха p.

2) S – общ. число успехов в этой серии.

Тогда:

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

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

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$$

Док-во

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

Тогда 1) $X_1, X_2, \ldots, X_n$ – независ. случ. велич.

2) $X_i \sim B(1, p), \; i \in N$ – одинаково распред.

3) $MX_i = p, \; DX_i = pq, \; i \in N.$

Т.е. последов. $X_1, X_2, \ldots$ удовл. усл. Т.П.

Кроме того, $S = \sum\limits_{i=1}^{n} X_i$

2) Т.о.

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

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

$$= P\left\{ \frac{\frac{k_1}{n} - p}{\frac{\sqrt{pq}}{\sqrt{n}}} \le \underbrace{\frac{\overline{X}_n - p}{\frac{\sqrt{DX_i}}{\sqrt{n}}}}_{Y_n} \le \frac{\frac{k_2}{n} - p}{\frac{\sqrt{pq}}{\sqrt{n}}} \right\} =$$

$$= P\left\{ \frac{k_1 - np}{\sqrt{npq}} \le Y_n \le \frac{k_2 - np}{\sqrt{npq}} \right\} \approx \left\{ \begin{array}{l} n >> 1 \Rightarrow \\ \Rightarrow Y_n \sim N(0,1) \end{array} \right\} \approx$$

$$\approx \Phi_0(x_2) - \Phi_0(x_1)$$

---

(7) Сформулировать определение случайной выборки и выборки, вариационного ряда. Записать и обосновать выражения для функций распределения случайной выборки и крайних членов вариационного ряда.

Опр. Множество возможных значений случ. велич. X с учётом её закона распределения, возможно неизвестного, будет назыв. генеральной совокупностью.

Опр. случайной выборкой объёма n из генеральной совокупности X назыв. случ. вектор:

$$\vec{X} = (X_1, \dots, X_n), \text{ где}$$

$X_i \sim X$, $i = \overline{1,n}$

$X_1, \dots, X_n$ – независимы в совокупности.

Опр. **Выборкой** объёма n из генеральной совокупности X наз. любая реализация случ. выборки объёма n из генеральной совокупности X:

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

Опр.

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

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

полученный из вектора $\vec{x}$ путём расположения его компонент в порядке неубывания, т. е. числа $x_1, \ldots, x_n$ располагаются в порядке неубывания:

$$x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)}$$

№ ✓

Пусть F — ф-ия распредел. случ. велич. X. Тогда **функция распределения случ. выборки** объёма n из генеральной совокупности X:

$$F_{\vec{X}}(t_1, \ldots, t_n) = P\{X_1 < t_1, \ldots, X_n < t_n\} =$$

$$= \left\{ \begin{matrix} \text{Все } X_i \text{ — независ.} \\ \text{в совокупности} \end{matrix} \right\} = P\{X_1 < t_1\} \cdot \ldots \cdot P\{X_n < t_n\} =$$

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

$$= F(t_1) \cdot \ldots \cdot F(t_n)$$

5

Опр.

Вариационным рядом, отвечающим случ. выборке:

$$\vec{X} = (X_1, \ldots, X_n),$$

называется случ. вектор

$$(X_{(1)}, \ldots, X_{(n)}), \text{ где}$$

$X_{(i)}$ – случ. величина, которая для каждой реализации $\vec{x}$ случ. в-ра $\vec{X}$ принимает значение, равное i-ому члену вариационного ряда, построенного по выборке $\vec{x}$.

Пусть $F(t)$ – ф-ия распред. генеральной совокупности X.

Тогда:

а) $F_{X_{(n)}}(t) = P\{X_{(n)} < t\} = P\{X_1 < t, X_2 < t, \ldots, X_n < t\} = \left\{\begin{matrix} X_i \text{ – независ.} \\ \text{в совокупности} \end{matrix}\right\} = P\{X_1 < t\} \cdot \ldots \cdot P\{X_n < t\} =$

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

$= [F(t)]^n$

б) $F_{X_{(1)}}(t) = P\{X_{(1)} < t\} = 1 - P\{X_{(1)} \geq t\} =$

$= 1 - P\{X_1 \geq t, \ldots, X_n \geq t\} = \{X_i \text{ – независ.}\} =$

$= 1 - P\{X_1 \geq t\} \cdot \ldots \cdot P\{X_n \geq t\} =$

$= 1 - [1 - P\{X_1 < t\}] \cdot \ldots \cdot [1 - P\{X_n < t\}] =$

$$= \{ X_i \sim X \Rightarrow F_{X_i} = F_X \} = 1 - [1 - F(t)]^n$$

---

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

<u>Опр.</u>

Начальным выборочным моментом порядка k назыв. статистика:

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

<u>Опр.</u>

Центральным выборочным моментом порядка k назыв. статистика:

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

<u>Опр.</u>

Выборочным средним назыв. статистика:

$$\underbrace{\hat{m}_1(\vec{X}) = \bar{X}}_{\text{обозн.}} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

<u>Опр.</u> Выборочной дисперсией назыв. статистика:

$$\underbrace{\hat{\mathring{m}}_2(\vec{X}) = \hat{\sigma}^2(\vec{X})}_{\text{обозн.}} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

Опр. Точечная оценка $\hat{\theta}$ парам. θ наз. несмещённой если

$$\exists M\left[\underbrace{\boxed{\hat{\theta}(\vec{X})}}_{\text{случ. велич.}}\right] = \underbrace{\theta}_{\text{теорет. знач. парам}}$$

случ. выборка

1) Пусть X – ген. совокупность

$$\theta = MX$$

Покажем, что $\hat{\theta}(\vec{X}) = \bar{X}$ явл. несмещённой оценкой для θ

$$M[\hat{\theta}(\vec{X})] = M\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} MX_i =$$

$$= \left\{\begin{array}{l} \text{из опр. случ. выб.} \Rightarrow \\ \Rightarrow X_i \sim X,\ i=\overline{1,n} \Rightarrow \\ \Rightarrow \exists MX_i = MX = \theta \end{array}\right\} =$$

$$= \frac{1}{n}\sum_{i=1}^{n}\theta = \frac{\theta \cdot n}{n} = \theta,$$

т.е. $\hat{\theta}_1$ – несмещ. оценка для θ.

2) Пусть X – ген. совокупность

$$\exists DX = \sigma^2$$

Покажем, что $\hat{\sigma}^2(\vec{X})$ является смещённой оценкой для σ^2

Обозн: $MX = m$

$$\hat{\sigma}^2(\vec{X}) = \underbrace{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}_{\text{выборочная дисперсия}}$$

На семинаре было доказано, что:

$$M[\sigma^2(\vec{X})] = \frac{n-1}{n}\sigma^2 + \sigma^2,$$ т.е.

выборочная дисперсия явл. смещ. оценкой дисперсии.

Выборочное среднее явл. несмещ. оценкой своего теор. аналога, а выборочная дисперсия — нет.

---

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

Пусть $\vec{x} = (x_1, \dots, x_n)$ — выборка из генер. совокупн. X.

Обознач.: $n(t, \vec{x})$ — число компонент в-ра $\vec{x}$, которые меньше, чем t.

<u>Опр.</u>

Эмпирической ф-ей распределения, построенной по выборке $\vec{x}$, назыв. функцию:

$$F_n : \mathbb{R} \to \mathbb{R},$$

определ. правилом:

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

Пусть $\vec{X}=(X_1,\dots,X_n)$ - случ. выборка из генер. совок. X

Обознач.: $n(t,\vec{X})$ - сл. вел., кот. для каждой реализ. $\vec{x}$ случ. выборки $\vec{X}$ принимает знач., равные $n(t,\vec{x})$

Опр.

Выборочной функцией распределения, отвечающ. случ. выборке $\vec{X}$, наз. отображ.:

$$\hat{F}: R \to R,$$

определ. правилом:

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

Теорема

Для любого фиксирован. $t_0 \in R$ последов. случ. величин

$$\hat{F}(t_0,\vec{X}_1), \hat{F}(t_0,\vec{X}_2), \dots, \hat{F}(t_0,X_n), \dots$$

сходится по вероятности к значению теоретической ф-ции распределения генер. совок. X, т.е.

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

Док-во

При фикс. $t_0 \in R$ случ. вел.

$$\hat{F}(t_0,\vec{X}_n)=\frac{n(t_0,\vec{X}_n)}{n}$$

равна относ. частоте успеха (т.е. события $\{X < t_0\}$)

в серии из n испыт. по схеме Бернулли

В соответствии с ЗБЧ в форме Бернулли:

$$\hat{F}(t_0, \vec{X}_n) \xrightarrow[n\to\infty]{P} p,$$ теорет. частота

но $p = P\{X < t_0\} = F(t_0)$, т.е.

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

---

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

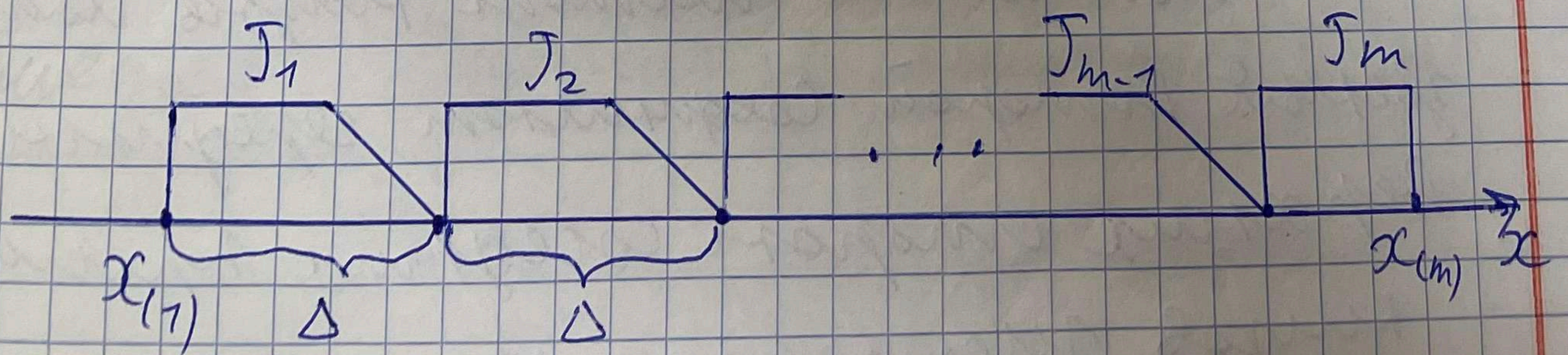
Если объём выборки достаточно велик ($n > 50$), то эл-ты выборки группируются в так называемый стат. ряд. Для этого отрезок $J = [x_{(1)}, x_{(n)}]$ разбивают на m равновеликих промежутков. Ширина каждого из них:

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

Далее полагают m:

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

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



Опр.
Интервальным статистическим рядом, отвечающим выборке $\vec{x}$, наз. таблица вида

J_1	J_2	$\dots$	J_m
n_1	n_2	$\dots$	n_m

Здесь n_i – число элементов выборки $\vec{x}$, попавших в промежуток J_i, $i=\overline{1;m}$
$\#$

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

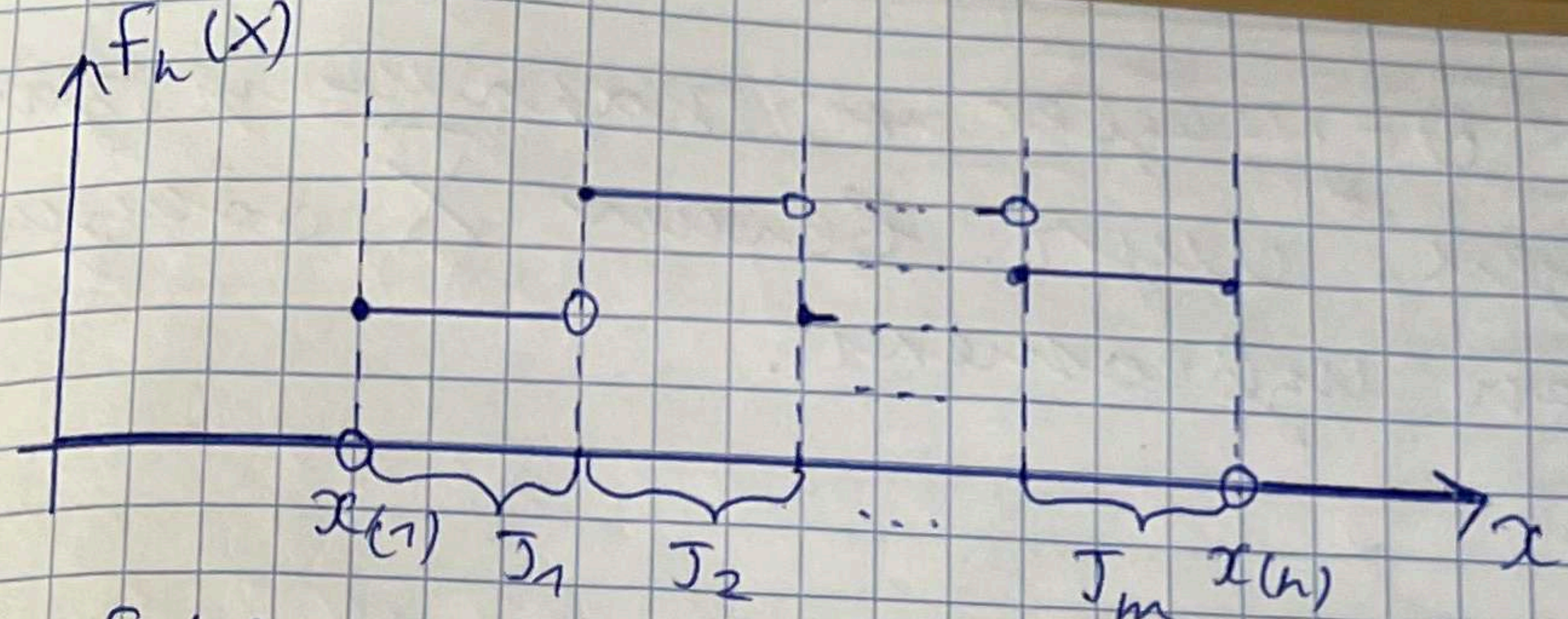
Опр.
Эмпирической плотностью распред. (соответствующ. выборке $\vec{x}$) называв. функция:

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

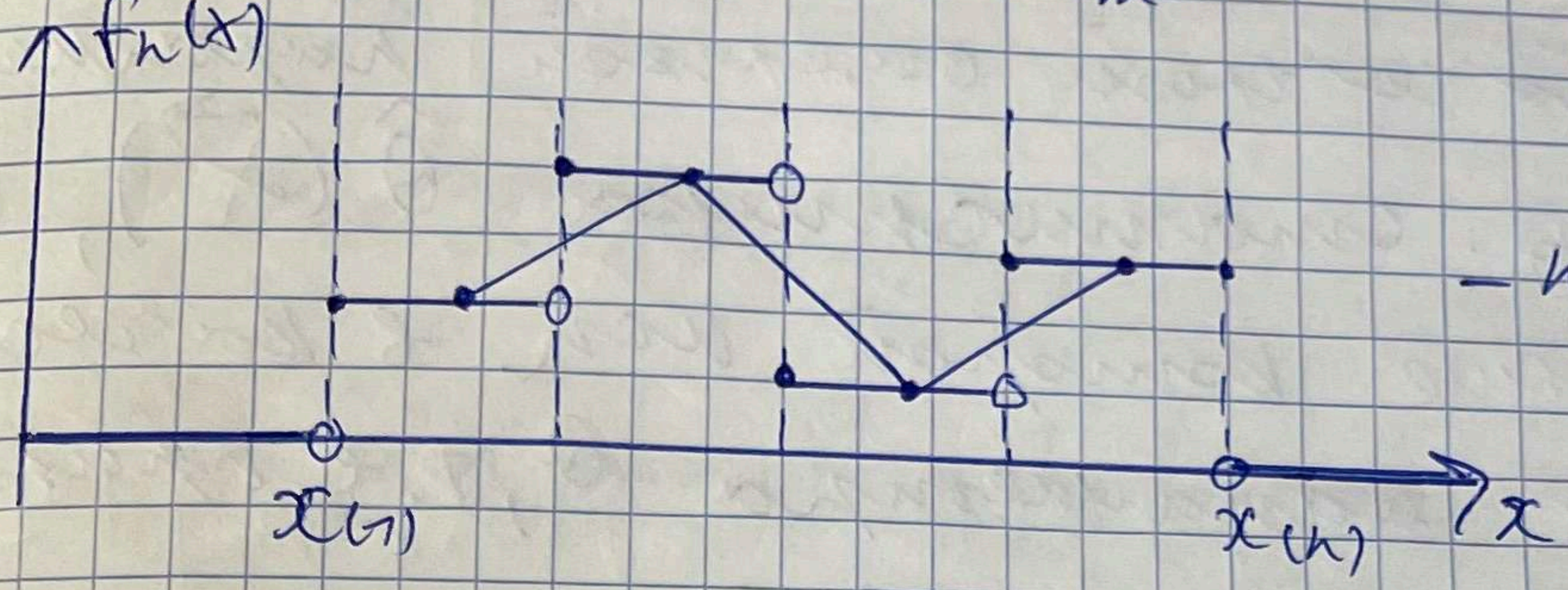
Опр.
График эмпирической функции плотности наз. гистограммой
$\#$

Пусть для некот. выборки $\vec{x}$ постр. гистогр.

Опр.
Полигоном частот называв. ломаная, звенья которой соединяют середины верхних сторон соседних прямоугольников гистограммы



– гистограмма



– полигон частот

---

(11) Постановка задачи идентификации неизвестных парам. закона распредел. случ. величины. Определение точечной оценки. Определение несмещённой точечной оценки. Показать, что выборочная дисперсия является смещённой оценкой дисперсии. Записать формулу для исправленной выборочной дисперсии.

Задача идентификации

Дано: X – случ. велич., закон распредел. которой известен с точностью до в-ра

$$\vec{\theta} = (\theta_1, \dots, \theta_r)$$

неизвестных параметров.

Требуется: Оценить (найти) в-р $\vec{\theta}$.

Пусть θ – неизвестн. параметр закона распредел. случ. величин. $\vec{X}$, общий вид которой известен.

<u>Опр.</u>

<u>Точечной оценкой параметра</u> θ наз. статистика $\hat{\theta}(\vec{X})$, выборочн. значение которой исп. в качестве знач. параметра θ, т.е. принимается равенство:

$$\theta = \hat{\theta}(\vec{X})$$

н

Пусть θ – неизв. парам. закона распред. случ. величин X;

$\hat{\theta}(\vec{X})$ – точечная оценка для θ

<u>Опр.</u>

<u>Точечная оценка</u> $\hat{\theta}$ парам. θ называется несмещённой, если

$$\exists M[\hat{\theta}(\vec{X})] = \theta$$

теор. знач. парам.

<u>Доказать</u>, что выборочная дисперсия $\hat{\sigma}^2(\vec{X})$ явл. смещ. оценкой дисперсии.

<u>Решение</u>

1) Обозначим:

$m = MX$, } возможно, неизвестно

$\sigma^2 = DX$ } возможно, что-то известно

$$\sigma^2(\vec{X}) = \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - m - (\overline{X}-m)\right)^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n}\Big[\underbrace{(X_i-m)}_{a} - \underbrace{\frac{1}{n}\sum_{j=1}^{n}(X_j-m)}_{b}\Big]^2 =$$

{ Нужно проверить, что $M[\hat{\sigma}^2(\vec{X})] \overset{?}{=} \sigma^2$

$$= \frac{1}{n}\sum_{i=1}^{n}\Big[(X_i-m)^2 + \frac{1}{n^2}\Big(\sum_{j=1}^{n}(X_j-m)\Big)^2 -$$

$$- \frac{2}{n}(X_i-m)\sum_{j=1}^{n}(X_j-m)\Big] =$$

$$= \frac{1}{n}\sum_{i=1}^{n}\Big[(X_i-m)^2 + \frac{1}{n}\sum_{j=1}^{n}(X_j-m)^2 +$$

$$+ \frac{2}{n^2}\sum_{1\le j<k\le n}(X_j-m)(X_k-m) -$$

$$- \frac{2}{n}\sum_{j=1}^{n}(X_i-m)(X_j-m)\Big] =$$

<u>Обозн.</u>: $A_i = (X_i-m)^2$

$$B = \frac{1}{n^2}\sum_{j=1}^{m}(X_j-m)^2$$

$$C = \frac{2}{n^2}\sum_{1\le j<k\le n}(X_j-m)(X_k-m)$$

$$D_i = -\frac{2}{n}\sum_{j=1}^{n}(X_i-m)(X_j-m)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left[A_i + B + C + D_i\right]$$

$$M[\hat{\sigma}^2(\vec{X})] = \frac{1}{n}\sum_{i=1}^{n}[MA_i + MB + MC + MD_i]$$

$$MA_i = M[(X_i - m)^2] = DX_i = \{X_i \sim X\} = DX = \sigma^2$$

$$MB = \frac{1}{n^2}\sum_{j=1}^{n}\underbrace{M[(X_j - m)^2]}_{DX_j = DX = \sigma^2} = \frac{1}{n^2}\sum_{j=1}^{n}\sigma^2 = \frac{\sigma^2}{n}$$

$$MC = \frac{2}{n^2}\sum_{1\le j<k\le n} M[(X_j - m)(X_k - m)] =$$

$$= \frac{2}{n^2}\sum_{j<k} cov(X_i, X_j) = \left\{\begin{array}{l} j<k \Rightarrow j \neq k \\ \Rightarrow X_j \text{ и } X_k \\ \text{независ. как элем.} \\ \text{случ. выборки} \Rightarrow \\ cov(X_j, X_k) = 0 \end{array}\right\} = 0.$$

$$MD_i = -\frac{2}{n}\sum_{j=1}^{n} M[(X_i - m)(X_j - m)] =$$

$$= -\frac{2}{n}\sum_{j=1}^{n} cov(X_i, X_j) =$$

$$= \left\{\begin{array}{l} cov(X_i, X_j) = 0 \text{ при } i = j \\ cov(X_i, X_i) = \sigma^2 \text{ при } i = j \end{array}\right\} = -\frac{2}{n}\sigma^2$$

Т.о. $M[\hat{\sigma}^2(\vec{X})] = \frac{\sigma^2}{n}\sum_{i=1}^{n}\left[1 + \frac{1}{n} + 0 - \frac{2}{n}\right] =$

$$= \frac{\sigma^2}{n}\cdot n\left[1 - \frac{1}{n}\right] = \frac{n-1}{n}\sigma^2 \neq \sigma^2$$

Опр.: Статистика

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

наз. исправленной выборочной дисперсией.

(12) Постановка задачи идентификации неизвестных параметров закона распределения случайной величины. Определение точечной оценки. Определение эффективной оценки. ...

Задача идентификации

Дано: X - случ. величина, закон распредел. которой известен с точностью до в-ра
$$\vec{\theta} = (\theta_1, \dots, \theta_r)$$
неизвестн. параметров

Требуется:
Оценить в-р $\vec{\theta}$

Пусть θ - неизв. парам. з-на распред. случ. велич. X, общий вид которой известен

Опр. Точечной оценкой параметра θ называв. статистика $\hat{\theta}(\vec{X})$, выборочное значение которой использ. в качестве знач. парам-ра θ, т.е. принимается равенство:
$$\theta = \hat{\theta}(\vec{X})$$

Пусть: 1) X - генеральная совокупность;
2) θ - параметр закона распредел. случ. вел. X
3) $\hat{\theta}(\vec{X_n})$ - точечная оценка парам-ра θ

Опр. Точечная оценка $\hat{\theta}(\vec{X_n})$ парам-ра θ называв. состоятельной, если:
$$\hat{\theta}(\vec{X_n}) \xrightarrow[n\to\infty]{P} \theta$$

(13) — // —

Определение эффективной оценки.

...

Опр. Оценка $\hat{\theta}$ назыв. эффективной оценкой парам-ра θ, если:

1) $\hat{\theta}$ является несмещ. оценкой для θ

2) $\hat{\theta}$ обладает наим. дисперсией среди всех несмещ. точечных оценок парам-ра θ

(14) — // —

Доказать теорему о единственности эффективной оценки.

Теорема

Пусть $\hat{\theta}_1(\vec{X})$ и $\hat{\theta}_2(\vec{X})$ - две эффективные ($\Rightarrow$ несмещённые) оценки парам-ра θ.

Тогда:

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

Док-во

Без док-ва.

—— // ——

(15) Определение кол-ва информации по Фишеру. Сформулировать теорему о неравенстве Рао-Крамера.

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

Опр. Кол-вом информ. по Фишеру, содерж. в случ. выборке $\vec{X}$ наз. число:

$$I(\theta)=M\left[\left(\frac{d\ln L}{d\theta}\right)^2\right],$$ где

L – функц. правдоподобия, отвечающ. случ. выборке $\vec{X}$:

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

$$p(X_i,\vec{\theta})=\begin{cases} f(X_1,\vec{\theta}), & \text{если } X \text{ – непр. случ. велич.}\\ p(X=X_i), & \text{если } X \text{ – дискр. случ. велич.}\end{cases}$$

(здесь f – функц. плотности распредел. вероятн. этой случ. велич. X)

Теорема (нер-во Рао-Крамера)

Пусть 1) Рассматриваемая модель явл. регулярной

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

Тогда:

$$D[\hat{\theta}]\geqslant\frac{1}{I(\theta)}$$

(16) —//— (17) (18)

...

(19-22) —— II ——

Сформулировать определ. γ-довер. интервала. Сформулировать определение центральной статистики.

Интерв. оценка

Опр. Интервальн. оценкой парам. θ уровня $\gamma \in (0,1)$ (γ-интерв. оценкой) назыв. пара статистик:

$$\underline{\theta}(\vec{X}) \text{ и } \overline{\theta}(\vec{X}) \text{ таких, что}$$

$$P\{\theta \in (\underline{\theta}(\vec{X}), \overline{\theta}(\vec{X}))\} = \gamma$$

Довер. интервал

Опр. γ-довер. интерв. (доверительным интервалом уровня γ) для парам. θ наз. реализацию интерв. оценки уровня γ для этого парам-ра, т.е. интервал

$$(\underline{\theta}(\vec{x}), \overline{\theta}(\vec{x}))$$

с детерминированными границами.

Опред. центр. статистики

Пусть X - случ. величина, з-н распредел. которой известен с точностью до значения парам. $\theta \in R$

Опр. Статистика $g(\vec{X}, \theta)$ назыв. центральной, если закон её распред. не зависит от неизв. парам-ра θ.