Математическая статистика

Домашнее задание №1.

Вариант: 22

ФИО: Трунов Андрей Романович

Группа: ИУ7-62Б

Задача №1. (Предельные теоремы теории вероятностей)

Условие: Известно, что 80% изготовленных заводом электроламп выдерживают гарантийный срок службы. Найти вероятность того, что в партии из 500 электроламп число выдержавших гарантийный срок службы находится в пределах от 380 до 420. Использовать неравенство Чебышева и интегральную теорему Муавра - Лапласа

Решение: 1) Формализуем задачу:

а) Пусть X_i - случайная величина, принимающая значения:
$\begin{cases} 1, \text{ если } i\text{-я электролампа выдержала гарантийный срок службы} \\ 0, \text{ иначе} \end{cases}$
, где $i = \overline{1, n}$

б) Пусть n - количество электроламп в партии. Из условия $n = 500$.

в) Пусть p - вероятность того, что электролампа выдержала гарантийный срок службы. Из условия $p = 0{,}8$.

$q = 1 - p = 1 - 0{,}8 = 0{,}2$

2) $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ и $X_i \sim B(n, p)$, тогда $M[\overline{X}_n] = \frac{1}{n}\sum_{i=1}^{n} M[X_i] = \frac{1}{n}\cdot np = p = 0{,}8$

$D[\overline{X}_n] = \frac{1}{n^2}\sum_{i=1}^{n} D[X_i] = \frac{1}{n^2} npq = \frac{pq}{n} = \frac{0{,}8\cdot 0{,}2}{500} = 0{,}00032$

3) Воспользуемся 1-ым неравенством Чебышева

$\forall \varepsilon > 0 \quad \boxed{P\{X \geq \varepsilon\} \leq \frac{M[X]}{\varepsilon}} \Rightarrow 1 - P\{X < \varepsilon\} \leq \frac{M[X]}{\varepsilon} \Rightarrow \boxed{P\{X < \varepsilon\} \geq 1 - \frac{M[X]}{\varepsilon}}$

Найдем $P\{\frac{380}{500} \leq \overline{X}_n \leq \frac{420}{500}\}$:

$P\{\frac{380}{500} \leq \overline{X}_n \leq \frac{420}{500}\} \leq P\{\overline{X}_n \geq \frac{380}{500}\} \leq \frac{0{,}8}{\frac{380}{500}} = \frac{0{,}8}{0{,}76} \approx 1{,}05 \Rightarrow$ тривиальная оценка (св-во вероятности)

$P\{\frac{380}{500} \leq \overline{X}_n \leq \frac{420}{500}\} \leq P\{\overline{X}_n \leq \frac{420}{500}\} \geq 1 - \frac{0{,}8}{0{,}76} \Rightarrow$ бесполезная оценка из-за знаков нерав-в

4) Воспользуемся 2-ым неравенством Чебышева

$\forall \varepsilon > 0 \quad \boxed{P\{|X - M[X]| \geq \varepsilon\} \leq \frac{D[X]}{\varepsilon^2}} \Rightarrow \boxed{P\{|X - M[X]| < \varepsilon\} \geq 1 - \frac{D[X]}{\varepsilon^2}}$

Найдем $P\{\frac{380}{500} \leq \overline{X}_n \leq \frac{420}{500}\}$:

$P\{\frac{380}{500} \leq \overline{X}_n \leq \frac{420}{500}\} = P\{\frac{380}{500} - 0{,}8 \leq \overline{X}_n - M[\overline{X}_n] \leq \frac{420}{500} - 0{,}8\} = P\{|\overline{X}_n - M[\overline{X}_n]| \leq 0{,}04\} \geq$

$P\{|\overline{X}_n - M[\overline{X}_n]| < 0{,}04\} \geq 1 - \frac{0{,}00032}{0{,}04^2} = 1 - \frac{0{,}00032}{0{,}0016} = 1 - 0{,}2 = 0{,}8$

$\boxed{P\{\frac{380}{500} \leq \overline{X}_n \leq \frac{420}{500}\} \geq 0{,}8}$

$S = n \cdot \overline{X}_n$ — общее количество электроламп, выдержавших гарантийный срок службы

5) Воспользуемся интегральной теоремой Муавра-Лапласа

$$\boxed{P\{k_1 \le S \le k_2\} \approx \Phi_0(x_2) - \Phi_0(x_1)}, \text{ где } x_i = \frac{k_i - np}{\sqrt{npq}}, \; i = 1,2, \; q = 1-p$$

$k_1 = 380$, $k_2 = 420$ (из условия)

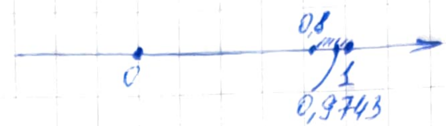
Найдем $x_1 = \dfrac{380 - 500 \cdot 0{,}8}{\sqrt{500 \cdot 0{,}8 \cdot 0{,}2}} = -\sqrt{5} \approx -2{,}23$, $x_2 = \dfrac{420 - 500 \cdot 0{,}8}{\sqrt{500 \cdot 0{,}8 \cdot 0{,}2}} = \sqrt{5} \approx 2{,}23$

Тогда $P\{380 \le S \le 420\} = P\{380 \le n\,\overline{X}_n \le 420\} \approx \Phi_0(2{,}23) - \Phi_0(-2{,}23) =$

$= 2\Phi_0(2{,}23) \approx 2 \cdot 0{,}48715 = 0{,}9743$

$$\boxed{P\{380 \le n \cdot \overline{X}_n \le 420\} = 0{,}9743}$$

6) На прямой



Ответ: вероятность а) с использованием 2-го неравенства Чебышева
$P\{380 \le n \cdot \overline{X}_n \le 420\} \ge 0{,}8$

б) с использованием интегральной теоремы Муавра-Лапласа
$P\{380 \le n \cdot \overline{X}_n \le 420\} = 0{,}9743$

Задача №2. (метод моментов)

Условие: С использованием метода моментов для случайной выборки $\vec{X}=(X_1,\dots,X_n)$ из генеральной совокупности X найти точечные оценки указанных параметров заданного закона распределения.

$$f_X(x)=\frac{1}{4^\theta\Gamma(\theta)}x^{\theta-1}e^{-\frac{x}{4}},\ x>0.$$

Решение: 1) Т.к. в задаче 1 неизвестный параметр $\Rightarrow r=1\Rightarrow$ потребуется одно уравнение.

Запишем уравнение относительно начального момента 1-го порядка:

$m_1(\theta)=\hat{m}_1(\vec{X})$

2) $$m_1(\theta)=M[X]=\int_{-\infty}^{+\infty}x f_X(x)dx=\int_0^{+\infty}x\cdot\frac{1}{4^\theta\Gamma(\theta)}x^{\theta-1}e^{-\frac{x}{4}}dx=$$

$$=\frac{1}{4^\theta\Gamma(\theta)}\int_0^{+\infty}x^\theta e^{-\frac{x}{4}}dx=\left\{\begin{matrix}t=\frac{x}{4}\Rightarrow x=4t\\ dt=\frac{dx}{4}\Rightarrow dx=4dt\end{matrix}\right\}=\frac{1}{4^\theta\Gamma(\theta)}\int_0^{+\infty}(4t)^\theta\cdot e^{-\frac{4t}{4}}\cdot dt=$$

$$=\frac{\cancel{4^\theta}\cdot 4}{\cancel{4^\theta}\Gamma(\theta)}\underbrace{\int_0^{+\infty}t^\theta\cdot e^{-t}dt}_{=\Gamma(\theta+1)}=\frac{4\cdot\Gamma(\theta+1)}{\Gamma(\theta)}=\frac{4\cdot\theta\cancel{\Gamma(\theta)}}{\cancel{\Gamma(\theta)}}=4\theta$$

3) $$m_1(\vec{X})=\frac{1}{n}\sum_{i=1}^{n}X_i=\overline{X}_n$$

Тогда $\theta=\frac{\overline{X}_n}{4}$, т.о. $\hat{\theta}(\vec{X})=\frac{\overline{X}_n}{4}$

$$\boxed{\text{Ответ: }\ \hat{\theta}(\vec{X})=\frac{\overline{X}_n}{4}}$$

Задача №3 (метод максимального правдоподобия)

Условие: С использованием метода максимального правдоподобия для случайной выборки $\vec{X}=(X_1,\dots,X_n)$ из генеральной совокупности X найти точечные оценки параметров заданного закона распределения. Вычислить выборочные значения найденных оценок для выборки $\vec{x}_5=(x_1,\dots,x_5)$.

$$f_X(x)=\frac{\theta^5}{4!}x^4e^{-\theta x},\ x>0 \qquad \vec{x}_5=(7,4,11,5,3)$$

Решение: 1) Составим функцию правдоподобия L

$$L(\vec{X},\theta)=p(X_1,\theta)\cdot\ldots\cdot p(X_n,\theta)=\left\{\begin{matrix}X\text{ - непрерывная}\\\text{случайная}\\\text{величина}\end{matrix}\right\}=$$

$$=f_X(X_1,\theta)\cdot\ldots\cdot f_X(X_n,\theta)=\left(\frac{\theta^5}{4!}\right)^n\cdot\left(X_1^4\cdot\ldots\cdot X_n^4\right)e^{-\theta\sum_{i=1}^{n}X_i}\longrightarrow\max_{\theta}$$

св-ва логарифмов:
$\log_a bc=\log_a b+\log_a c$
$\log_a\frac{b}{c}=\log_a b-\log_a c$
$\log_a a^m=m$
$\log_a b^m=m\log_a b$

св-ва степеней:
$(a^m)^n=a^{m\cdot n}$

Вместо этой задачи будем решать эквивалентную:

$$\ln(L(\vec{X},\theta))\longrightarrow\max_{\theta}$$

Тогда $\ln(L(\vec{X},\theta))=\ln(\theta^{5n})-\ln((4!)^n)+\sum_{i=1}^{n}\ln(X_i^4)+\ln(e^{-\theta\sum_{i=1}^{n}X_i})=$

$$=5n\ln(\theta)-n\ln(24)+4^n\sum_{i=1}^{n}\ln(X_i)-\theta\sum_{i=1}^{n}X_i\longrightarrow\max_{\theta}$$

2) Уравнение правдоподобия:

$$\frac{\delta(\ln(L))}{\delta(\theta)}=\frac{5n}{\theta}-0+0-\sum_{i=1}^{n}X_i=0$$

Тогда $\theta=\frac{5n}{\sum_{i=1}^{n}X_i}=\frac{5\cancel{n}}{\cancel{n}\cdot\overline{X}_n}=\frac{5}{\overline{X}_n}$, т.о $\hat{\theta}=\frac{5}{\frac{(7+4+11+5+3)}{5}}=\frac{25}{30}=\frac{5}{6}$

Проверим достаточное условие max

$$\left.\frac{\delta^2(\ln(L))}{\delta(\theta^2)}\right|_{\theta=\hat{\theta}}<0$$

$$\frac{\delta^2(\ln(L))}{\delta(\theta^2)}=-\frac{\overset{>0}{5n}}{\underset{>0}{\theta^2}}<0\implies\theta=\frac{5}{\overline{X}_n},\text{ точка max}$$

$$\boxed{\text{Ответ: }\theta=\frac{5}{\overline{X}_n};\ \hat{\theta}=\frac{5}{6}}$$

Задача №4. (доверительные интервалы)

Условие: После $n=8$ измерений давления в баке с горючим получены следующие результаты: 3,25; 2,82; 3,07; 3,12; 2,93; 2,87; 3,09; 3,17.

Считая ошибки измерений подчиненными нормальному закону, построить 90%-ные доверительные интервалы для математического ожидания и среднего квадратичного отклонения (СКО) давления в баке.

Решение: 1) Формализуем задачу.

а) Пусть a – давление в баке с горючим; детерминированная, но неизвестная величина.

б) Пусть ε – случайная величина, принимающая значения равные ошибкам измерения давления в баке с горючим. Из условия $\varepsilon \sim N(0, \sigma^2)$

систематические ошибки примем равными 0

в) Пусть $X = a + \varepsilon$ – случайная величина, принимающая значения, равные результатам измерений давления в баке с горючим.

Тогда $X = const + \underbrace{\varepsilon}_{\sim N(\ldots)} \Rightarrow X \sim N(M_X, \sigma_X^2)$

г) Найдем $M_X = M[X] = M[a+\varepsilon] = a + \overbrace{M[\varepsilon]}^{0} = a$

$\sigma_X^2 = D[X] = D[a+\varepsilon] = D[\varepsilon] = \sigma^2$

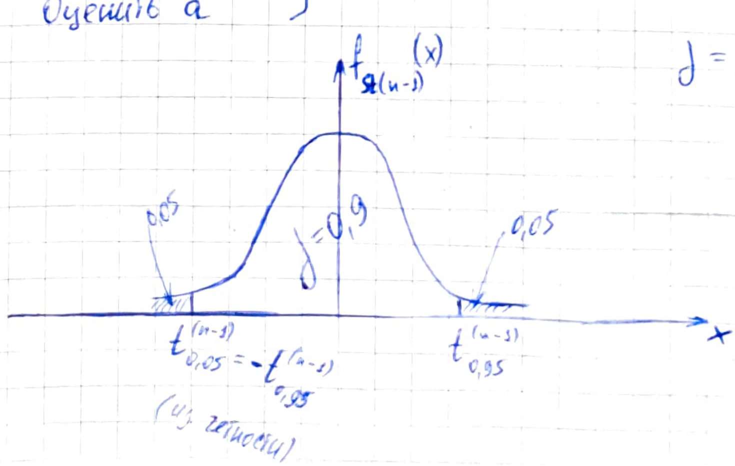
Т.о. $X \sim N(a, \sigma^2)$, где a и σ^2 – неизвестные величины.

2) Доверительный интервал для a.

$$\left.\begin{array}{l} X \sim N(a,\sigma^2) \\ a \text{ – неизвестно} \\ \sigma^2 \text{ – неизвестно} \\ \text{Оценить } a \end{array}\right\} \Rightarrow T(\vec{X}, a) = \frac{a - \bar{X}}{s(\vec{X})}\sqrt{n} \sim St(n-1)$$

$\gamma = 0,9$ (из условия)

$t_{0,95}^{(n-1)}$ – квантиль уровня 0,95 распределения $St(n-1)$



$$P\left\{-t^{(n-1)}_{0,95} < \underbrace{\frac{a-\overline{X}}{S(\vec{X})}\sqrt{n}}_{T(\vec{X},a)} < t^{(n-1)}_{0,95}\right\} = 0,9$$

Выразим a: $P\left\{\underbrace{\overline{X} - \frac{s(\vec{X})\,t^{(n-1)}_{0,95}}{\sqrt{n}}}_{\underline{a}(\vec{X})} < a < \underbrace{\overline{X} + \frac{s(\vec{X})\,t^{(n-1)}_{0,95}}{\sqrt{n}}}_{\overline{a}(\vec{X})}\right\} = 0,9$

Доверительный интервал для a: $(\underline{a}(\vec{X}); \overline{a}(\vec{X}))$

Найдем доверительный интервал для a:

а) $n-1 = 8-1 = 7$

б) $t^{(7)}_{0,95} = 1,8946$ (из таблицы)

в) $\overline{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i = \frac{1}{8}(3,25+2,82+3,07+3,12+2,93+2,87+3,09+3,17) = 3,04$

г) $S^2(\vec{X}) = \frac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2 = \frac{1}{7}\cdot(0,21^2+(-0,22)^2+0,03^2+0,08^2+(-0,11)^2+(-0,17)^2+$
$+0,05^2+0,13^2) = \frac{801}{35000} \Rightarrow S(\vec{X}) = \sqrt{\frac{801}{35000}} \approx 0,15128$

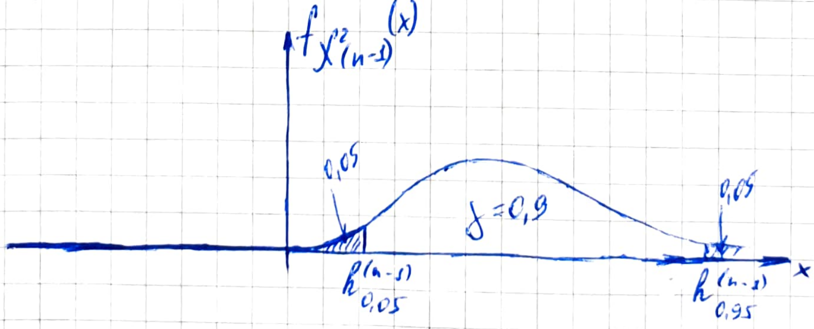
д) $\underline{a}(\vec{X}) = 3,04 - \frac{0,15128\cdot 1,8946}{\sqrt{8}} \approx 2,93867$

е) $\overline{a}(\vec{X}) = 3,04 + \frac{0,15128\cdot 1,8946}{\sqrt{8}} \approx 3,14133$

$$\boxed{M_X = a \in (2,93867; 3,14133)}$$

3) Доверительный интервал для σ.

$\left.\begin{array}{l} X \sim N(a,\sigma^2) \\ \sigma^2 \text{ – неизвестно} \\ \text{Оценить } \sigma \end{array}\right\} \Rightarrow T(\vec{X},\sigma) = \frac{(n-1)S^2(\vec{X})}{\sigma^2} \sim \chi^2(n-1)$



$h^{(n-1)}_{0,05}$, $h^{(n-1)}_{0,95}$ – квантиль уровня 0,05 и 0,95 соответственно распределения $\chi^2(n-1)$.

$$P\left\{ h_{0,05}^{(n-1)} < \frac{(n-1)S^2(\vec{X})}{\sigma^2} < h_{0,95}^{(n-1)} \right\} = 0,9$$

Выразим σ : $P\left\{ \underbrace{\sqrt{\frac{(n-1)S^2(\vec{X})}{h_{0,05}^{(n-1)}}}}_{\underline{\sigma}(\vec{X})} < \sigma < \underbrace{\sqrt{\frac{(n-1)S^2(\vec{X})}{h_{0,95}^{(n-1)}}}}_{\overline{\sigma}(\vec{X})} \right\} = 0,9$

Доверительный интервал для σ : $(\underline{\sigma}(\vec{X}); \overline{\sigma}(\vec{X}))$.

Найдем доверительный интервал для σ :

а) $n-1 = 8-1 = 7$

б) $h_{0,05}^{7} = 14,067$, $h_{0,95}^{7} = 2,167$ (из таблицы)

в) $S^2(\vec{X}) = \{$посчитали в предыдущем пункте решения$\} \approx 0,15128$

г) $\overline{\sigma}(\vec{X}) = \sqrt{\frac{7 \cdot 0,15128}{2,167}} \approx \underline{\underline{0,699053}}$

$\underline{\sigma}(\vec{X}) = \sqrt{\frac{7 \cdot 0,15128}{14,067}} \approx \underline{\underline{0,274372}}$

$$\boxed{\sigma_X = \sigma \in (0,274372; 0,699053)}$$

Ответ: Доверительный интервал для математического ожидания давления в баке: $(2,93867; 3,14133)$.

Доверительный интервал для СКО давления в баке: $(0,274372; 0,699053)$.