

**Федеральное государственное автономное образовательное
учреждение высшего образования «Национальный исследовательский
университет ИТМО»
Факультет программной инженерии и компьютерной техники**

Математическая Статистика

РГР 1

Вариант 2

Выполнили:

Кобик Никита Алексеевич

Маликов Глеб Игоревич

Группа № Р3224

Преподаватель:

Кудрявцева Ирина Владимировна

г. Санкт-Петербург

2023

Оглавление

Задание.....	3
Исходные данные	4
Решение	5
Размах варьирования.....	5
Статический ряд распределения частот СВ X.....	6
Интервальные статистические ряды частот и относительных частот	7
Полигон и гистограмма относительных частот.....	8
Эмпирическая функция распределения и её график.....	9
Выборочные значения числовых характеристик СВ X: математическое ожидание, дисперсия, среднее квадратическое отклонение	9
Выбор закона распределения СВ X по виду полигона и гистограммы относительных частот	10
Точечная оценка параметров предполагаемого распределения и функции распределения и плотности вероятности СВ X.....	10
Проверка согласия эмпирической функции распределения с теоретической при помощи критерия согласия Пирсона	11
Интервальные оценки параметров распределения.....	13
Проверка нулевой гипотезы о математическом ожидании при альтернативной гипотезе	14
Проверка нулевой гипотезы о дисперсии против альтернативной гипотезы.....	14
МНК	16
Задание 1.....	16
Дополнительное задание 1.....	18
Задание 2.....	20
Дополнительное задание 2.....	21

Задание

Требуется:

1. определить размах варьирования R ;
2. составить статистический ряд распределения частот $CB X$;
3. составить интервальные статистические ряды частот и относительных частот;
4. построить полигон и гистограмму относительных частот;
5. найти эмпирическую функцию распределения и построить её график;
6. вычислить выборочные значения числовых характеристик $CB X$:
математического ожидания $M(X)$, дисперсии $D(X)$, среднего квадратического отклонения $\sigma(X)$;
7. по виду полигона и гистограммы относительных частот сделать выбор закона распределения $CB X$;
8. найти точечные оценки параметров предполагаемого распределения, записать функцию распределения и плотность вероятности $CB X$;
9. проверить согласие эмпирической функции распределения $F^*(x)$ с теоретической $F(x)$ при помощи критерия согласия χ^2 —Пирсона или λ —Колмогорова.

В случае нормального распределения $CB X$ по заданному уровню значимости α :

1. найти интервальные оценки параметров распределения;
2. проверить нулевую гипотезу $H_0: M(X) = a = a_0$ о математическом ожидании и при альтернативной гипотезе $H_a: a \neq a_0 (a > a_0; a < a_0)$;
3. проверить нулевую гипотезу $H_0: D(x) = \sigma^2 = \sigma_0^2$ о дисперсии против альтернативной гипотезы $H_a: \sigma^2 \neq \sigma_0^2 (\sigma^2 < \sigma_0^2; \sigma^2 > \sigma_0^2)$.

Исходные данные

Вариант 2

Химический анализ 110 проб угля для определения воды дал следующие результаты (в процентах):

14	13	17	18	22	15	21	22	14	15	17
15	18	14	14	19	16	23	13	12	16	19
15	19	14	13	18	18	15	16	14	22	9
12	14	11	18	16	15	20	15	16	24	16
22	15	13	18	21	13	12	16	20	18	15
21	15	13	22	18	16	15	10	19	18	16
14	15	16	11	16	13	10	11	9	20	14
19	21	20	17	15	17	13	23	15	17	17
24	16	15	13	27	21	19	14	9	18	16
18	20	10	17	12	23	16	24	17	24	21

Таблица 1 - Исходные данные

Уровень значимости – 0.1

Критерий согласия – χ^2

$$\alpha_0 = \alpha_1$$

$$\sigma_0^2 = \sigma_1^2$$

Гипотезы $H_a - \alpha > \alpha_0, \sigma^2 \neq \sigma_0^2$

Решение

Размах варьирования

Расположим значения результатов эксперимента в порядке возрастания для получения вариационного ряда:

9	9	9	10	10	10	11	11	11	12	12
12	12	13	13	13	13	13	13	13	13	13
14	14	14	14	14	14	14	14	14	14	15
15	15	15	15	15	15	15	15	15	15	15
15	15	15	16	16	16	16	16	16	16	16
16	16	16	16	16	16	17	17	17	17	17
17	17	17	18	18	18	18	18	18	18	18
18	18	18	19	19	19	19	19	19	20	20
20	20	20	21	21	21	21	21	21	22	22
22	22	22	23	23	23	24	24	24	24	27

Таблица 2 - Данные в порядке возрастания

Определим минимальное и максимальное значения x , а также сам размах R как разность между крайними значениями вариант вариационного ряда

$$x_{min} = 9$$

$$x_{max} = 27$$

$$R = x_{max} - x_{min} = 18$$

Статический ряд распределения частот СВ X

Составим таблицу вариационного ряда и их распределения частот

Проценты анализа воды x_i , %	Частота m_i
9	3
10	3
11	3
12	4
13	9
14	10
15	15
16	14
17	8
18	11
19	6
20	5
21	6
22	5
23	3
24	4
27	1

Таблица 3 - Статический ряд распределения частот

Контроль: $\sum_i n_i = 110$

Интервальные статистические ряды частот и относительных частот

Рассчитаем длину интервала

$$h = \frac{R}{1 + 3.2 \lg n} = \frac{18}{1 + 3.2 \lg 110} = 2.571$$

Число интервалов k равно округленному до целого $\frac{R}{h}$

$$k = \frac{R}{h} = 1 + 3.2 \lg 110 \approx 7$$

Составим интервальные статистические ряды распределения частот и относительных частот

№	Интервал	Середина	n_i	w_i	N_i	w_i	$\frac{n_i}{h}$	$\frac{w_i}{h}$
1	[9; 11,57)	10,286	9	0,0818	9	0,0818	3,5	0,0318
2	[11,57; 14,14)	12,857	23	0,209	32	0,291	8,944	0,0813
3	[14,14; 16,71)	15,429	29	0,264	61	0,555	11,278	0,103
4	[16,71; 19,29)	18	25	0,227	86	0,782	9,722	0,0884
5	[19,29; 21,86)	20,571	11	0,1	97	0,882	4,278	0,0389
6	[21,86; 24,43)	23,143	12	0,109	109	0,991	4,667	0,0424
7	[24,43; 27]	25,714	1	0,00909	110	1	0,389	0,00354
Σ			110	1				

Полигон и гистограмма относительных частот

Строим полигон относительных частот (рисунок 1) по данным столбцов таблицы 4



Рисунок 1 - Полигон относительных частот

Строим гистограмму относительных частот используя как основания границы интервалов.

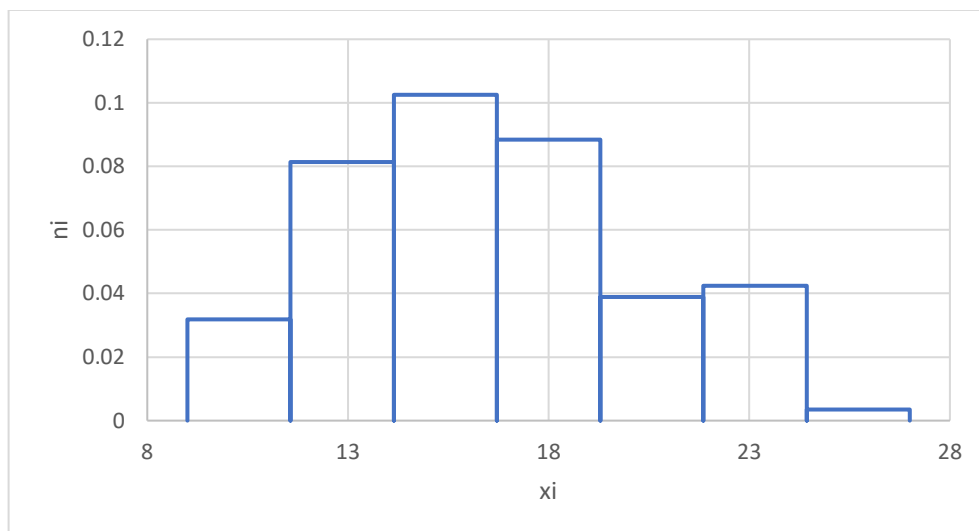


Рисунок 2 - Гистограмма относительных частот

Эмпирическая функция распределения и её график

По данным таблицы 4 найдём эмпирическую функцию распределения

$$F^*(x) = W(X < x) = \frac{n_x}{n}$$

X	≤ 9	11.571	14.14286	16.714	19.286	21.857	24.429	> 27
$F^*(x)$	0	0.082	0.291	0.554	0.782	0.882	0.991	1

Таблица 4 - Эмпирическая функция распределения

Построим график $F^*(x)$ с соединёнными точками, учитывая непрерывность функции

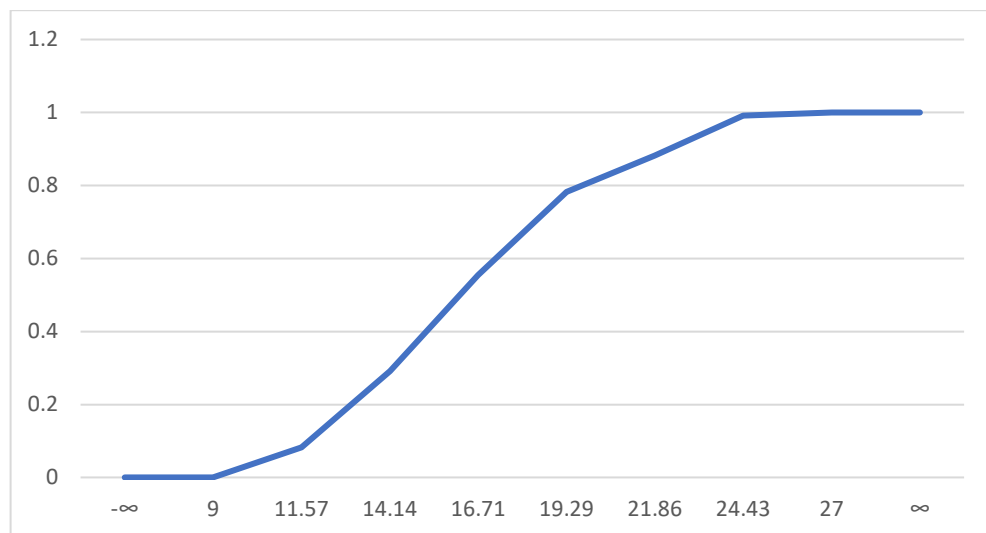


Рисунок 3 - Эмпирическая функция распределения

Выборочные значения числовых характеристик СВ X: математическое ожидание, дисперсия, среднее квадратическое отклонение

Найдём математическое ожидание по формуле

$$M_B(X) = \bar{X} = \frac{1}{n} \sum_{i=1}^k y_i n_i = \frac{1}{110} \sum_{i=1}^7 y_i n_i \approx 16.504$$

Тогда дисперсия находится как

$$D_B(X) = \frac{1}{n} \sum_{i=1}^k n_i (y_i - \bar{X})^2 = \frac{1}{110} \sum_{i=1}^7 n_i (y_i - 16.504)^2 \approx 13.99$$

Зная дисперсию выборки, можем найти среднее квадратическое отклонение

$$\sigma_B(X) = \sqrt{D_B(X)} = \sqrt{14.12} \approx 3.74$$

Выбор закона распределения СВ X по виду полигона и гистограммы относительных частот

Согласно рисункам 1 и 2, можем предположить о нормальном распределении СВ X.

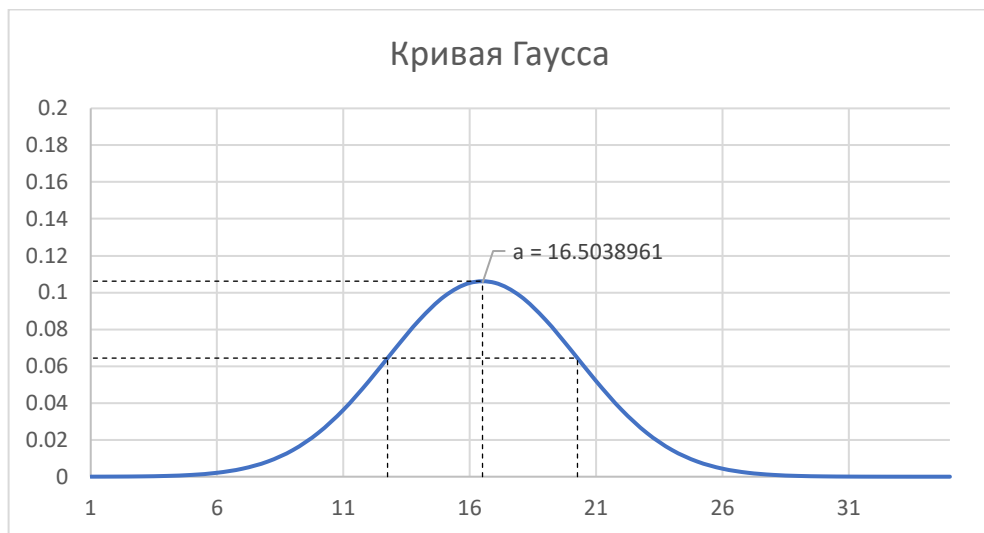


Рисунок 4 - Предположение о нормальном распределении

Точечная оценка параметров предполагаемого распределения и функции распределения и плотности вероятности СВ X

Плотность вероятности и функция распределения СВ X, распределенной по нормальному закону, имеют вид

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-a)^2}{2\sigma^2}} dt$$

Найдем точные оценки параметров $a = M(X)$ и $\sigma = \sigma(X)$ от нормального распределения

$$\tilde{a} = \bar{X} \approx 16.504$$

$$\sigma_B = \sqrt{\tilde{D}(X)} = \sqrt{\frac{n}{n-1} D_B(X)} = \sqrt{\frac{110}{109}} 13.99 \approx 3.76$$

Получается, плотность вероятности предполагаемого распределения $N(a, \sigma)$ имеет вид

$$f(x) = \frac{1}{3.76\sqrt{2\pi}} e^{-\frac{(x-16.504)^2}{28.28}}$$

И функция распределения

$$F(x) = \frac{1}{3.76\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-16.504)^2}{28.28}} dt$$

Проверка согласия эмпирической функции распределения с теоретической при помощи критерия согласия Пирсона

Вероятность попадания СВ X , распределенной по закону $N(a, \sigma)$, в интервал $(\alpha; \beta)$ найдём по формуле

$$p = P(\alpha < x < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right)$$

Где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$ – функция Лапласа

Вероятность попадания СВ X в первый частичный интервал $(-\infty; 11.57)$ равна (учтем, что $\Phi(-x) = -\Phi(x)$)

$$\begin{aligned} p_1 = P(-\infty < X < 11.57) &= \frac{1}{2} \left(\Phi\left(\frac{11.57 - 16.504}{3.76}\right) - \Phi\left(\frac{-\infty - 16.504}{3.76}\right) \right) \\ &= \frac{1}{2} \left(\Phi\left(\frac{11.57 - 16.504}{3.76}\right) - \Phi\left(\frac{-\infty - 16.504}{3.76}\right) \right) = \frac{1}{2} (-\Phi(1.31) + \Phi(\infty)) \\ &= \frac{1}{2} (-0.8098 + 1) = 0.0951 \end{aligned}$$

$$\begin{aligned} p_2 = P(11.57 < X < 14.14) &= \frac{1}{2} \left(\Phi\left(\frac{14.14 - 16.504}{3.76}\right) - \Phi\left(\frac{11.57 - 16.504}{3.76}\right) \right) \\ &= \frac{1}{2} (-\Phi(0.63) + \Phi(1.31)) = \frac{1}{2} (-0.47131 + 0.8098) = 0.169245 \end{aligned}$$

$$\begin{aligned} p_3 = P(14.14 < X < 16.71) &= \frac{1}{2} \left(\Phi\left(\frac{16.71 - 16.504}{3.76}\right) - \Phi\left(\frac{14.14 - 16.504}{3.76}\right) \right) \\ &= \frac{1}{2} (\Phi(0.05) + \Phi(0.62)) = \frac{1}{2} (0.03988 + 0.46474) = 0.25231 \end{aligned}$$

$$p_4 = P(16.71 < X < 19.29) = \frac{1}{2} \left(\Phi \left(\frac{19.29 - 16.504}{3.76} \right) - \Phi \left(\frac{16.71 - 16.504}{3.76} \right) \right) \\ = \frac{1}{2} (\Phi(0.74) - \Phi(0.05)) = \frac{1}{2} (0.5407 - 0.03988) = 0.25041$$

$$p_5 = P(19.29 < X < 21.86) = \frac{1}{2} \left(\Phi \left(\frac{21.86 - 16.504}{3.76} \right) - \Phi \left(\frac{19.29 - 16.504}{3.76} \right) \right) \\ = \frac{1}{2} (\Phi(1.42) - \Phi(0.74)) = \frac{1}{2} (0.84439 - 0.5407) = 0.151845$$

$$p_6 = P(21.86 < X < 24.43) = \frac{1}{2} \left(\Phi \left(\frac{24.43 - 16.504}{3.76} \right) - \Phi \left(\frac{21.86 - 16.504}{3.76} \right) \right) \\ = \frac{1}{2} (\Phi(2.11) - \Phi(1.42)) = \frac{1}{2} (0.96514 - 0.84439) = 0.060375$$

$$p_7 = P(24.43 < X < 27) = \frac{1}{2} \left(\Phi \left(\frac{27 - 16.504}{3.76} \right) - \Phi \left(\frac{24.43 - 16.504}{3.76} \right) \right) \\ = \frac{1}{2} (\Phi(2.79) - \Phi(2.11)) = \frac{1}{2} (0.99473 - 0.96514) = 0.014795$$

Интервал наблюдаемых значений СВ X	Частота n_i	p_i	np_i	$(n_i - np_i)^2$	$\frac{(n_i - np_i)^2}{np_i}$
$(-\infty; 11.57)$	9	0.0946	10.4117	1.993	0.1914
$[11.57; 14.14)$	23	0.1702	18.7267	18.2606	0.9751
$[14.14; 16.71)$	29	0.2574	28.3172	0.4662	0.0165
$[16.71; 19.29)$	25	0.2481	27.2929	5.2574	0.1926
$[19.29; 21.86)$	11	0.1524	16.7666	33.2542	1.9833
$[21.86; 24.43)$	12	0.0596	6.56244	29.5669	4.5055
$[24.43; \infty)$	1	0.0174	1.92227	0.8506	0.4425
Σ	110	1	110		8.3069

Таблица 5 - Определение значения критерия Пирсона

Находим табличное значение $\chi^2_{\text{табл.}}$ используя уровень значимости $\alpha = 0.1$ и число степеней свободы $\nu = k - r - 1 = 4$:

$$\chi^2_{0.1;4} = 7.7794$$

Нулевая гипотеза отклоняется, так как $8.3069 > 7.7794$

Интервальные оценки параметров распределения

Доверительный интервал, накрывающий математическое ожидание СВ

X с надежностью $P = 1 - \alpha$ имеет вид:

$$\tilde{a} - t \frac{\sigma_B}{\sqrt{n}} < a < \tilde{a} + t \frac{\sigma_B}{\sqrt{n}}$$

По таблице квантилей распределения Стьюдента по заданному уровню значимости

$$\alpha = 0.1$$

И числу степеней свободы $\nu = n - 1 = 109$ найдем квантиль

$$t = t_{\frac{\alpha}{2}, \nu} = t_{0.05; 109} = 1.659$$

Вычислим точность оценки

$$\delta = t \frac{\sigma_B}{\sqrt{n}} = 1.659 \cdot \frac{3.76}{\sqrt{110}} = 0.595$$

Искомый доверительный интервал для $M(X) = a$

$$16.504 - 0.595 < a \leq 16.504 + 0.595$$

$$15.909 < a \leq 17.099$$

Полуинтервал $(15.909; 17.099]$ накрывает неизвестное $M(X)$ с вероятностью $p = 0.9$

Доверительный интервал, накрывающий среднее квадратическое отклонение СВ X с надежностью $P = 1 - \alpha$:

$$\sqrt{\frac{n-1}{\chi^2_{\frac{\alpha}{2}, \nu}}} \sigma_B < \sigma < \sqrt{\frac{n-1}{\chi^2_{1-\frac{\alpha}{2}, \nu}}} \sigma_B$$

$$0.9 * 3.76 < \sigma < 1.13 * 3.76$$

$$3.384 < \sigma < 4.2488$$

Интервал $(3.384; 4.2488)$ накрывает неизвестное $\sigma(X)$ с вероятностью $p = 0.9$

Проверка нулевой гипотезы о математическом ожидании при альтернативной гипотезе

$$H_0: M(X) = a = a_0$$

$$H_a: a > a_0$$

Согласно варианту, имеем $a_0 = a_1 = 17.098$. Надо проверить нулевую гипотезу $H_0: a = a_0 = 17.098$ против альтернативной $H_a: a > a_0$

Воспользуемся правилом $U_{\text{набл.}} < U_{\text{крит.}}$ для подтверждения или отклонения нулевой гипотезы, где:

$$U_{\text{набл.}} = \frac{(\bar{x} - a_0)\sqrt{n}}{\sigma_{\text{в}}}$$

$$\Phi(U_{\text{крит.}}) = \frac{1 - 2\alpha}{2}$$

Используя уровень значимости $\alpha = 0.1$ находим

$$\Phi(U_{\text{крит.}}) = \frac{1 - 2 * 0.1}{2} = 0.4, \quad U_{\text{крит.}} = 1.2815$$

$$U_{\text{набл.}} = \frac{(16.504 - 17.098)\sqrt{110}}{3.76} \approx -1.6589$$

Так как $U_{\text{набл.}} < U_{\text{крит.}}$, то нулевую гипотезу принимаем.

Проверка нулевой гипотезы о дисперсии против альтернативной гипотезы

$$H_0: D(X) = \sigma^2 = \sigma_0^2$$

$$H_a: \sigma^2 \neq \sigma_0^2$$

Согласно варианту, имеем $\sigma_0 = \sigma_1 = 4.233$. Надо проверить нулевую гипотезу $H_0: \sigma^2 = \sigma_0^2 = 17.917$ против альтернативной $H_a: \sigma^2 \neq \sigma_0^2$

Воспользуемся правилом $\chi_{\text{лев.кр.}}^2 < \chi_{\text{набл.}}^2 < \chi_{\text{прав.кр.}}^2$ для подтверждения или отклонения нулевой гипотезы, где:

$$\chi_{\text{набл.}}^2 = \frac{(n - 1)\sigma_{\text{в}}^2}{\sigma_0^2}$$

$$\chi_{\text{лев.кр.}}^2 = \chi_{1 - \frac{\alpha}{2}, v}^2$$

$$\chi_{\text{прав.кр.}}^2 = \chi_{\frac{\alpha}{2}, v}^2$$

Так же воспользуемся равенством Уилсона-Гильферти, применяемое при степенях свободы свыше 30:

$$\chi_{\alpha,v}^2 = v \left(1 - \frac{2}{9v} + Z_{\alpha} \sqrt{\frac{2}{9v}} \right)^3$$

где z_{α} находится через функцию Лапласа $\Phi(z_{\alpha}) = \frac{1-2\alpha}{2}$

Подставляя нужные значения получаем:

$$\chi_{\text{набл.}}^2 = \frac{109 \cdot 3.758^2}{4.233^2} \approx 85.90300773$$

$$\Phi\left(z_{1-\frac{\alpha}{2}}\right) = \Phi(z_{0.95}) = \frac{1 - 2 \cdot 0.95}{2}, z_{\alpha} = -1.645$$

$$\Phi\left(z_{\frac{\alpha}{2}}\right) = \Phi(z_{0.05}) = \frac{1 - 2 \cdot 0.05}{2}, z_{\alpha} = 1.645$$

$$\chi_{\text{лев.кр.}}^2 = \chi_{0.95,109}^2 = 109 \left(1 - \frac{2}{9 \cdot 109} - 1.645 \sqrt{\frac{2}{9 \cdot 109}} \right)^3 \approx 85.9030087$$

$$\chi_{\text{прав.кр.}}^2 = \chi_{0.05,109}^2 = 109 \left(1 - \frac{2}{9 \cdot 109} + 1.645 \sqrt{\frac{2}{9 \cdot 109}} \right)^3 \approx 134.366$$

$\chi_{\text{набл.}}^2 < \chi_{\text{лев.кр.}}^2 < \chi_{\text{прав.кр.}}^2$ поэтому гипотезу H_0 отвергаем, принимаем гипотезу $H_a: \sigma^2 \neq 17.917$

МНК

Задание 1

В таблице приведены данные о сроке службы колеса вагона в годах (x) и износа толщины обода колеса, (y , мм).

x_i	0,5	1	1,5	2	2,5	3	3,5	4	4,5	5
y_i	0,4	0,7	1,2	1,7	1,9	2,2	2,6	3	3,5	3,8

Таблица 6 - Значения задачи 1

В предположении, что между x и y существует линейная зависимость, определить параметры линейной регрессии $y = kx + b$ методом наименьших квадратов. Сделать выводы об износе толщины обода колеса через 5,5 лет.

Задание 2

В боковой стенке высокого цилиндрического бака у самого дна закреплен кран. После его открытия вода начинает вытекать из бака. В таблице приведены данные об изменении высоты (h , м) и времени (t , мин).

t_i	1	2	4	6	8	10	12	15	18	20
h_i	3.6	3.2	2.57	1.95	1.45	1.09	0.9	0.6	0.3	0.1

Таблица 7 - Значения задачи 2

В предположении, что между t и h существует квадратичная зависимость, определить параметры регрессии $h = a_2 t^2 + a_1 t + a_0$ методом наименьших квадратов. Спрогнозировать время, когда бак опустеет.

Дополнительное задание

В задании 1 определить также параметры квадратичной регрессии $y = a_2 x^2 + a_1 x + a_0$, вычислить сумму квадратов отклонений, сравнить с результатом полученным в задании 1 и сделать вывод.

В задании 2 определить также параметры линейной регрессии $y = kx + b$, вычислить сумму квадратов отклонений, сравнить с результатом полученным в задании 2 и сделать вывод.

Задание 1

Составляем таблицу 8 и заполняем столбцы N , x_i , y_i , x_i^2 , $x_i y_i$.

Составляем систему уравнений для нахождения коэффициентов k и b .

$$\begin{cases} k \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n y_i x_i \\ k \sum_{i=1}^n x_i + bn = \sum_{i=1}^n y_i \end{cases}$$

$$\begin{cases} 96.25k + 27.5b = 73.3 \\ 27.5k + 10b = 21 \end{cases}$$

Найдём неизвестные k и b по формулам Крамера:

$$k = \frac{\Delta_1}{\Delta}, b = \frac{\Delta_2}{\Delta}$$

$$\Delta = \begin{vmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{vmatrix} = \sum_{i=1}^n x_i^2 \cdot n - \left(\sum_{i=1}^n x_i \right)^2$$

$$\Delta_1 = \begin{vmatrix} \sum_{i=1}^n y_i x_i & \sum_{i=1}^n x_i \\ \sum_{i=1}^n y_i & n \end{vmatrix} = \sum_{i=1}^n y_i x_i \cdot n - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i$$

$$\Delta_2 = \begin{vmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n y_i x_i \\ \sum_{i=1}^n y_i & \sum_{i=1}^n y_i \end{vmatrix} = \sum_{i=1}^n x_i^2 \cdot \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i x_i$$

$$\Delta = \begin{vmatrix} 96.25 & 27.5 \\ 27.5 & 10 \end{vmatrix} = 206.25$$

$$\Delta_1 = \begin{vmatrix} 73.3 & 27.5 \\ 21 & 10 \end{vmatrix} = 155.5$$

$$\Delta_2 = \begin{vmatrix} 96.25 & 73.3 \\ 27.5 & 21 \end{vmatrix} = 5.5$$

Окончательно получим

$$k = 0.7539, b = 0.0267$$

То есть уравнение

$$y = 0.7539x + 0.0267$$

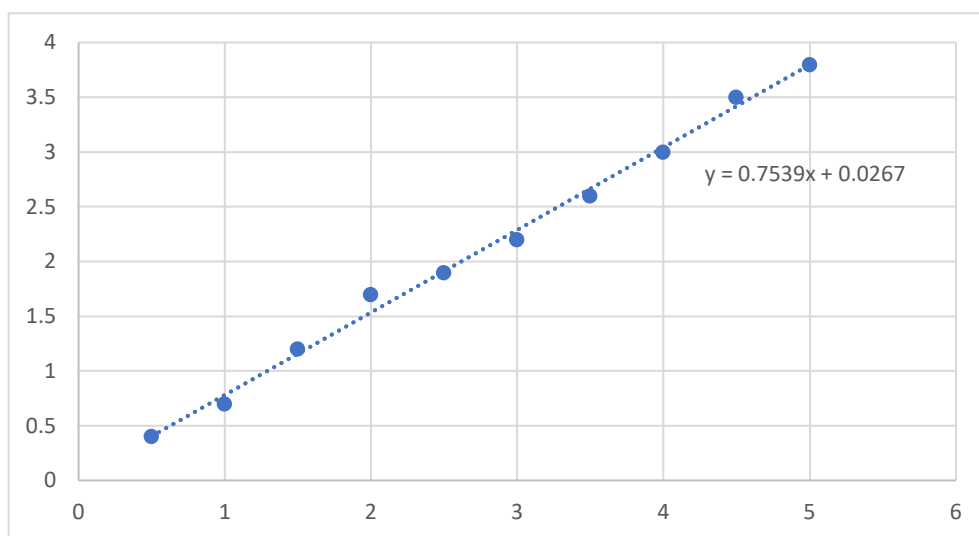


Рисунок 5 - Линейная регрессия задания 1

N	x_i	y_i	x_i^2	$x_i y_i$	$y_{\text{расч. } i}$	δ_i	δ_i^2
1	0.5	0.4	0.25	0.2	0.4036	-0.0036	0.000013
2	1	0.7	1	0.7	0.7806	-0.0806	0.006497
3	1.5	1.2	2.25	1.8	1.1576	0.0424	0.001800
4	2	1.7	4	3.4	1.5345	0.1655	0.027375
5	2.5	1.9	6.25	4.75	1.9115	-0.0115	0.000133
6	3	2.2	9	6.6	2.2885	-0.0885	0.007830
7	3.5	2.6	12.25	9.1	2.6655	-0.0655	0.004284
8	4	3	16	12	3.0424	-0.0424	0.001800
9	4.5	3.5	20.25	15.75	3.4194	0.0806	0.006497
10	5	3.8	25	19	3.7964	0.0036	0.000013
Σ	27.5	21	96.25	73.3			0.056242

Таблица 8 - Данные и промежуточные значения задания 1

Износ толщины обода колеса при 5.5 лет составляет:

$$0.7539 \cdot 5.5 + 0.0267 = 4.1733 \text{ мм}$$

Дополнительное задание 1

Составляем таблицу 9 и заполняем столбцы N , x_i , y_i , x_i^2 , x_i^3 , x_i^4 , $x_i y_i$, $x_i^2 y_i$

Составляем систему уравнений для нахождения коэффициентов a_2 , a_1 и a_0 .

N	x_i	y_i	x_i^2	x_i^3	x_i^4	$x_i y_i$	$x_i^2 y_i$	$y_{\text{расч. } i}^2$	δ_i	δ_i^2
1	0.5	0.4	0.25	0.125	0.0625	0.2	0.1	0.4082	-0.0082	0.00007
2	1	0.7	1	1	1	0.7	0.7	0.7821	-0.0821	0.00674
3	1.5	1.2	2.25	3.375	5.0625	1.8	2.7	1.1568	0.0432	0.00186
4	2	1.7	4	8	16	3.4	6.8	1.5323	0.1677	0.02813
5	2.5	1.9	6.25	15.625	39.0625	4.75	11.875	1.9085	-0.0085	0.00007
6	3	2.2	9	27	81	6.6	19.8	2.2855	-0.0855	0.00730
7	3.5	2.6	12.25	42.875	150.0625	9.1	31.85	2.6632	-0.0632	0.00399
8	4	3	16	64	256	12	48	3.0417	-0.0417	0.00174
9	4.5	3.5	20.25	91.125	410.0625	15.75	70.875	3.4209	0.0791	0.00626
10	5	3.8	25	125	625	19	95	3.8009	-0.0009	0.00000
Σ	27.5	21	96.25	378.125	1583.3125	73.3	287.7			0.05617

Таблица 9 - Данные и промежуточные значения дополнительного задания 1

$$\begin{cases} n \cdot a_0 + \sum_{i=1}^n x_i \cdot a_1 + \sum_{i=1}^n x_i^2 \cdot a_2 = \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i \cdot a_0 + \sum_{i=1}^n x_i^2 \cdot a_1 + \sum_{i=1}^n x_i^3 \cdot a_2 = \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n x_i^2 \cdot a_0 + \sum_{i=1}^n x_i^3 \cdot a_1 + \sum_{i=1}^n x_i^4 \cdot a_2 = \sum_{i=1}^n x_i^2 y_i \end{cases}$$

$$\begin{cases} 10 \cdot a_0 + 27.5 \cdot a_1 + 96.25 \cdot a_2 = 21 \\ 27.5 \cdot a_0 + 96.25 \cdot a_1 + 378.125 \cdot a_2 = 73.3 \\ 96.25 \cdot a_0 + 378.125 \cdot a_1 + 1583.313 \cdot a_2 = 287.7 \end{cases}$$

Решим систему как матричное уравнение

$$X = A^{-1}B$$

$$A = \begin{pmatrix} 10 & 27.5 & 96.25 \\ 27.5 & 96.25 & 378.125 \\ 96.25 & 378.125 & 1583.313 \end{pmatrix} B = \begin{pmatrix} 21 \\ 73.3 \\ 287.7 \end{pmatrix}$$

$$A^{-1} = \begin{pmatrix} 1.383 & -1.05 & 0.167 \\ -1.05 & 0.965 & -0.167 \\ 0.167 & -0.167 & 0.03 \end{pmatrix} X = \begin{pmatrix} 0.035 \\ 0.7456 \\ 0.0015 \end{pmatrix}$$

Окончательно получим

$$a_0 = 0.035, a_1 = 0.7456, a_2 = 0.0015$$

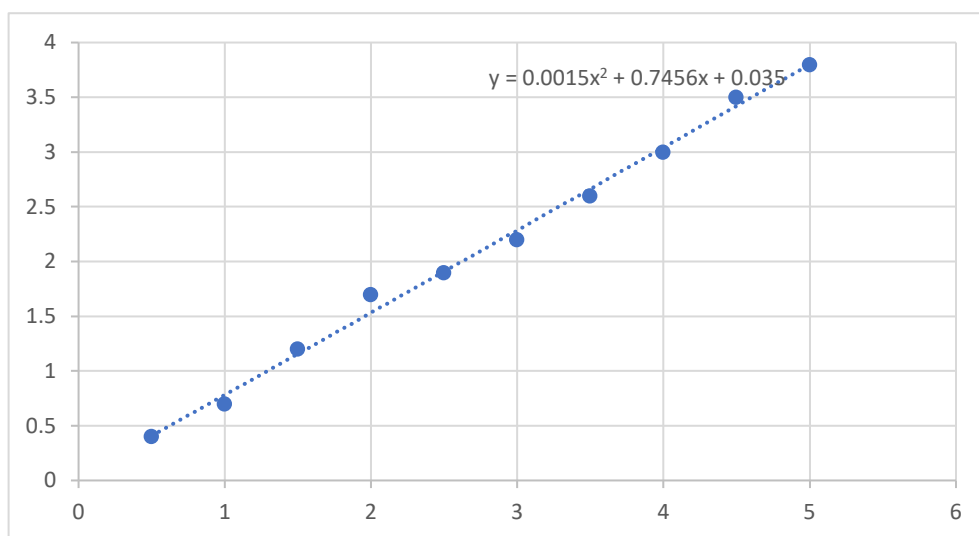


Рисунок 6 - Квадратичная регрессия задания 1

δ_i^2 полученный квадратичной регрессией оказался не на много меньше δ_i^2 полученный линейной регрессией.

$$0.056167 < 0.056242$$

Износ толщины обода колеса при 5.5 лет составляет:

$$0.0015 \cdot 5.5^2 + 0.7456 \cdot 5.5 + 0.035 = 4.1817 \text{ мм}$$

Задание 2

Составляем таблицу 10 и заполняем столбцы N , x_i , y_i , x_i^2 , x_i^3 , x_i^4 , $x_i y_i$, $x_i^2 y_i$

Составляем систему уравнений для нахождения коэффициентов a_2 , a_1 и a_0 .

N	x_i	y_i	x_i^2	x_i^3	x_i^4	$x_i y_i$	$x_i^2 y_i$	$y_{\text{расч. } i}^2$	δ_i	δ_i^2
1	1	3.6	1	1	1	3.6	3.6	3.5305	0.0695	0.00483
2	2	3.2	4	8	16	6.4	12.8	3.1936	0.0064	0.00004
3	4	2.57	16	64	256	10.28	41.12	2.5740	-0.0040	0.00002
4	6	1.95	36	216	1296	11.7	70.2	2.0265	-0.0765	0.00585
5	8	1.45	64	512	4096	11.6	92.8	1.5510	-0.1010	0.01021
6	10	1.09	100	1000	10000	10.9	109	1.1476	-0.0576	0.00332
7	12	0.9	144	1728	20736	10.8	129.6	0.8163	0.0837	0.00701
8	15	0.6	225	3375	50625	9	135	0.4544	0.1456	0.02121
9	18	0.3	324	5832	104976	5.4	97.2	0.2546	0.0454	0.00206
10	20	0.1	400	8000	160000	2	40	0.2115	-0.1115	0.01243
Σ	96	15.76	1314	20736	352002	81.68	731.32			0.06699

Таблица 10 - Данные и промежуточные значения задания 2

$$\begin{cases} 10 \cdot a_0 + 96 \cdot a_1 + 1314 \cdot a_2 = 15.76 \\ 96 \cdot a_0 + 1314 \cdot a_1 + 20736 \cdot a_2 = 81.68 \\ 1314 \cdot a_0 + 20736 \cdot a_1 + 352002 \cdot a_2 = 731.32 \end{cases}$$

Решим систему как матричное уравнение

$$X = A^{-1}B$$

$$A = \begin{pmatrix} 10 & 96 & 96.25 \\ 96 & 1314 & 20736 \\ 1314 & 20736 & 352002 \end{pmatrix} B = \begin{pmatrix} 15.76 \\ 81.68 \\ 731.32 \end{pmatrix}$$

$$A^{-1} = \begin{pmatrix} 0.737 & -0.148 & 0.0059 \\ -0.148 & 0.0406 & -0.0018 \\ 0.0059 & -0.0018 & 0.000089 \end{pmatrix} X = \begin{pmatrix} 3.885 \\ -0.3638 \\ 0.009 \end{pmatrix}$$

Окончательно получим

$$a_0 = 3.885, a_1 = -0.3638, a_2 = 0.009$$

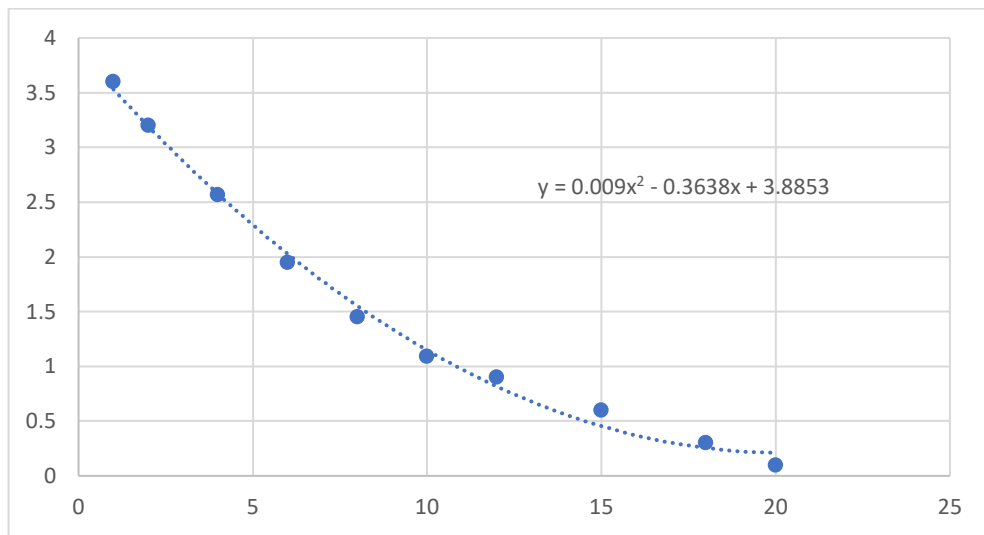


Рисунок 7 - Квадратичная регрессия задания 2

Полученная функция:

$$y = 0.009x^2 - 0.3638x + 3.885$$

Найдем время, когда бак опустеет

$$y = 0 \Rightarrow 0.009x^2 - 0.3638x + 3.885 = 0$$

Решим квадратное уравнение:

$$D = b^2 - 4ac = (-0.3638)^2 - 4 \cdot 0.009 \cdot 3.885 = -0.00761$$

Дискриминант отрицательный, следовательно нет действительных решений.

Дополнительное задание 2

Составляем таблицу 11 и заполняем столбцы N , x_i , y_i , x_i^2 , $x_i y_i$.

Составляем систему уравнений для нахождения коэффициентов k и b .

$$\begin{cases} k \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n y_i x_i \\ k \sum_{i=1}^n x_i + bn = \sum_{i=1}^n y_i \end{cases}$$

$$\begin{cases} 1314k + 96b = 81.68 \\ 96k + 10b = 15.76 \end{cases}$$

Найдём неизвестные k и b по формулам Крамера:

$$k = \frac{\Delta_1}{\Delta}, b = \frac{\Delta_2}{\Delta}$$

$$\Delta = \begin{vmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{vmatrix} = \sum_{i=1}^n x_i^2 \cdot n - \left(\sum_{i=1}^n x_i \right)^2$$

$$\Delta_1 = \begin{vmatrix} \sum_{i=1}^n y_i x_i & \sum_{i=1}^n x_i \\ \sum_{i=1}^n y_i & n \end{vmatrix} = \sum_{i=1}^n y_i x_i \cdot n - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i$$

$$\Delta_2 = \begin{vmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n y_i x_i \\ \sum_{i=1}^n y_i & \sum_{i=1}^n y_i \end{vmatrix} = \sum_{i=1}^n x_i^2 \cdot \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i x_i$$

$$\Delta = \begin{vmatrix} 1314 & 96 \\ 96 & 10 \end{vmatrix} = 3924$$

$$\Delta_1 = \begin{vmatrix} 81.68 & 96 \\ 15.76 & 10 \end{vmatrix} = -696.16$$

$$\Delta_2 = \begin{vmatrix} 1314 & 81.68 \\ 96 & 15.76 \end{vmatrix} = 12867.36$$

Окончательно получим

$$k = -0.17741, b = 3.27914$$

То есть уравнение

$$y = -0.17741x + 3.27914$$

δ_t^2 полученный линейной регрессией оказался на много больше δ_t^2 полученный квадратичной регрессией.

$$0.979609 > 0.066986$$

Найдем время, когда бак опустеет:

$$y = 0$$

$$x = \frac{3.27914}{-0.17741} = 18.48334$$

Как видно, такое значение не может быть верным так как согласно реальным данным даже при $t = 20$, бак не пустой.