

## Вариант 22.

**Задача 1. (2 балла)** Имеется выборка:  $X = \{7, 3, 5, 7, 8, 11, 8, 7\}$ .

Необходимо:

1. Построить дискретный вариационный ряд (с указанием  $m, m^c, w, w^c$ );
2. Вычислить средние порядка  $p = -1,2$ ; проверить правило мажорантности средних.
3. Найти размах, моду, медиану;
4. Вычислить дисперсию и относительный коэффициент вариации  $V_\sigma$ ; сделать вывод.

**Задача 2. (2 балла)** Представить в виде интервального вариационного ряда выборку

$$X = \{5, 1, 3, 15, 1, 19, 0, 5, 7, 1, 8, 12, 2, 6, 8, 3, 25, 3, 13, 8\}.$$

Необходимо:

- 1) Число интервалов  $k$  определить по формуле Стерджеса;
- 2) Построить кумулятивную кривую относительных частот;
- 3) Вычислить среднее арифметическое и дисперсию методом условных моментов (относительно  $Mo$ );
- 4) Найти моду и медиану;

**Задача 3. (2 балла)** Пусть выборка  $X$  разбита на две группы:  $X = X_1 \cup X_2$ , где

$$X_1 = \{9, 1, 5, 2, 6, 7\}, \quad X_2 = \{1, 3, 1, 4, 5, 6, 4, 2, 1\}.$$

Найти:

1. Средние арифметические выборок  $X_1, X_2$  и  $X$ ;
2. Дисперсии выборок  $X_1, X_2$ ;
3. Среднюю из внутригрупповых дисперсий;
4. Межгрупповую дисперсию;
5. Общую дисперсию выборки  $X$ .

**Задача 4. (2 балла)** Имеются данные об уровне инфляции и курсе доллара:

Период времени	1	2	3	4	5	6	7
Изменение темпа инфляции ( $X, \%$ )	10	-3	4	5	-2	4	3
Изменение курса доллара ( $Y, \%$ )	4	0,9	1,7	1,5	-1,5	1,4	-1

- 1) Установить тесноту связи между признаками  $x$  и  $y$ ;
- 2) Методом наименьших квадратов найти оценки параметров  $a, b$  уравнения линейной модели тренда;
- 3) Вычислить коэффициент детерминации; сделать вывод;

**Задача 5. (2 балла)** Имеются следующие данные о производстве электроэнергии (млрд.квт.ч.):

Годы	2017	2018	2019	2020	2021	2022	2023
Производство электроэнергии	33	39	38	37	33	32	30

Необходимо:

- 1) Применить метод, основанный на медиане, для проверки наличия тренда;
- 2) Вычислить коэффициенты модели тренда;
- 3) Сделать прогноз на 2026 год.

Борис Тимофеев 1 группа 22 Вариант

1. Исходная выборка

$$X = \{7, 3, 5, 7, 8, 8, 8, 7\}$$

1) Построим вариацион. ряд.

1) Упорядочим зн-я по возрастанию

3 5 7 7 7 8 8 11

2) Вычислим абс. частоты (число повторений зн-я)  $m_i$  в выборке

1 1 3 3 3 2 2 1

3) Вычислим отн. частоты  $w_i$

Т.к.  $n = 8$  - кол-во зн-тов, то отн. частоты

$\frac{1}{8} \frac{1}{8} \frac{3}{8} \frac{3}{8} \frac{3}{8} \frac{2}{8} \frac{2}{8} \frac{1}{8}$

Таким образом, убрав дубликаты, получим дискр. вариацион. ряд (ДВР)

$x_i$	3	5	7	8	11	
$m_i$	<del>1</del>	<del>1</del>	<del>3</del>	<del>2</del>	<del>1</del>	$\sum m_i = 8$
$w_i$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	$\sum w_i = 1$
=						
$m_i^{cum}$	1	2	5	7	8	
$w_i^{cum}$	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{5}{8}$	$\frac{7}{8}$	1	

Добавим к числ. накопительные частоты  $m_i^{cum}, w_i^{cum} \Rightarrow$

2) Вычислим среднее порядков  $p = \{-1, 0, 1, 2\}$

$$\overline{x}_2 = \sqrt{\frac{1}{8} \sum_{i=1}^8 x_i^2} = \left( \frac{1}{8} \cdot (9 + 25 + 49 + 49 + 49 + 64 + 64 + 121) \right)^{\frac{1}{2}} \approx 7.33$$

$$\overline{x}_1 = \frac{1}{8} \sum_{i=1}^8 x_i = \frac{1}{8} (3 + 5 + 7 \cdot 3 + 8 \cdot 2 + 11) = 7$$

$$\overline{x}_0 = \sqrt[8]{3 \cdot 5 \cdot 7^3 \cdot 8^2 \cdot 11} \approx 6.60$$

1



$$\bar{x}_{-1} = \frac{8}{\sum_{i=1}^6 \frac{1}{x_i}} = \frac{8}{\left(\frac{1}{3} + \frac{1}{5} + \frac{3}{7} + \frac{2}{8} + \frac{1}{11}\right)} \approx 6.15$$

Уточн. значение

$$\bar{x}_{-1} < \bar{x}_0 < \bar{x}_1 < \bar{x}_2 \Rightarrow \text{последовательность монотонно возрастает}$$

$$6.15 < 6.60 < 7 < 7.33$$

3) Найти размах, моду и медиану:

$$1) R = x_{\max} - x_{\min} = x_{(8)} - x_{(1)} = 8 - 3 = 5$$

2) Mode = 7, т.к. 7 встречается чаще всего (3 раза)

$$3) \text{Median} = 3 \ 5 \ 7 \ 7 \ 7 \ 8 \ 8 \ 11 \Rightarrow$$

$$\text{Median} = \frac{7+7}{2} = 7$$

4) Найти дисперсию

$$\sigma^2 = \frac{1}{8} \left( (3-7)^2 + (5-7)^2 + (7-7)^2 \cdot 3 + (8-7)^2 \cdot 2 + (11-7)^2 \right) =$$

$$= 4.75$$

Относ. коэфф. вариации равен

$$V_{\sigma} = \frac{\sigma}{\bar{x}} = \frac{\sqrt{4.75}}{7} \cdot 100\% = 31.1\%$$

Таким образом, поскольку значение относительного коэффициента вариации меньше, чем  $V_{\sigma} < 33\%$ .

Поскольку Mode = Median =  $\bar{x} \Rightarrow$  распределение симметрично

2. Пусть  $X = \{5, 8, 3, 15, 1, 19, 0, 5, 7, 1, 8, 12, 2, 6, 8, 3, 25, 3, 13, 8\}$

1) Вычислим  $k$  по ф-ле Гершмеса:

$k = 1 + \lfloor \log_2 n \rfloor$ , где  $n = 20$  - размер выборки  $\Rightarrow$

$$k = 1 + \lfloor \log_2 20 \rfloor = 5$$

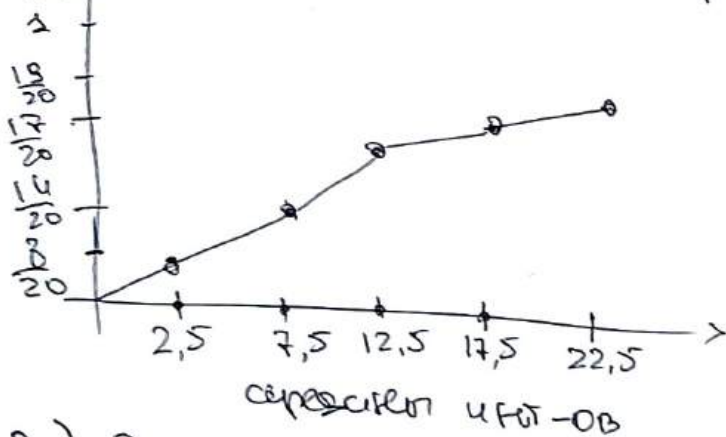
Построим ЧБР

Поскольку  $R = x_{\max} - x_{\min} = 25 - 0 = 25$ , то

$h = \frac{25}{5} = 5 \Rightarrow$  получаем интервалы с частотами:

	$[0; 5)$	$[5; 10)$	$[10; 15)$	$[15; 20)$	$[20; 25)$	
$m_i$	8	6	3	2	1	} ЧБР
$w_i$	$\frac{8}{20}$	$\frac{6}{20}$	$\frac{3}{20}$	$\frac{2}{20}$	$\frac{1}{20}$	
$M_i^{\text{cum}}$	8	14	17	19	20	
$W_i^{\text{cum}}$	$\frac{8}{20}$	$\frac{14}{20}$	$\frac{17}{20}$	$\frac{19}{20}$	1	

2) Построим кумулянт. крив. отн. частот, где в каждой точке берем середину отрезков (интервалов)



3) Вычислим среднее и дисперсию методом усн. моментов по mode

$$\bar{x} = x_{\text{усн}} + h \frac{\sum_{i=1}^k z_i m_i}{n} = x_{\text{усн}} + h \mu_{1,0}$$

любой интервал:  $[0; 5)$  (частота канторовская)  $\Rightarrow$

$$x_{\text{med}} = 2.5$$

$$\mu_{3,0} = \frac{1}{n} \sum_{i=1}^n (x_i - x_{\text{med}}) = \frac{1}{20} ((5-2.5) + (1-2.5) + \dots) = 4.75 \Rightarrow$$

$$\bar{x} = 2.5 + 5 \cdot 4.75 = 26.25$$

Второй шаг

$$\sigma^2 = h^2 (\mu_{2,0} - \mu_{3,0}^2)$$

$$\mu_{2,0} = \frac{1}{20} ((5-2.5)^2 + (1-2.5)^2 + \dots) \approx 64.25 \Rightarrow$$

$$\sigma^2 = 25 \cdot (64.25 - 22.5625) = 1.042$$

$$4) \mu_c = x_{\text{med}} + h \frac{\frac{1}{2} - W_{\text{med}} - 1}{W_{\text{med}}} = 10 + 5 \cdot \frac{\frac{1}{2} - \frac{6}{20}}{\frac{3}{20}} = 16.7$$

$$\begin{aligned} \mu_b &= x_{\text{med}} + h \frac{n_{\text{med}} - n_{\text{med}-1}}{2n_{\text{med}} - n_{\text{med}-1} - n_{\text{med}+1}} \\ &= 10 + 5 \cdot \frac{\frac{3}{20}}{2 \cdot \frac{3}{20} - \frac{6}{20}} = 4 \end{aligned}$$

3. Пример

$$X_1 = \{9, 8, 5, 2, 6, 7\}$$

$$X_2 = \{8, 3, 1, 4, 5, 6, 4, 2, 1\}$$

1) Найти средние значения

$$\bar{x}_1 = \frac{1}{6} (9 + 8 + 5 + 2 + 6 + 7) = \frac{30}{6} = 5$$

$$\bar{x}_2 = \frac{1}{9} (8 + 3 + 1 + 4 + 5 + 6 + 4 + 2 + 1) = \frac{27}{9} = 3$$

$$\bar{x} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n} = \frac{6 \cdot 5 + 9 \cdot 3}{15} = \frac{30 + 27}{15} = 3.8$$



2) Кадмеев. Реч. Морфология

$$\sigma_1^2 = \frac{1}{6} ((9-5)^2 + \dots) \approx 7,7$$

$$\sigma_2^2 = \frac{1}{9} ((1-3)^2 + \dots) \approx 3,1$$

$$3) \sigma_{12}^2 = \frac{\sigma_1^2 n_1 + \sigma_2^2 n_2}{n} = 4,94$$

$$4) S^2 = \frac{n_1 (\bar{x}_1 - \bar{x})^2 + n_2 (\bar{x}_2 - \bar{x})^2}{n} = 0,96$$

$$5) \sigma^2 = \sigma_{12}^2 + S^2 = 5,9 - \text{общая дисперсия}$$

4.

x	10	-3	4	5	-2	4	3
y	4	0,9	1,7	1,5	-1,5	1,4	-1

Оценим поPearson коэффициенту

$$r_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} - \text{к.к. Пирсона}$$

$$\bar{x} = \frac{1}{7} (10 - 3 + 4 + 5 - 2 + 4 + 3) = 3$$

$$\bar{y} = \frac{1}{7} (4 + 0,9 + 1,7 + 1,5 - 1,5 + 1,4 - 1) = 0,8$$

$$\sigma_x^2 = \frac{1}{7} ((10-3)^2 + \dots) = 16,5 \quad \sigma_x^2 = \overline{x^2} - (\bar{x})^2$$

$$\sigma_y^2 =$$

$$\sigma_y^2 = \overline{y^2} - (\bar{y})^2$$

$$\overline{x^2} = 25,57 \Rightarrow \sigma_x^2 = 18,57$$

$$\overline{y^2} = 3,88 \Rightarrow \sigma_y^2 = 2,88$$

$$\sigma_{xy} = \overline{xy} - \bar{x} \cdot \bar{y}$$

$$\overline{xy} = \frac{1}{n} \sum: \frac{1}{7} (10 \cdot 4 - 3 \cdot 0,9 + \dots) = 8,17$$

$$\Rightarrow \sigma_{xy} = 5,17$$

5

$r_{xy} = 0.74 \Rightarrow$  существует сильная положительная линейная зависимость между признаками

2) По мин

$$y = ax + b, \quad a = \frac{\sigma_{xy}}{\sigma_x^2} \approx 0.31, \quad b = \bar{y} - a\bar{x} = 0.07 \Rightarrow$$

Ур-е линейной регрессии

$$y = 0.31x + 0.07$$

3) Коэфф. детерминации  $R^2 = r_{xy}^2 = 0.5476$ , т.е.

54% дисперсии нек. ряда объясняется моделью, поскольку  $R^2 < 0.7$ , то модель не совсем адекватна.

Но, с другой стороны, т.н.

$$\sqrt{\sigma_y^2(1 - r_{xy}^2)} = 0.14 < 1.13 = 0.67\sigma_y, \text{ то}$$

коэф. и критерии, но мы можем считать модель достаточно хорошей для прогнозирования,

5.

2017	2018	2019	2020	2021	2022	2023
33	35	36	39	33	32	30

4) Определим ме:

$$T = 7$$

30 32 33 | 33 | 37 38 39

ме = 33

-3	-2	-1	0	1	2	3
2017	2018	2019	2020	2021	2022	2023
+	-	-	+	+	-	-

 $\Rightarrow$ 

$$V(T) = 3, \quad L(T) = 6$$

$$3 = V(T) \stackrel{?}{>} \frac{1}{2}(T+1 - 1.96\sqrt{T-1}) = \frac{1}{2}(8 - 1.96\sqrt{6}) = 1.4$$

Верно  
 $6 = L(T) < \log_2(T+3) = \log_2 8 = 3 \Rightarrow$  критерий выполняется  
неверно



2) No error

$$a = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{\overline{ty} - \bar{t}\bar{y}}{\sigma_t^2}$$

$$\bar{t} = 0$$

$$\bar{y} = 34,57$$

$$\overline{ty} = \frac{1}{7} \cdot (-3 \cdot 33 + \dots) = -4$$

$\Rightarrow$

$$\sigma_t^2 = \frac{1}{7} (9 + 4 + 1 + 0 + 1 + 4 + 9) = 4$$

$$a = -1$$

$$b = \bar{y} - a\bar{t} = 34,57 \Rightarrow$$

$$y = -t + 34,57 - \text{линейная модель тренда}$$

$$3) y_{2026} = -6 + 34,57 = 28,57 \Rightarrow$$

В 2026 году производство

должно быть равно 28,57

7