

Московский университет им. С.Ю. Витте
Расчетно-аналитическая работа

Статистика

Вариант 11

Задание

- 1) Скопировать данные своего варианта.
- 2) Ранжировать ряд данных сортировкой по значениям от минимального к максимальному.
- 3) Рассчитать количество интервалов по формуле Стерджеса, округлив вверх до целых единиц.
- 4) Рассчитать величину интервала h , округлить до десятков.
- 5) Рассчитать границы интервалов:
- 6) Подсчитать количество единиц совокупности, принадлежащих каждому из интервалов.
- 7) Построить интервальный вариационный ряд в виде таблицы
- 8) Построить гистограмму распределения для интервалов и полигон распределения для вариантов, кумуляту.
- 9) Вычислить среднее арифметическое, моду, медиану, квартили, децили.
- 10) Вычислить показатели вариации: R , d_{cp} , s^2 , s , V_r , V_d , V . Вычислить асимметрию и эксцесс.
- 11) Сделать вывод об однородности вариационного ряда, о симметричности и остро- или плоско-вершинности распределения.

Таблица 1 – Исходные данные

48	72	189	57	86	226
151	143	45	181	171	54
49	98	101	58	117	121
65	35	164	78	42	196

85	178	89	102	213	106
53	74	105	63	88	126
63	121	81	75	145	97
25	86	122	30	103	146
68	105	154	81	126	184
111	39	41	133	46	49
58					

Решение:

1. Ранжируем ряд распределения. Результат ранжирования представлен в табл. 2, сверху вниз, слева направо.

Таблица 2 – Ранжированный ряд распределения

25	49	72	89	117	151
30	53	74	97	121	154
35	54	75	98	121	164
39	57	78	101	122	171
41	58	81	102	126	178
42	58	81	103	126	181
45	63	85	105	133	184
46	63	86	105	143	189
48	65	86	106	145	196
49	68	88	111	146	213
					226

2. Определяем число групп разбиения по формуле Стерджесса:

$$N = 1 + 3.322 \lg n = 1 + 3.322 \lg 61 \approx 6.93 = 7 \text{ групп.}$$

3. Рассчитываем величину интервала.

$$x_{\min} = 25, x_{\max} = 226, h_x = 226 - 25 = 201.$$

$$\text{Величина интервала } l = \frac{h_x}{N} = \frac{201}{7} \approx 28.7 \approx 29.$$

4. Расширим границы интервалов снизу до 24.

В результате получим следующие интервалы:

24 – 53; 53 – 82; 82 – 111; 111 – 140; 140 – 169; 169 – 198; 198 – 227.

5. Рассчитаем количество единиц совокупности, принадлежащих каждому из интервалов. При попадании значения на границу интервала будем относить его к предыдущему интервалу; если на границу попадает четное количество, то разделяем значения пополам между соседними интервалами. В результате получаем следующий интервальный ряд распределения, представленный в табл. 3. В таблице сразу рассчитаем значения вариант (середины интервалов) для построения полигона и накопленную частоту для построения кумуляты.

Таблица 3 – Интервальный ряд распределения

Интервал	Варианта (середина интервала), x_i	Частота, n_i	Накопленная частота, S_i
24 – 53	38.5	12	12
53 – 82	67.5	14	26
82 – 111	96.5	14	40
111 – 140	125.5	7	47
140 – 169	154.5	6	53
169 – 198	183.5	6	59
198 – 227	212.5	2	61
Сумма	–	61	–

6. Строим гистограмму, полигон и кумуляту распределения на основе данных, представленных в табл. 3.

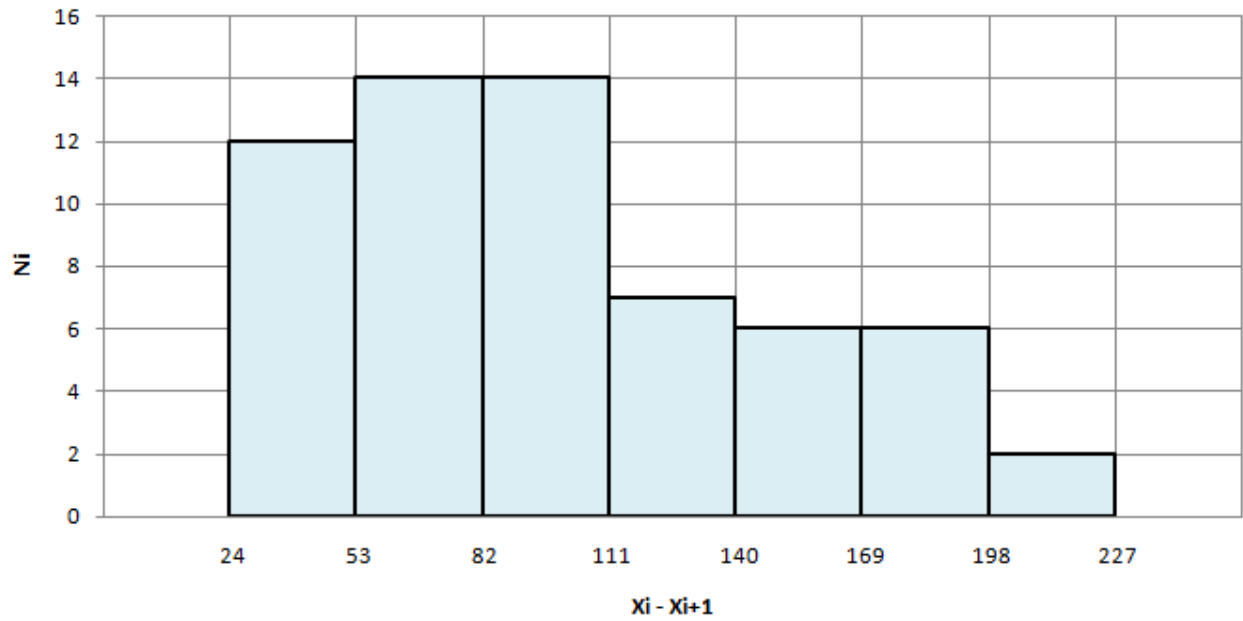


Рисунок 1 – Гистограмма распределения

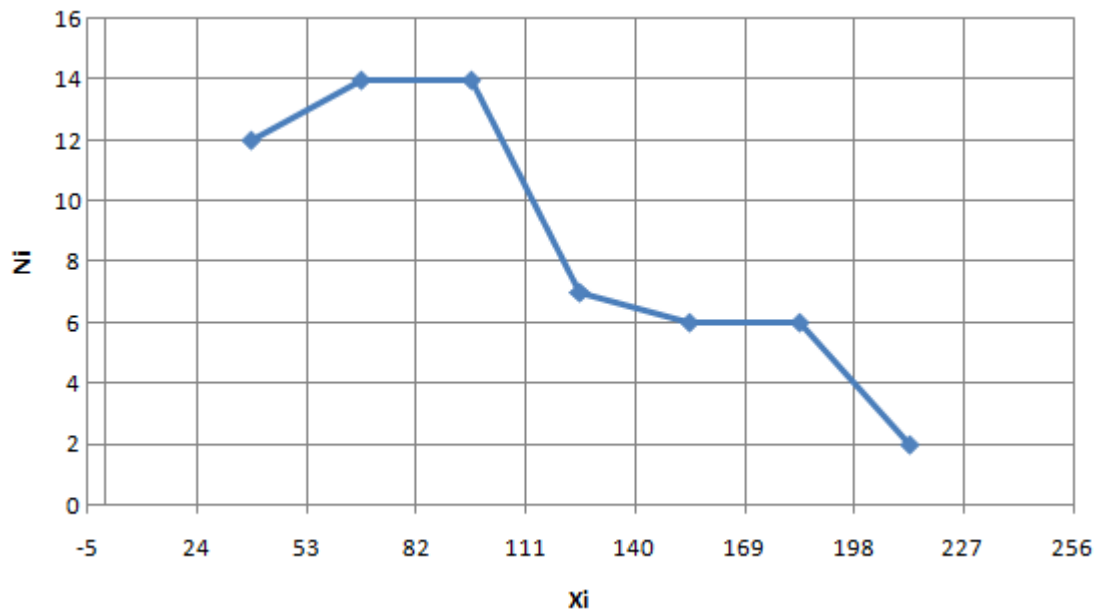


Рисунок 2 – Полигон распределения

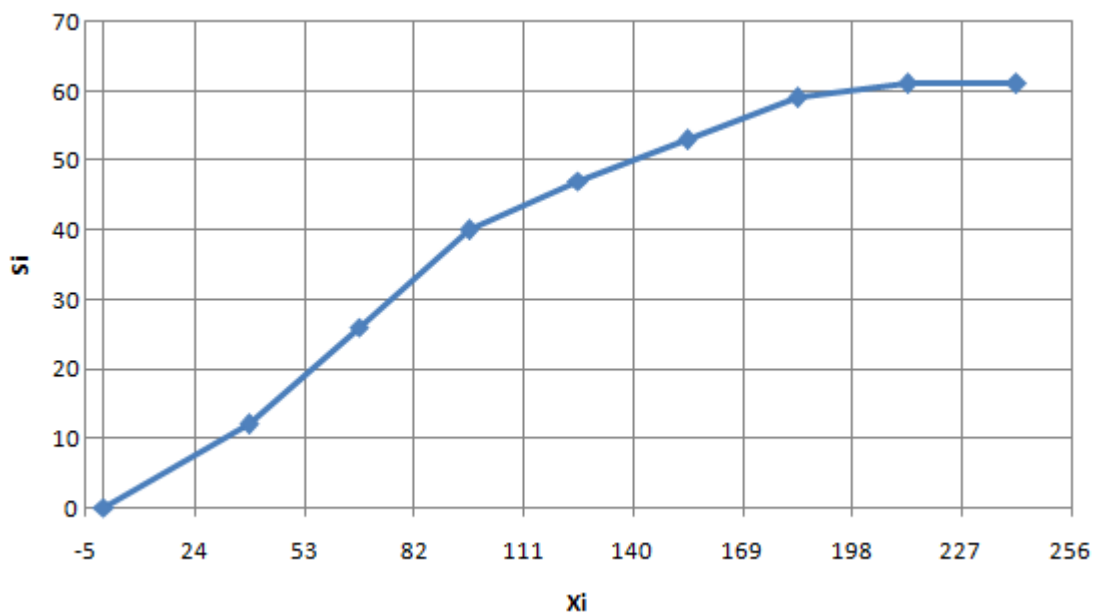


Рисунок 3 – Кумулята распределения

7. Для вычисления средних показателей и показателей вариации составим расчетную таблицу.

Таблица 4 – Расчетная таблица

Варианта, x_i	Частота, n_i	$x_i \cdot n_i$	$(x - \bar{x})^2 \cdot n_i$	$(x - \bar{x})^3 \cdot n_i$	$(x - \bar{x})^4 \cdot n_i$	$ x_i - \bar{x} \cdot n_i$
38.5	12	462	45133.29	-2767928.5	169751155.34	735.93
67.5	14	945	14631.28	-472997.95	15291015.84	452.59
96.5	14	1351	155.05	-515.97	1717.09	46.59
125.5	7	878.5	4613.41	118436.02	3040505.06	179.70
154.5	6	927	17934.25	980503.75	53606229.75	328.03
183.5	6	1101	42006.15	3514744.36	294086150.90	502.03
212.5	2	425	25390.02	2860747.47	322326513.99	225.34
Сумма	61	6089.5	149863.44	4232989.2	858103288	2470.23

Выборочная средняя: $\bar{x}_v = \frac{\sum x_i n_i}{\sum n_i} = \frac{6089.5}{61} \approx 99.8.$

Мода:

$$\begin{aligned} Mo &= x_0 + i \cdot \frac{(n_{M_o} - n_{M_o-1})}{(n_{M_o} - n_{M_o-1}) + (n_{M_o} - n_{M_o+1})} = \\ &= 53 + 58 \cdot \frac{(28 - 12)}{(28 - 12) + (28 - 7)} \approx 78.1, \end{aligned}$$

где $x_0 = 53$ и $i = 58$ – соответственно нижняя граница и величина модального интервала (модальный интервал 53 – 111); $n_{M_o} = 28$, $n_{M_o-1} = 12$, $n_{M_o+1} = 7$ – частоты модального, премодального и послемодального интервалов.

Медиана:

$$Me = x_0 + i \cdot \frac{0.5 \cdot \sum_1^m n_i - S_{Me-1}}{n_{Me}} = 82 + 29 \cdot \frac{0.5 \cdot 61 - 26}{14} \approx 91.3,$$

где $x_0 = 82$ и $i = 29$ – соответственно нижняя граница и величина медианного интервала (медианный интервал 82-111); $n_{Me} = 14$ – частота медианного интервала, $S_{Me-1} = 26$ – накопленная частота предмедианного интервалов.

Квартили:

$$\begin{aligned} Q_1 &= X_{Q_1} + i \cdot \frac{\frac{1}{4} \sum n - S_{Q_1-1}}{n_{Q_1}} = 53 + 29 \cdot \frac{\frac{1}{4} \cdot 61 - 12}{14} \approx 59.7, \\ Q_3 &= X_{Q_3} + i \cdot \frac{\frac{3}{4} \sum n - S_{Q_3-1}}{n_{Q_3}} = 111 + 29 \cdot \frac{\frac{3}{4} \cdot 61 - 40}{7} \approx 134.8, \end{aligned}$$

где $X_{Q_1} = 53$ и $i = 29$ – соответственно нижняя граница и величина интервала, содержащего нижний квартиль (это интервал 53 – 82); $n_{Q_1} = 14$ – частота интервала, $S_{Q_1-1} = 12$ накопленная частота интервала, предшествующего интервалу, содержащему нижний квартиль;

$X_{Q_3} = 111$ и $i = 29$ – соответственно нижняя граница и величина интервала, содержащего верхний квартиль (этот интервал 111 – 140); $n_{Q_3} = 7$ – частота

интервала, $S_{Q_{3-1}} = 40$ – накопленная частота интервала, предшествующего интервалу, содержащему верхний квартиль.

Средний квартиль есть медиана: $Q_2 = M_e = 91.3$.

Децили:

$$d_1 = X_{d1} + i \cdot \frac{\frac{1}{10} \sum n \cdot S_{d1-1}}{n_{d1}} = 24 + 29 \cdot \frac{\frac{1}{10} \cdot 61 - 0}{12} \approx 38.7,,$$

где $X_{d1} = 24$ и $i = 29$ – соответственно нижняя граница и величина интервала, содержащего первый дециль (интервал 24 – 53); $n_{d1} = 12$ – частота интервала, $S_{d1-1} = 0$ – накопленная частота интервала, предшествующего интервалу, содержащему первый дециль.

Аналогично находим и другие децили:

$$d_2 = X_{d2} + i \cdot \frac{\frac{2}{10} \sum n \cdot S_{d2-1}}{n_{d2}} = 53 + 29 \cdot \frac{\frac{2}{10} \cdot 61 - 12}{14} \approx 53.4,$$

$$d_3 = X_{d3} + i \cdot \frac{\frac{3}{10} \sum n \cdot S_{d3-1}}{n_{d3}} = 53 + 29 \cdot \frac{\frac{3}{10} \cdot 61 - 12}{14} = 66.1,$$

$$d_4 = X_{d4} + i \cdot \frac{\frac{4}{10} \sum n \cdot S_{d4-1}}{n_{d4}} = 53 + 29 \cdot \frac{\frac{4}{10} \cdot 61 - 12}{14} = 78.7,$$

$$d_5 = M_e = 91.3,$$

$$d_6 = X_{d6} + i \cdot \frac{\frac{6}{10} \sum n \cdot S_{d6-1}}{n_{d6}} = 82 + 29 \cdot \frac{\frac{6}{10} \cdot 61 - 26}{14} \approx 104,$$

$$d_7 = X_{d7} + i \cdot \frac{\frac{7}{10} \sum n \cdot S_{d7-1}}{n_{d7}} = 111 + 29 \cdot \frac{\frac{7}{10} \cdot 61 - 40}{7} \approx 122.2,$$

$$d_8 = X_{d8} + i \cdot \frac{\frac{8}{10} \sum n \cdot S_{d8-1}}{n_{d8}} = 140 + 29 \cdot \frac{\frac{8}{10} \cdot 61 - 47}{6} \approx 148.7,$$

$$d_9 = X_{d9} + i \cdot \frac{\frac{9}{10} \sum n \cdot S_{d9-1}}{n_{d9}} = 169 + 29 \cdot \frac{\frac{9}{10} \cdot 61 - 53}{6} \approx 178.2.$$

8. Рассчитываем показатели вариации.

$$\text{Размах вариации } R = x_{\max} - x_{\min} = 226 - 25 = 201.$$

Среднее линейное отклонение:

$$d = \frac{\sum |x_i - \bar{x}| \cdot n_i}{\sum n_i} = \frac{2470.23}{61} \approx 40.5.$$

$$\text{Дисперсия: } s_g^2 = \frac{\sum (x_i - \bar{x})^2 \cdot n_i}{\sum n_i - 1} = \frac{149863.44}{60} \approx 2497.7.$$

Исправленное среднее квадратическое отклонение:

$$s_g = \sqrt{s_g^2} = \sqrt{2497.7} \approx 50.$$

$$\text{Коэффициент осцилляции: } V_r = \frac{R}{x} = \frac{201}{99.8} \approx 2.$$

$$\text{Относительное линейное отклонение: } V_d = \frac{d}{x} = \frac{40.5}{99.8} \approx 0.4.$$

$$\text{Коэффициент вариации: } V_\sigma = \frac{s_g}{x} \cdot 100\% = \frac{50}{99.8} \cdot 100\% \approx 50\%.$$

Коэффициент асимметрии:

$$A_s = \frac{\left(\sum (x_i - \bar{x})^3 \cdot n_i \right) / \sum n_i}{s_g^3} = \frac{4232989.2 / 61}{50^3} \approx 0.56.$$

Коэффициент эксцесса:

$$E = \frac{\left(\sum (x_i - \bar{x})^4 \cdot n_i \right) / \sum n_i}{\sigma_e^4} - 3 = \frac{858103288 / 61}{50^4} - 3 = -0.75.$$

Выводы:

Выборочная средняя равна 99.8, мода распределения равна 78.1, а медиана 91.3.

Так как $Mo < M_e < \bar{x}$, то имеет место правосторонняя асимметрия (правая ветвь гистограммы распределения более вытянута). Данный вывод подтверждается и коэффициентом асимметрии, который выше 0.

Коэффициент эксцесса ниже 0, что характеризует плосковершинное распределение.

Среднее квадратичное отклонение равно 50, коэффициент вариации выше 33 %, что говорит о неоднородности рассматриваемой совокупности.

Список литературы:

1) Кильдеева Р.А. Практикум по статистике. Методическое пособие для студентов экономических специальностей очной и заочной форм обучения. – Петропавловск-Камчатский: КамчатГТУ, 2005. – 240 с.

2) Гусаров В.М. Статистика: Учебное пособие для вузов. – М.: ЮНИТИ-ДАНА, 2001.

3) Статистика: Учебное пособие/Харченко Л-П., Долженкова В.Г., Ионин В.Г. и др., Под ред. В.Г.Ионина. - Изд. 2-е, перераб. и доп. - М.: ИНФРА-М.2003.