

**Практическое занятие № 1**

**ПОСТРОЕНИЕ ВАРИАЦИОННОГО РЯДА**

Использование выборок большого объема позволяет снизить величину статистической ошибки, но получаемый при этом значительный массив данных неудобен для проведения статистической характеристики. Для группировки наблюдений в большой выборке, где число значений варьирующего количественного признака 30 и более, используются вариационные ряды.

**Пример.** При изучении признака «*Масса плодов перца сорта Золотой Юбилей*» были получены следующие результаты, г:

92	96	86	102	93	77	98	95	83	90
97	82	104	94	99	95	86	105	91	110
93	89	103	75	91	87	100	83	88	94
99	105	93	96	84	98	92	105	107	94
85	87	104	99	91	103	89	93	88	100

**Необходимо построить вариационный ряд и гистограмму распределения частот.**

При проведении наблюдения был получен достаточно большой объем данных:  $n=50$ . Построение вариационного ряда начинается с определения количества интервалов, на которые надо разбить выборку. Для определения оптимального количества интервалов  $k$  из общего числа наблюдений  $n$  извлекается квадратный корень. Если полученное значение не целое, его следует округлить до целого, т.к. количество интервалов не может быть выражено дробным числом. В нашем случае

$$k = \sqrt{n} = 7$$

Интервалы, кроме первого и последнего, должны быть одинаковых размеров. Величина первого и последнего интервалов может быть немного больше или меньше в связи с тем, что в них попадает меньше значений признака.

Для успешной разбивки выборки на интервалы рассчитывают шаг интервала  $i$ . Для этого находят самое большое  $X_{max}$  и самое маленькое  $X_{min}$  значение признака, находят между ними разность, которую делят на количе-

ство интервалов  $k$ . Полученное частное округляем до того же уровня точности, которое мы использовали для записи значений признака. Т.е. если значения признака даны в целых числах, округляем до целых; если значения признака даны с точностью до десятых, то округляем до десятых и т.д. В нашем случае

$$i = \frac{X_{\max} - X_{\min}}{k} = \frac{110 - 75}{7} = 5$$

Разбиваем весь массив данных на интервалы. Интервал представляет собой ряд данных, для более краткой записи интервала используют только первое и последнее число (нижняя и верхняя границы) между которыми ставят знак интервала  $\div$ . Нижней границей первого интервала является  $X_{\min}$ , для определения верхней границы первого интервала к  $X_{\min}$  прибавляют шаг интервала  $i$ . В нашем случае первый интервал имеет следующий вид  $75 \div 80$ . Второй интервал начинается с числа, идущего после верхней границы интервала, т.е. нижней границей второго интервала будет число  $81$ . Для определения верхней границы второго интервала к нижней границе мы прибавляем « $i$  - единица точности». В нашем примере значения признака даны в целых, т.е. единицей точности является 1. Таким образом, верхнюю границу второго интервала мы определяем как  $81 + (5 - 1) = 81 + 4 = 85$ . Этот же алгоритм мы используем для определения последующих интервалов.

Мы получили следующие интервалы:

- I.  $75 \div 80$
- II.  $81 \div 85$
- III.  $86 \div 90$
- IV.  $91 \div 95$
- V.  $96 \div 100$
- VI.  $101 \div 105$
- VII.  $106 \div 110$

На следующем этапе необходимо определить сколько значений признака входит в каждый интервал. Для того чтобы не переписывать еще раз значения признака, используют метод штрихов или конвертиков. Более

простым является метод штрихов. Сущность этого метода заключается в том, что вместо значения признака напротив соответствующего интервала мы рисуем штрих, затем подсчитываем количество штрихов и определяем, сколько значений признака находятся в полученных интервалах, т.е. определяем частоту встречаемости признака **f**. Сумма частот  $\sum f$  должна быть равна общему числу значений признака **n**, т.е.

$$\sum f = n = 50$$

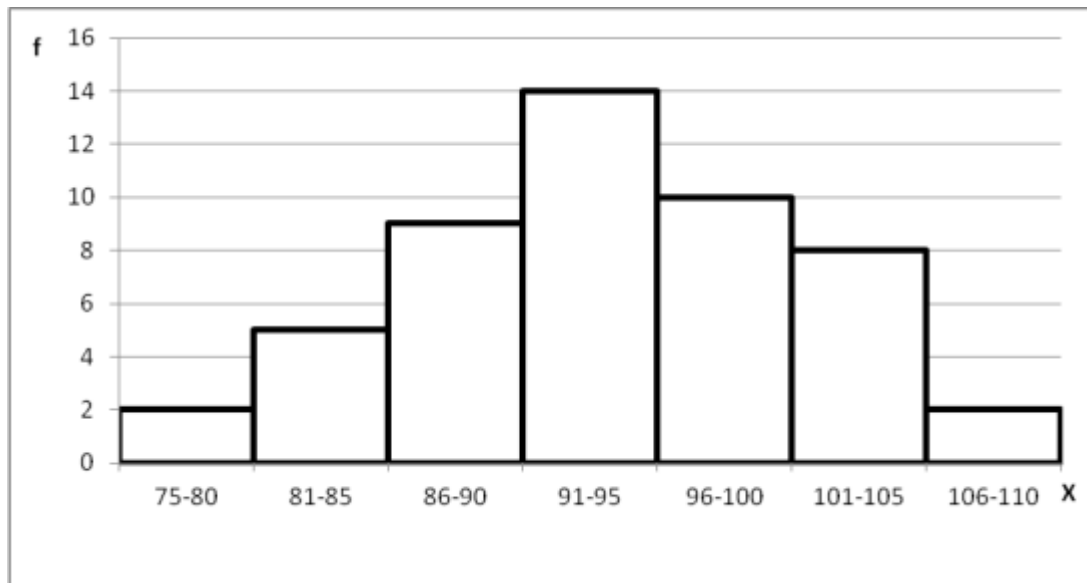
Заполняем таблицу:

Интервал	Разноска значений признака	Частота, f	X
75 ÷ 80	II	2	78
81 ÷ 85	IIII	5	83
86 ÷ 90	IIII III	9	88
91 ÷ 95	IIII IIII III	14	93
96 ÷ 100	IIII IIII	10	98
101 ÷ 105	IIII III	8	103
106 ÷ 110	II	2	108
<b>Сумма</b>	-	<b>50</b>	-

X – середина интервала, для определения X складываем верхнюю и нижнюю границы интервала и полученное слагаемое делим на 2. Так, для первого интервала  $X=(75+80):2=155:2=77,5$  округляем до целых 78; для второго  $X=(81+85):2=166:2=83$  и т.д.

*Таким образом, заполнив столбцы «X» и «f», мы получаем вариационный ряд, где каждому значению признака X соответствует частота f.*

Строим гистограмму распределения частот:



**Задание.** Согласно N примера и варианта построить вариационный ряд и гистограмму распределения частот результатов наблюдения.

*Шифр задания к работе*

№ варианта	Номера колонок	№ варианта	Номера колонок	№ варианта	Номера колонок
<b>1</b>	1 2 3 4 5	<b>11</b>	1 3 4 7 9	<b>21</b>	3 7 8 9 10
<b>2</b>	2 3 4 5 6	<b>12</b>	4 5 8 9 10	<b>22</b>	1 4 5 6 9
<b>3</b>	3 4 5 6 7	<b>13</b>	1 2 3 6 9	<b>23</b>	1 5 6 7 10
<b>4</b>	4 5 6 7 8	<b>14</b>	2 3 4 8 10	<b>24</b>	2 4 5 7 8
<b>5</b>	5 6 7 8 9	<b>15</b>	3 4 5 9 10	<b>25</b>	1 5 7 9 10
<b>6</b>	6 7 8 9 10	<b>16</b>	1 3 5 6 8	<b>26</b>	2 3 4 8 10
<b>7</b>	1 3 5 7 9	<b>17</b>	4 6 7 9 10	<b>27</b>	3 5 6 8 9
<b>8</b>	2 4 6 8 10	<b>18</b>	2 5 6 8 9	<b>28</b>	2 4 6 7 9
<b>9</b>	1 2 5 6 9	<b>19</b>	3 5 6 8 10	<b>29</b>	3 4 5 8 9
<b>10</b>	2 3 5 7 10	<b>20</b>	1 2 4 5 7	<b>30</b>	4 6 7 8 9

**Пример 1. Масса плодов сливы сорта Венгерка итальянская, г**

Значение признака по колонкам

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
54	45	44	46	43	38	46	39	49	40
39	37	42	47	35	43	47	36	40	42
31	48	45	57	45	40	40	41	44	60
43	40	58	41	49	35	33	48	52	43
38	35	46	36	52	39	44	45	42	44
45	56	51	45	46	54	42	38	50	35
49	44	41	32	47	41	37	45	43	39
46	53	44	35	39	47	40	48	34	51
34	50	45	42	46	59	53	40	41	47
45	39	36	54	42	46	43	62	46	41

**Пример 2. Масса плодов томата сорта Марьюшка, г**

Значение признака по колонкам

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
55	59	50	75	57	57	70	81	63	67
62	68	61	80	65	48	65	75	65	87
83	49	77	64	51	58	79	64	77	74
67	56	66	60	69	62	56	58	53	56
48	71	40	52	74	80	44	88	72	66
75	63	63	68	76	66	49	64	59	76
79	60	89	73	59	69	63	55	68	65
70	53	57	45	66	86	57	50	58	62
58	67	72	65	79	60	70	78	45	73
81	78	66	58	87	73	61	71	51	59

**Выполнение работы.**

1. Записать название примера, номер варианта.
2. Заполнить рабочую таблицу.

*Рабочая таблица*

Номера колонок	Значение признака, X

3. Подсчитать количество значений признака  $n =$
4. Определить количество интервалов  $k = \sqrt{n}$
5. Рассчитать величину интервала  $i = \frac{X_{\max} - X_{\min}}{k}$
6. Заполнить таблицу

Интервал	Разноска значений признака	Частота, f	X
<b>Сумма</b>	-		-

7. Построить гистограмму распределения частот: