

Техника построения вариационных рядов



Пример: На основании многолетних клинических наблюдений, проводившихся в Сухумском питомнике обезьян, составлена следующая выборка, включающая 100 анализов на содержание кальция (мг %) в сыворотке крови низших обезьян (павианов-гамдерилов).

Приведены данные.

Нужно сгруппировать эти данные в вариационный ряд. В данном случае признак варьирует непрерывно в пределах от 9,0 до 14,7 мг %.

Устанавливаем величину классового интервала:

$$h = \frac{14,7 - 9,0}{1 + 3,321g100} = \frac{5,7}{7,6} = 0,75 \approx 0,8$$

Определяем нижнюю границу первого класса:

$$X_H = 9,0 - \frac{0,8}{2} = 8,6$$

Затем намечаем следующие классовые интервалы:

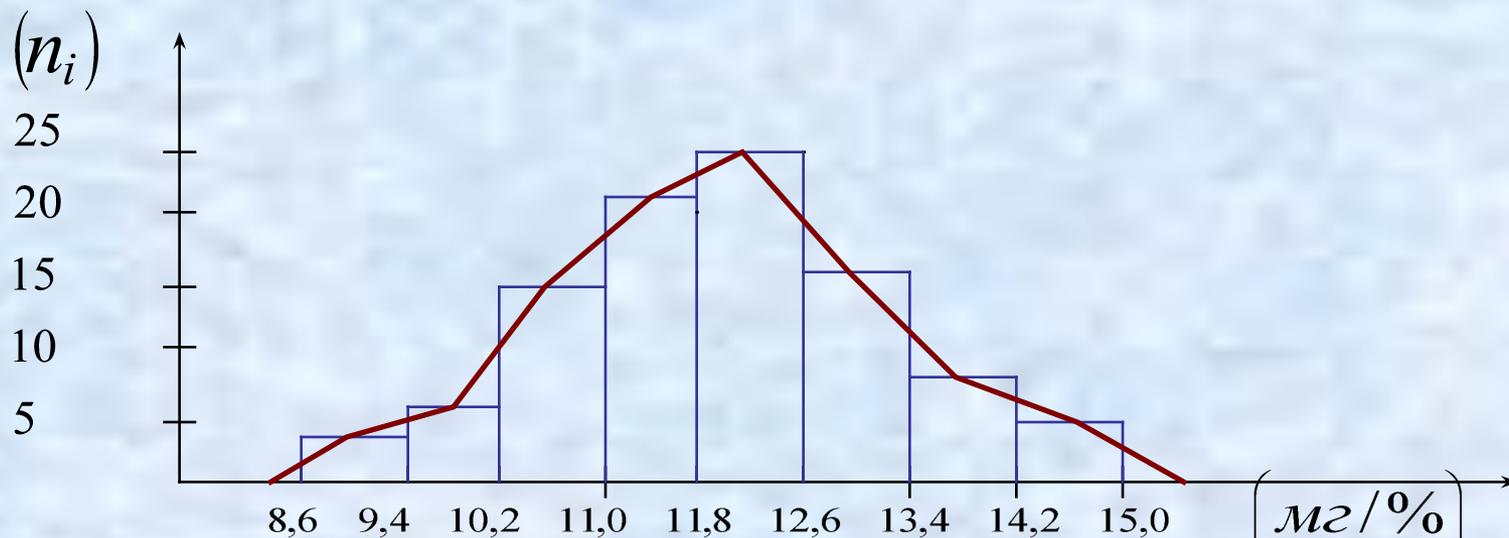
8,6 – 9,4 – 10,2 – 11,0 – 11,8 – 12,6 – 13,4 – 14,2 – 15,0

Получилось 8 интервалов.

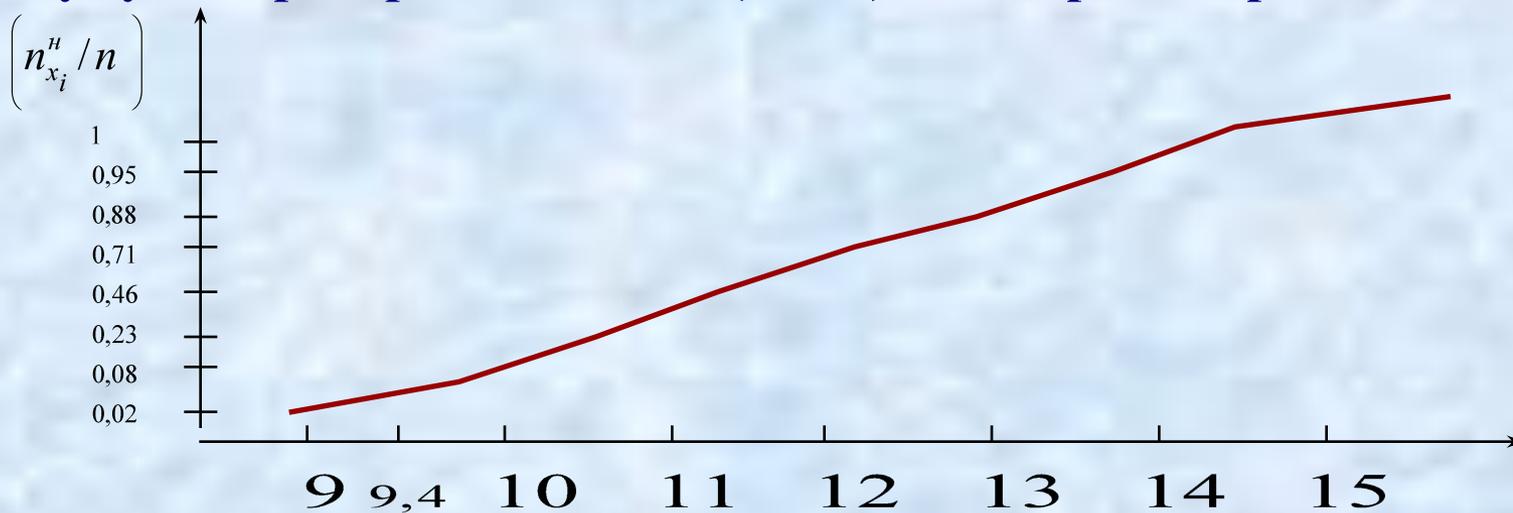
Строим вспомогательную таблицу и разносим все 100 вариаций по намеченным классовым интервалам.

Классы по уровню кальция в сыворотке крови, мг %	Срединные значения классов <i>x_i</i>	Частоты <i>n_i</i>	Накопл. частота <i>n_i^+</i>	$\frac{n_{x_i}^H}{n}$
8,6-9,4	9,0	2	2	0,02
9,4-10,2	9,8	6	8	0,08
10,2-11,0	10,6	15	23	0,23
11,0-11,8	11,4	23	46	0,46
11,8-12,6	12,2	25	71	0,71
12,6-13,4	13,0	17	88	0,88
13,4-14,2	13,8	7	95	0,95
14,2-15,0	14,6	5	100	1
Сумма		100		

Гистограмма распределения Са (мг %) в сыворотке крови обезьян.



Кумулята распределения Са (мг %) в сыворотке крови обезьян.



Основные характеристики варьирующих объектов.

Среднее значение выборки, мода, медиана,
размах.

Вариационные ряды и их графики дают наглядное представление о варьировании признаков, но они недостаточны для полного описания варьирующих объектов. Для этой цели служат особые логически и теоретически обоснованные числовые показатели, называемые статистическими характеристиками.

К ним относятся прежде всего средние величины и показатели вариации.

Средние величины могут характеризовать только однородную совокупность вариантов.

Средняя арифметическая \bar{X} может быть простой и взвешенной.

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} \quad x: x_1, x_2, x_3, \dots, x_n$$

Когда отдельные варианты повторяются:

$x: x_1 - n_1$ раз
 $x_2 - n_2$ раз
 $x_k - n_k$ раз

$$\bar{X} = \frac{\sum_{i=1}^k n_i x_i}{n}$$

Некоторые свойства сумм:

$$1. \sum_{i=1}^n x_i = n\bar{x} \quad (\text{из определения})$$

$$2. \sum_{i=1}^n Cx_i = C \sum_{i=1}^n x_i$$

$$3. \sum_{i=1}^n (x_i \pm C) = \sum_{i=1}^n x_i \pm nC$$

$$4. \sum_{i=1}^n (x_i \pm y_i) = \sum_{i=1}^n x_i \pm \sum_{i=1}^n y_i$$

Свойства среднего:

$$1. \quad \overline{x \pm C} = \frac{\sum_{i=1}^n (x_i \pm C)}{n} = \overline{x} \pm C$$

$$2. \quad \overline{Cx} = \frac{\sum_{i=1}^n Cx_i}{n} = \frac{C \sum_{i=1}^n x_i}{n} = C \overline{x}$$

$$3. \quad \sum_{i=1}^n (x_i - \overline{x}) = (x_1 - \overline{x}) + (x_2 - \overline{x}) + \dots + (x_n - \overline{x}) =$$
$$= (x_1 + x_2 + \dots + x_n) - n\overline{x} = \sum_{i=1}^n x_i - n\overline{x} = n\overline{x} - n\overline{x} = 0$$

Структурные средние – это величины обычно представляют собой конкретные варианты имеющейся совокупности, которые занимают особое место в ряду распределения.

Def: Медиана – это значение варианты, которая делит ранжированный ряд на равные по числу вариант части.

4 7 12 8 9

$$Me = 12$$

5 7 13 15

$$Me = \frac{7+13}{2} = 10$$

медианному интервалу соответствует первая накопленная частота, превосходящая $n/2$.

$$Me = x_{нме} + h \cdot \frac{\frac{n}{2} - n_{x_{ме-1}}^H}{n_{ме}^H}$$

где:

$x_{нме}$ – нижняя граница медианного интервала.

h – шаг разбиения, ширина класса.

$n_{ме}^H$ – абсолютная частота медианного интервала.

$n_{x_{ме-1}}^H$ – накопленная частота интервала, предшествующего медианному интервалу.

Пример (для таблицы 1):

$$n/2 = 50 \Rightarrow n_{xi}^H = 71 \Rightarrow$$

медианный интервал : 11,8–12,6

$$x_{нме} = 11,8, \quad h = 0,8$$

$$n_{ме-1}^H = 46, \quad \cancel{n_{ме}^H} = 25$$

$$Me = 11,8 + 0,8 \cdot \frac{50 - 46}{25} = 11,8 + 0,128 \approx 11,93.$$

Модой называется величина, наиболее часто встречающаяся в данной совокупности. Класс с наибольшей частотой называется модальным. Для определения моды интервальных рядов служит формула

$$M_o = x_{нмo} + h \cdot \frac{n_{M_o} - n_{M_o - 1}}{2n_{M_o} - n_{M_o - 1} - n_{M_o + 1}}$$

$x_{нмo}$ – нижняя граница модального интервала.

h – ширина класса

n_{M_o} – абсолютная частота модального интервала.

$n_{M_o - 1}$ – абсолютная частота интервала, предшествующего модальному.

$n_{M_o + 1}$ – абсолютная частота интервала, следующего за модальным.

Классы по уровню кальция в сыворотке крови, мг %	Срединные значения классов	Частоты n_i	Накопл. частота	$\frac{n_{x_i}^H}{n}$
8,6-9,4	9,0	2	2	0,02
9,4-10,2	9,8	6	8	0,08
10,2-11,0	10,6	15	23	0,23
11,0-11,8	11,4	23	46	0,46
11,8-12,6	12,2	25	71	0,71
12,6-13,4	13,0	17	88	0,88
13,4-14,2	13,8	7	95	0,95
14,2-15,0	14,6	5	100	1
<i>Сумма</i>		<i>100</i>		

Характеристики рассеяния:

размах, дисперсия.

Размах R_x

$$R_x = x_{\max} - x_{\min}$$