Дисциплины:
«Теория вероятностей», «Математическая статистика», «Теория вероятностей и математическая статистика»

Тема: Вариационный ряд

МТИ
Московский
Технологический
Институт

Основные определения

- **Математическая статистика** изучает случайные события и случайные величины по результатам наблюдений.
- **Статистическая совокупность** – это совокупность предметов или явлений, объединенных каким-либо общим признаком.
- **Статистические данные** – это результат наблюдений над статистической совокупностью – это сведения о том, какие значения принял в итоге наблюдения изучаемый признак.
- Функция, характеризующая наблюдаемую случайную величину, называется **статистикой**. Она каждому набору наблюдаемых значений признака ставит в соответствие определенное действительное число.
- **Генеральная совокупность** – это совокупность объектов или наблюдений, все элементы которой подлежат изучению при статистическом анализе. Число объектов в генеральной совокупности называется ее **объемом**.

Основные определения

- Часть объектов генеральной совокупности, используемая для исследования, называется **выборочной совокупностью** или **выборкой.**
- Для того чтобы по выборке можно было адекватно судить о случайной величине, она должна быть **репрезентативной**.
- Существует два способа образования выборки:
- **1) повторная выборка**, когда каждый элемент, случайно отобранный и исследованный, возвращается в общую совокупность и может быть отобран повторно;
- **2) бесповторная выборка**, когда отобранный элемент не возвращается в общую совокупность

МТИ Московский Технологический Институт

Основные определения

Пусть n число проведенных наблюдений случайной величины X, среди которых r различных вариантов, каждое из которых фиксировалось ni раз:

$$\begin{pmatrix} x_1 & x_2 & \dots & x_r \\ n_1 & n_2 & \dots & n_r \end{pmatrix} \qquad n = \sum_{i=1}^{r} n_i$$

ni - называются **частотами** варианты xi,

$w_i = \dfrac{n_i}{n}$ - **частости** (доли, относительные частоты);

Частоты и частости называются **весами**.

В бесповторной выборке все ni=1, а r=n.

Московский Технологический Институт

Основные определения

- Ряд вариант, расположенных в порядке возрастания их значений, с соответствующими им весами называется **вариационным рядом**.
- Вариационный ряд называется **дискретным**, если он представляет собой выборку значений дискретной случайной величины

Варианты xi	x1	x2	…	xr
Частоты ni	n1	n2	…	nr

- Вариационный ряд называется **непрерывным (интервальным)**, если он представляет собой выборку значений непрерывной случайной величины

Варианты xi	[a1,a2)	[a2,a3)	…	[ar,ar+1)
Частоты ni	n1	n2	…	nr

Московский Технологический Институт

Основные определения

Для наглядности представления рядов используют:

1) полигоны по точкам (xi,ni) или (ci,ni), где ci - середины интервалов интервальных рядов,
2) Гистограммы (столбиковые) на интервалах,
3) Кумулянты по точкам (xi,mi) или (ci,mi), где mi – это накопленные частоты.

Эмпирической функцией распределения вариационного ряда называется функция, равная накопленным частостям:

$$F_n(x) = w_x = \frac{m_x}{n} = \frac{1}{n}\sum_{x_i < x} n_i$$

Эмпирической плотностью распределения непрерывного вариационного ряда называется функция равная

$$f_n(x) = \frac{n_i}{n\Delta}$$

внутри интервалов и равная нулю за указанными интервалами.

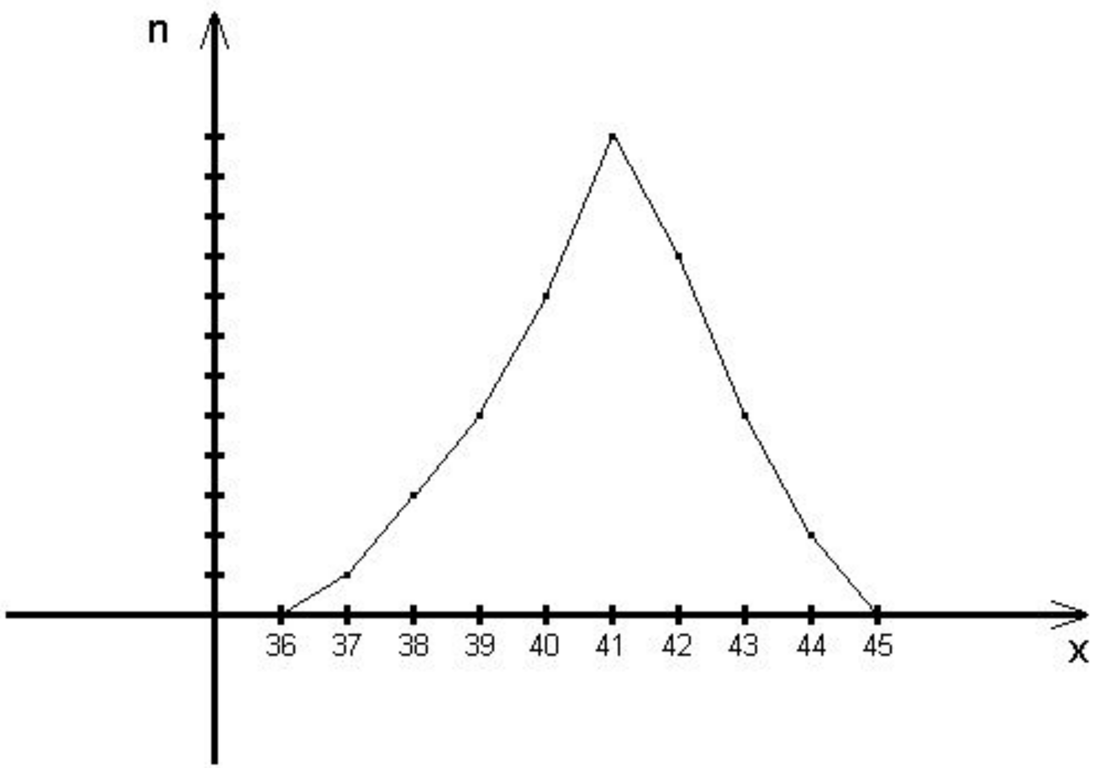
Пример 1.

- В магазине за день было продано 45 пар мужской обуви. Имеется выборка значений случайной величины *X* – размера обуви:
- 39, 41, 40, 42, 41, 40, 42, 44, 40, 43, 42, 41, 43, 39, 42,
- 41, 42, 39, 41, 37, 43, 41, 38, 43, 42, 41, 40, 41, 38, 44,
- 40, 39, 41, 40, 42, 40, 41, 42, 40, 43, 38, 39, 41, 41, 42.
- Построить дискретный вариационный ряд, полигон, кумулянту и эмпирическую функцию распределения.

Решение. Для построения вариационного ряда различные значения признака располагаем в порядке их возрастания и под каждым из этих значений записываем его частоту:

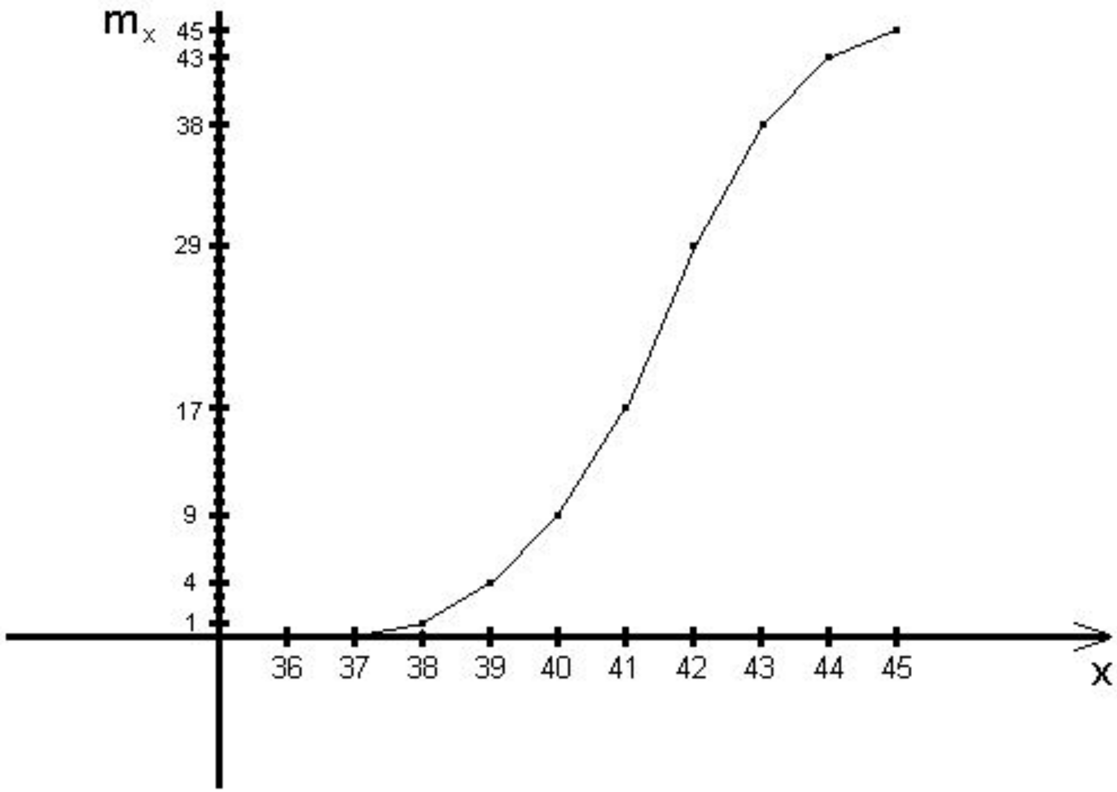
xi	37	38	39	40	41	42	43	44
ni	1	3	5	8	12	9	5	2

- Построим для этого ряда полигон.
- Сначала отметим на графике точки (x_i, n_i)



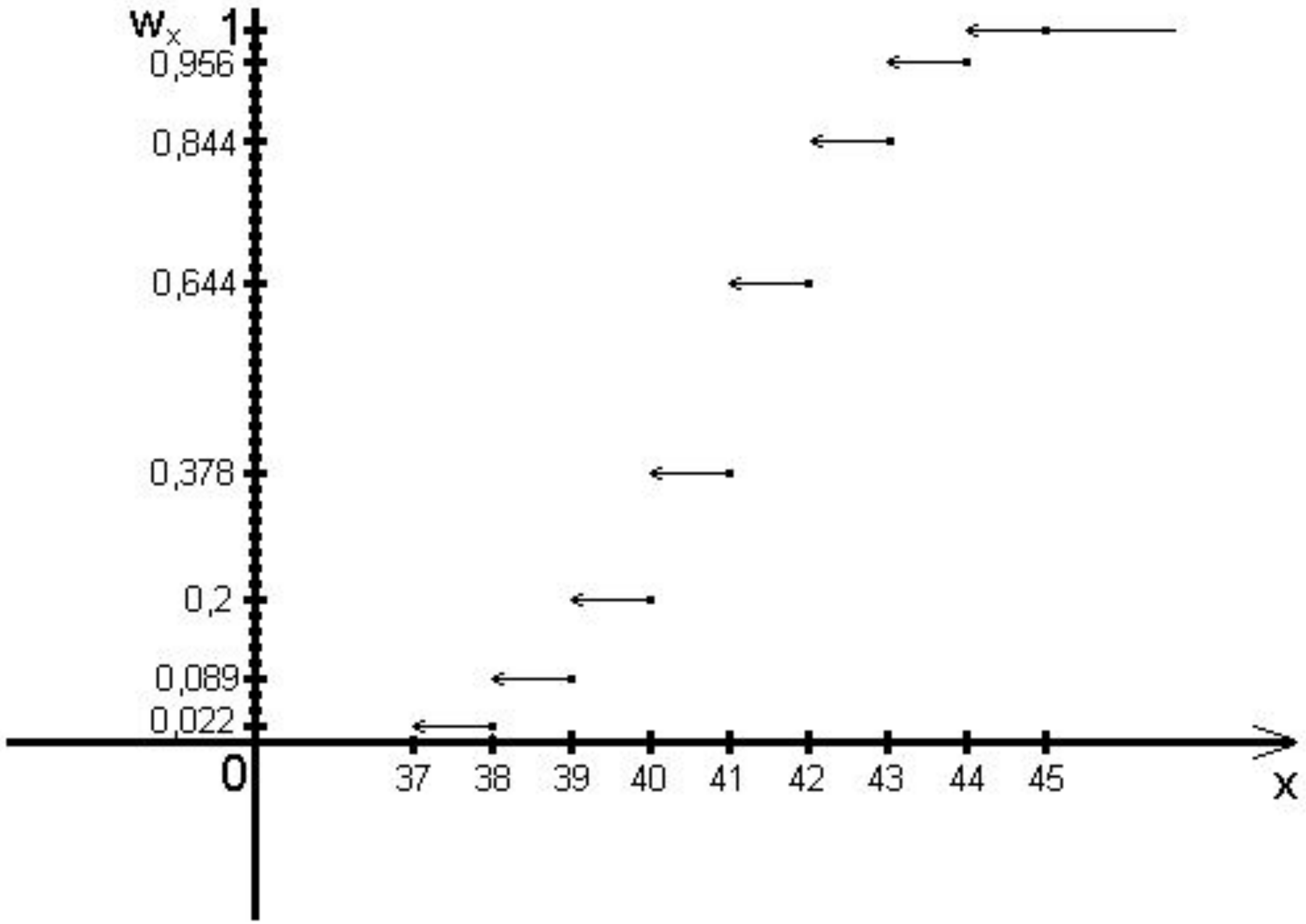
- По построенной выше таблице распределения найдем накопленные частоты и частости

xi	37	38	39	40	41	42	43	44	45
ni	1	3	5	8	12	9	5	2	0
m	0	1	4	9	17	29	38	43	45
w	0	1/45= 0,022	4/45= 0,089	9/45= 0,2	17/45= 0,375	29/45= 0,644	38/45= 0,844	43/45= 0,956	45/45= 1



Построим кумулянту

- Построим эмпирическую функцию распределения



Числовые характеристики вариационного ряда:

1. **Выборочное среднее** – средняя арифметическая наблюдаемых вариант признака

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{или} \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{r} x_i n_i$$

Свойства средней:

$\bar{c} = c$, C – const

$\overline{kx} = k \cdot \bar{x}$, K - const

$\overline{x - \bar{x}} = 0$

$\overline{x \pm y} = \bar{x} \pm \bar{y}$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{l} \overline{x_i} m_i$$

$\overline{x_i}$ Групповые средние

m_i Объемы групп

l Число групп

Числовые характеристики вариационного ряда:

2. Вариационный размах: $R = x_{\max} - x_{\min}$

3. Выборочное среднее линейное отклонение

$$d = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - \bar{x}\right| \quad \text{или} \quad d = \frac{1}{n}\sum_{i=1}^{r}\left|x_i - \bar{x}\right| n_i$$

4. Выборочная дисперсия

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \quad \text{или} \quad S^2 = \frac{1}{n}\sum_{i=1}^{r}(x_i - \bar{x})^2 n_i \quad \text{или} \quad S^2 = \overline{x^2} - \bar{x}^2$$

5. **Выборочное среднее квадратическое отклонение** $S = \sqrt{S^2}$

6. **Выборочный коэффициент вариации** $v = \dfrac{S}{|\bar{x}|} \cdot 100\%$

7. **Исправленная выборочная дисперсия** $\overline{S^2} = \dfrac{n \cdot S^2}{n-1}$

8. **Выборочное исправленное среднее квадратическое отклонение** $\bar{S} = \sqrt{\overline{S^2}}$

Московский Технологический Институт

Свойства дисперсии:

1. дисперсия постоянной величины равна нулю;
2. если ко всем вариантам случайной величины добавить постоянное число, то дисперсия не изменится;
3. если все варианты случайной величины умножить на одно и то же число k, то дисперсия умножится на k^2
4. Правило сложения дисперсий: $S^2 = \overline{S_i^2} + \delta^2$

$$S_i^2 = \frac{1}{n_i}\sum_{j=1}^{n_i}(x_{ij} - \overline{x_i})^2 n_{ij}$$ - групповые дисперсии

$$\overline{S_i^2} = \frac{1}{n}\sum_{i=1}^{l} S_i^2 n_i$$ - средняя групповых дисперсий

$$\delta^2 = \frac{1}{n}\sum_{i=1}^{l}(\overline{x_i} - \overline{x})^2 n_i$$ - межгрупповая дисперсия

5. Если $U = \dfrac{X - C}{h}$, то $\overline{x} = h \cdot \overline{U} + C$, $S_x^2 = h^2 S_u^2$

Московский Технологический Институт

Пример 2

В условии **примера 1** вычислим числовые характеристики полученного вариационного ряда: n=45

$$\bar{x} = \frac{37\cdot 1 + 38\cdot 3 + 39\cdot 5 + 40\cdot 8 + 41\cdot 12 + 42\cdot 9 + 43\cdot 5 + 44\cdot 2}{45} =$$

$$= \frac{1839}{45} = 40{,}867.$$

$$\overline{x^2} = \frac{37^2\cdot 1 + 38^2\cdot 3 + 39^2\cdot 5 + 40^2\cdot 8 + 41^2\cdot 12 + 42^2\cdot 9 + 43^2\cdot 5 + 44^2\cdot 2}{45} =$$

$$= \frac{75271}{45} = 1672{,}689,$$

$$S^2 = \overline{x^2} - \bar{x}^2 = 1672{,}689 - (40{,}867)^2 = 1672{,}689 - 1670{,}112 = 2{,}577$$

$$S = \sqrt{S^2} = \sqrt{2{,}577} = 1{,}605, \qquad R = x_{\max} - x_{\min} = 44 - 37 = 7$$

$$v = \frac{S}{|\bar{x}|}\cdot 100\% = \frac{1{,}605}{40{,}867}100\% = 3{,}93\%$$

Московский Технологический Институт

Пример 3

Приведены данные об урожайности ржи на различных участках поля:

Урожайность, ц/га x_i	[9-12]	[12-15]	[15-18]	[18-21]	[21-24]	[24-27]
Доля участка в общей площади, %	6	12	33	22	19	8

Найти выборочную среднюю, дисперсию, коэффициент вариации и размах урожайности ржи.

Решение. Так как имеем интервальный ряд, то расчеты будут производиться по серединам интервалов. Преобразуем исходную таблицу к виду:

xi	[9-12]	[12-15]	[15-18]	[18-21]	[21-24]	[24-27]
ci	10,5	13,5	16,5	19,5	22,5	25,5
ni	6	12	33	22	19	8

Московский Технологический Институт

$$\bar{x} = \frac{10{,}5 \cdot 6 + 13{,}5 \cdot 12 + 16{,}5 \cdot 33 + 19{,}5 \cdot 22 + 22{,}5 \cdot 19 + 25{,}5 \cdot 8}{6 + 12 + 33 + 22 + 19 + 8} =$$

$$= \frac{1830}{100} = 18{,}3.$$

$$\overline{x^2} = \frac{10{,}5^2 \cdot 6 + 13{,}5^2 \cdot 12 + 16{,}5^2 \cdot 33 + 19{,}5^2 \cdot 22 + 22{,}5^2 \cdot 19 + 25{,}5^2 \cdot 8}{6 + 12 + 33 + 22 + 19 + 8} =$$

$$= \frac{35019}{100} = 350{,}19$$

$$S^2 = \overline{x^2} - \bar{x}^2 = 350{,}19 - (18{,}3)^2 = 350{,}19 - 334{,}89 = 15{,}3$$

$$S = \sqrt{15{,}3} = 3{,}91$$

$$v = \frac{3{,}91}{18{,}3} \cdot 100\% = 21{,}366\%$$

$$R = 27 - 9 = 18$$

Пример 4

В таблице приведено распределение n=50 рабочих по производительности труда X (единиц за смену), разделенных на две группы. Найти общие и групповые средние и проверить «правило сложения дисперсий».

	1 группа					2 группа				
xij	34	85	96	102	103	63	69	83	89	106
nij	5	2	11	8	4	2	6	8	3	1

Решение. Введем новые варианты $U_{ij} = \dfrac{x_{ij} - C}{h} = \dfrac{x_{ij} - 80}{10}$, тогда

	1 группа					2 группа				
Uij	-4,6	0,5	1,6	2,2	2,3	-1,7	-1,1	0,3	0,9	2,6
nij	5	2	11	8	4	2	6	8	3	1

Определим групповые средние и дисперсии:

$$n_1 = 5+2+11+8+4 = 30, \quad n_2 = 2+6+8+3+1 = 20, \quad l = 2$$

$$\overline{U_1} = \frac{-4{,}6\cdot 5 + 0{,}5\cdot 2 + 1{,}6\cdot 11 + 2{,}2\cdot 8 + 2{,}3\cdot 4}{30} = \frac{22{,}4}{30} = 0{,}7467$$

$$\overline{U_2} = \frac{-1{,}7\cdot 2 - 1{,}1\cdot 6 + 0{,}3\cdot 8 + 0{,}9\cdot 3 + 2{,}6\cdot 1}{20} = \frac{-2{,}3}{20} = -0{,}115$$

$$\overline{U_1^2} = \frac{(-4{,}6)^2\cdot 5 + 0{,}5^2\cdot 2 + 1{,}6^2\cdot 11 + 2{,}2^2\cdot 8 + 2{,}3^2\cdot 4}{30} = \frac{194{,}34}{30} = 6{,}478$$

$$\overline{U_2^2} = \frac{(-1{,}7)^2\cdot 2 + (-1{,}1)^2\cdot 6 + 0{,}3^2\cdot 8 + 0{,}9^2\cdot 3 + 2{,}6^2\cdot 1}{20} = \frac{22{,}93}{20} = 1{,}1475$$

$$S_{U_1}^2 = \overline{U_1^2} - \left(\overline{U_1}\right)^2 = 6{,}478 - 0{,}7467^2 = 5{,}9205$$

$$S_{U_2}^2 = \overline{U_2^2} - \left(\overline{U_2}\right)^2 = 1{,}1475 - (-0{,}115)^2 = 1{,}1343$$

Вернемся к переменной X по формулам: $\bar{x} = h\cdot\overline{U} + C$ и $S_x^2 = h^2 S_U^2$

$$\bar{x}_1 = h\cdot\overline{U}_1 + C = 10\cdot 0{,}7467 + 80 = 87{,}467$$

$$S_{x_1}^2 = h^2 S_{U_1}^2 = 100\cdot 5{,}9205 = 592{,}05$$

$$\bar{x}_2 = h\cdot\overline{U}_2 + C = 10\cdot(-0{,}115) + 80 = 78{,}85$$

$$S_{x_2}^2 = h^2 S_{U_2}^2 = 100\cdot 1{,}1343 = 113{,}43$$

Московский
Технологический
Институт

Найдем общую среднюю и дисперсию:

$$\overline{U}=(-4{,}6\cdot5+0{,}5\cdot2+1{,}6\cdot11+2{,}2\cdot8+2{,}3\cdot4-$$

$$-1{,}7\cdot2-1{,}1\cdot6+0{,}3\cdot8+0{,}9\cdot3+2{,}6\cdot1)\frac{1}{50}=\frac{20{,}1}{50}=0{,}402$$

или $\overline{U}=\dfrac{\overline{U_1}\cdot30+\overline{U_2}\cdot20}{50}=\dfrac{22{,}4-2{,}3}{50}=\dfrac{20{,}1}{50}=0{,}402$

$$\overline{U^2}=((-4{,}6)^2\cdot5+0{,}5^2\cdot2+1{,}6^2\cdot11+2{,}2^2\cdot8+2{,}3^2\cdot4+$$

$$+(-1{,}7)^2\cdot2+(-1{,}1)^2\cdot6+0{,}3^2\cdot8+0{,}9^2\cdot3+2{,}6^2\cdot1)\frac{1}{50}=\frac{217{,}29}{50}=4{,}3458$$

или $\overline{U^2}=\dfrac{\overline{U_1^2}\cdot30+\overline{U_2^2}\cdot20}{50}=\dfrac{194{,}34+22{,}95}{50}=\dfrac{217{,}29}{50}=4{,}3458$

$$S_U^2=\overline{U^2}-\left(\overline{U}\right)^2=4{,}3458-0{,}402^2=4{,}1842$$

Вернемся к переменной X:

$$\bar{x}=h\cdot\overline{U}+C=10\cdot0{,}402+80=84{,}02$$

$$S_x^2=h^2S_U^2=100\cdot4{,}1842=418{,}42$$

Найдем среднюю арифметическую групповых дисперсий:

$$\overline{S_i^2} = \frac{1}{n}\sum_{i=1}^{l} S_i^2 n_i = \frac{1}{50}(S_{x_1}^2 \cdot 30 + S_{x_2}^2 \cdot 20) = \frac{592{,}04 \cdot 30 + 113{,}43 \cdot 20}{50} = \frac{20029{,}3}{50} = 400{,}6$$

Найдем межгрупповую дисперсию:

$$\delta^2 = \frac{1}{n}\sum_{i=1}^{l} (\overline{x_i} - \bar{x})^2 n_i = \frac{1}{50}\left((\overline{x_1} - \bar{x})^2 \cdot 30 + (\overline{x_2} - \bar{x})^2 \cdot 20\right) =$$

$$= \frac{(87{,}467 - 84{,}02)^2 \cdot 30 + (78{,}85 - 84{,}02)^2 \cdot 20}{50} =$$

$$= \frac{11{,}87951 \cdot 30 + 26{,}7289 \cdot 20}{50} = \frac{890{,}9633}{50} = 17{,}82$$

Проверим правило: $S^2 = \overline{S_i^2} + \delta^2$

400,6+17,82=418,42

ВЫПОЛНЯЕТСЯ

Тестовые вопросы:

1. В результате 10 опытов получены следующие выборочные значения: 3; 3; 4; 4; 4; 5; 5; 5; 6; 6. Вариационный ряд имеет вид:

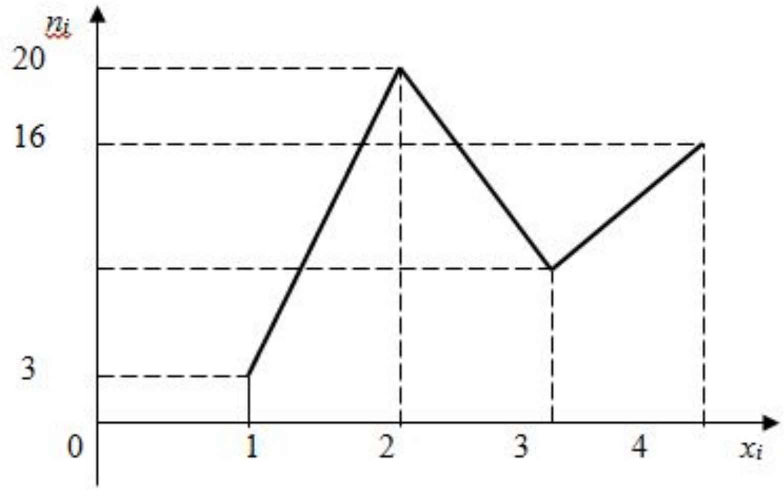
X				
n				

2. В результате 10 опытов получены следующие выборочные значения: 2; 3; 3; 4; 4; 4; 4; 5; 5; 5. Вариационный ряд имеет вид:

X				
n				

Тестовые вопросы:

3. Из генеральной совокупности извлечена выборка объемом n = 49, полигон частот которой



Тогда число выборочных значений (число вариант) для x=3 равно:

а) 11;

б) 9;

в) 10;

Тестовые вопросы:

4. Выборка задана в виде распределения частот ni:

Xi	4	7	8	12	17
ni	2	4	5	6	3

Распределение относительных частот wi имеет вид

а)

Xi	4	7	8	12	17
wi	0,1	0,2	0,25	0,3	0,15

б)

Xi	4	7	8	12	17
wi	0,1	0,2	0,3	0,2	0,2

в)

Xi	4	7	8	12	17
wi	0,2	0,2	0,2	0,2	0,2

г)

Xi	4	7	8	12	17
wi	0,2	0,1	0,3	0,3	0,1

Тестовые вопросы:

5. Показателем вариации признака статистической совокупности является:
а) мода;
б) медиана;
в) относительная частота;
г) дисперсия.

6. Если количественный признак принимает дискретные значения, то соответствующий вариационный ряд называется
а) дискретным
б) интервальным;
в) непрерывным;
г) атрибутивным.

7. Если количественный признак изменяется непрерывно или принимает много значений, то соответствующий вариационный ряд называется
а) дискретным;
б) интервальным;
в) качественным;
г) атрибутивным.

Тестовые вопросы:

8. Средним квадратичным отклонением называется
а) среднее отклонение вариантов от среднего значения.
б) максимальное отклонение вариантов от среднего значения.
в) размах значений признака.
г) минимальное отклонение вариантов от среднего значения

9. Выборочное наблюдение – это
а) сплошное наблюдение;
б) несплошное наблюдение;
в) наблюдение-опрос;
г) наблюдение всей генеральной совокупности

10. Вариант дискретного вариационного ряда, имеющий наибольшую частоту, называется
а) модой;
б) медианой;
в) средней геометрической величиной;
г) средней арифметической величиной

Тестовые вопросы:

11. Статистическое распределение выборки имеет вид

Xi	-2	1	3	4
ni	2	5	6	7

Тогда относительная частота варианты x=3 равна:

а) 0,3
б) 6
в) 0,25
г) 0,1

12. Статистическое распределение выборки имеет вид

Xi	-2	2	3	4
ni	6	4	3	7

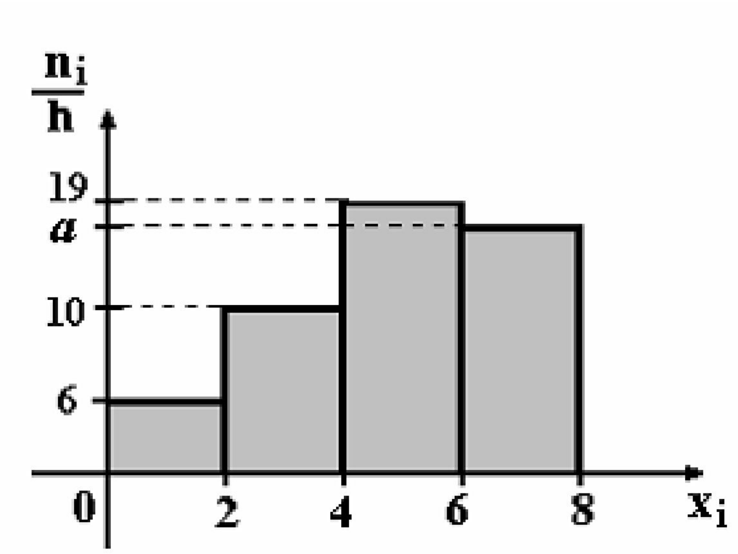
Тогда относительная частота варианты x=2 равна:

а) 0,5
б) 4
в) 0,65
г) 0,2

Московский
Технологический
Институт

Тестовые вопросы:

13. По выборке объема n=100 построена гистограмма частот



Тогда значение ***а*** равно…

а) 15

б) 65

в) 14

г) 16

МТИ
Московский Технологический Институт

Тестовые вопросы:

14. Мода вариационного ряда 5 , 8 , 8 , 9 , 10 , 11 , 13 равна …

а) 13
б) 5
в) 8
г) 9

15. Медиана вариационного ряда 1 , 2 , 5 , 6 , 7 , 7 , 10 равна …

а) 7
б) 1
в) 10
г) 6

16. Для вариационного ряда

Xi	1	3	7
ni	2	5	3

Найдем математическое ожидание, дисперсию, вариацию

Задача для самостоятельного решения:

В таблице приведен ряд моментов t срока работы электрической лампочки в годах. Построить интервальный вариационный ряд, найти среднее значение и дисперсию выборки, размах, коэффициент вариации, построить полигон и гистограмму, эмпирическую функцию распределения, эмпирическую плотность распределения:

0,001	0,001	0,003	0,012	0,046	0,169	0,763	1,620	0,007	1,728
2,067	1,824	0,187	0,646	0,389	1,046	4,672	0,113	1,295	0,500
1,754	0,543	0,074	1,075	0,370	2,612	1,504	0,396	3,245	1,751
0,369	0,534	1,227	0,724	1,307	1,353	2,157	1,000	1,800	0,382
0,500	0,697	0,636	0,365	0,916	1,871	1,134	0,606	0,975	1,043

Ответы к задачам задания 4:

1. Интервал движения автобуса равен 15 минутам. Какова вероятность того, что пассажир на остановке автобуса будет ждать его не более 5 минут? (1/3)

2. Пусть случайная величина X распределена равномерно на отрезке [0,5]. Найти $M(X^2)$. (25/12)

3. Пусть X – нормально распределенная случайная величина с параметрами a=1 и σ=2. Найти вероятности $p_1 = P(|x| > 1)$ и $p_2 = P(|x - 1| < 2)$. (p1=0,979, p2=0,242)

4. Пусть X – непрерывная случайная величина с функцией распределения F(x)=ax+b, сосредоточенной на отрезке (-1;4). Найти a и b. (a=1/5, b=1/5)

5. Суммарная месячная выручка 10 фирм в среднем равна 10000 руб. В 90% случаях эта выручка отклоняется от средней не более чем на 1000 руб. Найти вероятность того, что очередная месячная выручка не превосходит 9500 руб.

(0,2011)

6. Счетчик улавливает частицы, количество которых представляет собой пуассоновский поток с параметром λ=0,1. Чему равно время наблюдения частиц, чтобы с вероятностью 0,9 прибор уловил хотя бы одну частицу? (23)