Величко Арсений Александрович ИВТ 2 курс, 2 группа, 3 подгруппа
Адаев Роман Михайлович ИВТ 2 курс, 1 группа, 1 подгруппа
Предмет: Анализ данных и основы Data Science

Лабораторная работа №2

Содержание

Лабораторная работа №2 **1**
Содержание 1
Задание 1 1
Задание 2 2
Подзадание 1 2
Подзадание 2 2
Задание 3 3

Задание 1

Вариационный ряд при k = 7; шаг = 1.6; n = 60

a_k	a_{k+1}	x_i	m_i
4	5.6	4.8	5
5.6	7.2	6.4	17
7.2	8.8	8	9
8.8	10.4	9.6	15
10.4	12	12.8	4
13.6	15.2	14.4	3

Среднее значение признака:

$$\overline{x} = M = \frac{1}{n}\sum_{i=1}^{n} x_i m_i = 8.69$$

Дисперсия:

$$D = \frac{1}{n}\sum_{i=1}^{n} (x_i - \overline{x})^2 m_i = 6.51$$

Среднее квадратическое отклонение:

$$S = \sigma = \sqrt{D} = 2.55$$

Коэффициент вариации:

Величко Арсений Александрович ИВТ 2 курс, 2 группа, 3 подгруппа
Адаев Роман Михайлович ИВТ 2 курс, 1 группа, 1 подгруппа

$$\nu = \frac{S}{\bar{x}} \cdot 100\% = 29.36\%$$

Вариационный размах:

$$R = x_{max} - x_{min} = 14.4 - 4.8 = 9.6$$

Коэффициент асимметрии:

$$\Lambda = \frac{\sum_{i=1}^{n} (x_i - M)^3 m_i}{n \cdot S^3} = 0.45$$

Эксцесс:

$$E = \frac{\sum_{i=1}^{n} (x_i - M)^4 m_i}{n \cdot S^4} - 3 = - 0.56$$

Задание 2

Дан ряд распределения хозяйств по количеству рабочих на 100 га сельскохозяйственных угодий (n = 60):

12	6	8	6	11	10	11	7	10	12	8	7	7	6	7	8	6	11	9	11
9	10	11	9	15	10	7	8	8	8	11	9	8	7	5	9	7	7	14	11
9	8	7	4	6	7	5	5	10	7	7	5	8	10	10	15	10	10	13	12

Вариационный ряд при k = 7; шаг = 1.6; n = 60

a_k	a_{k+1}	x_i	m_i
4	5.6	4.8	5
5.6	7.2	6.4	17
7.2	8.8	8	9
8.8	10.4	9.6	15
10.4	12	12.8	4
13.6	15.2	14.4	3

Величко Арсений Александрович ИВТ 2 курс, 2 группа, 3 подгруппа
Адаев Роман Михайлович ИВТ 2 курс, 1 группа, 1 подгруппа
Предмет: Анализ данных и основы Data Science

Подзадание 1

Имеются ли различия в обеспеченности хозяйств рабочей силой?
Да. Хозяйства обеспечены рабочей силой достаточно неравномерно.
Эти отличия несущественные или они весьма большие?
Для ответа на этот вопрос рассчитаем коэффициент вариации данного ряда $v = \frac{\sigma}{\bar{x}}$.

В данном ряду $\bar{x} \approx 8,69$; $\sigma = 2,55$, следовательно $v \approx 29\%$.

Эти различия можно считать достаточно большими, поскольку коэффициент вариации данного ряда весьма велик.

Подзадание 2

Является ли рассматриваемое распределение симметричным?
В симметричном ряду значения моды ряда, медианы и среднее арифметическое значение равны, т.е. выполняется условие:

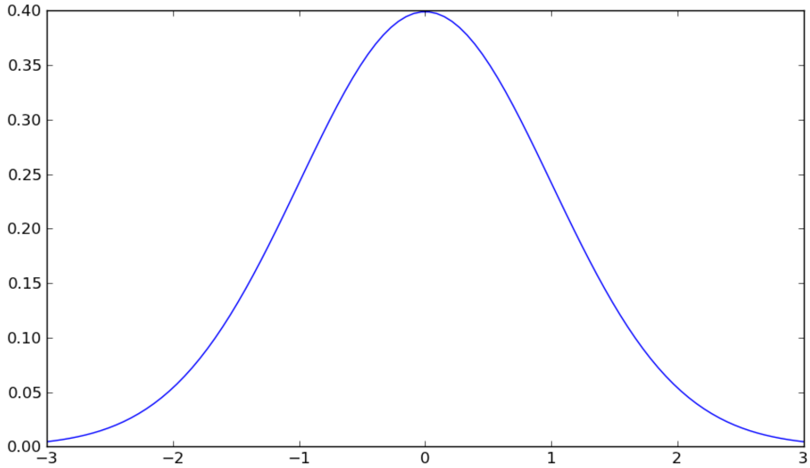
$$\bar{x} = Mo = Me.$$

В данном ряду $\bar{x} \approx 8,69$; $Mo = 7$; $Me = 8$, следовательно условие не выполняется, а значит распределение не является симметричным. Также стоит отметить, что коэффициент асимметрии данного ряда не равен нулю.

Можно ли его считать нормальным?

Для математического обоснования близости эмпирического распределения закону нормального распределения рассчитываются критерии согласия. К таким критериям относятся критерий Колмогорова, критерий Романовского и критерии Пирсона, однако в данном случае достаточно посмотреть на график частот в данном ряде.

Сравним этот график с кривой нормального распределения:



Это сравнение наглядно демонстрирует, что данный ряд не является нормальным, поскольку в ином случае графики имели бы общий вид. В пользу данного утверждения говорит и асимметрия данного ряда, поскольку нормальное распределение симметрично. Кроме того, эксцесс нормального распределения равен нулю, тогда как эксцесс исследуемого ряда равен -0,56.

Задание 3

Устройств с выходом в интернет на одного члена семьи									
0,85	1,59	0,48	0,85	1,48	0,51	1,2	0,84	0	1,16
1,44	1,78	1,03	0,01	0,16	0,2	0,26	1,59	1,12	1,77
1,92	0,32	0,46	1,1	0,36	1,52	0,12	0,75	1,86	1,4
0,53	1,75	0,09	0,24	1,49	1	1,94	0,45	0,53	1,9
1,07	0,49	1,13	1,03	0,86	1,22	1,72	1,84	0,89	0,74
1,84	0,49	0,93	0,01	0,57	0,57	1,08	0,83	0,71	1,77
1,56	0,41	1,55	0,26	0,08	0,24	1,09	1,01	1,79	0,98
0,31	1,02	1,52	0,34	1,45	0,04	0,65	1,14	1,68	1,11
0,83	1,19	0,18	1,58	1,73	1,63	1,81	1,51	1,62	0,05
0,24	1,41	1,01	0,24	0,23	0,3	0,33	1,4	1,69	1,34

Величко Арсений Александрович ИВТ 2 курс, 2 группа, 3 подгруппа
Адаев Роман Михайлович ИВТ 2 курс, 1 группа, 1 подгруппа
Предмет: Анализ данных и основы Data Science

Вычислим математические характеристики вариационного ряда:

X_i	[0; 0,25)	[0,25; 0,5)	[0,5; 0,75)	[0,75; 1)	[1; 1,25)	[1,25; 1,5)	[1,5; 1,75)	[1,75; 2)
m_i	16	14	8	10	18	8	14	12

Среднее значение признака:

$$\overline{x} = M = \frac{1}{n}\sum_{i=1}^{n} x_i m_i = 0,975$$

Дисперсия:

$$D = \frac{1}{n}\sum_{i=1}^{n} (x_i - \overline{x})^2 = 0,35$$

Среднее квадратическое отклонение:

$$S = \sigma = \sqrt{D} = 0,591$$

Коэффициент вариации:

$$\nu = \frac{S}{\overline{x}} \cdot 100\% = 60,67\%$$

Вариационный размах:

$$R = x_{max} - x_{min} = 1,9 - 0 = 1,94$$

Коэффициент асимметрии:

$$\Lambda = \frac{\sum_{i=1}^{n}(x_i - M)^3}{n \cdot S^3} = -0,0018$$

Эксцесс:

$$E = \frac{\sum_{i=1}^{n}(x_i - M)^4}{n \cdot S^4} - 3 = -1,3$$