

Анализ Данных

Лабораторная работа 3: «Вариационный ряд»

Тема: «Числовые характеристики вариационных рядов»

Цель: Вычислить математические характеристики вариационного ряда.

Используемое оборудование: ПК, табличный процессор Excel

Постановка задачи:

1. Ознакомиться с материалами Лекции «Математические характеристики вариационных рядов».
2. Выполнить задания № 1, 2, 3 лабораторной работы

Задание № 1.

Задание № 1.
В качестве изучаемого признака рассматривается число продаж каждого из 26 случайно выбранных продавцов универсама:

16, 12, 15, 15, 23, 9, 15, 13, 14, 14, 21, 15, 14, 17, 27, 15, 16, 12, 16, 19, 14, 16, 17, 13, 14, 14.

Необходимо сделать следующее:

- 1) Построить вариационный ряд;
- 2) Провести анализ построенного вариационного ряда;

Решение:

	A	B	C	D	E	F
1	Ряд	n	Варианты	Частоты	Накопленные частоты	Накопленные частоты
2		9	31	9	1	0
3		12		12	2	0,032258065
4		12		13	3	0,096774194
5		13		14	8	0,193548387
6		13		15	6	0,451612903
7		13		16	4	0,64516129
8		14		17	2	0,774193548
9		14		19	1	0,838709677
10		14		21	2	0,870967742
11		14		23	1	0,935483871
12		14		27	1	0,967741935
13		14		28	0	1
14		14				
15		14				
16		15				
17		15	Среднее арифметическое			
18		15	15,58064516			
19		15				
20		15				
21		15	Дисперсия		Разности для дисперсии	Разности для асимметрии
22		16	11,98543184		43,30489074	-284,9741197
23		16			12,82101977	-45,90752241
24		16	Среднее квадратичное отклонение		12,82101977	-45,90752241
25		16	3,461998244		6,659729448	-17,18639858
26		17			6,659729448	-17,18639858
27		17	Коэффициент вариации		6,659729448	-17,18639858
28		19	22%		2,498439126	-3,949145715
29		21			2,498439126	-3,949145715
30		21	Коэффициент асимметрии		2,498439126	-3,949145715
31		23	0,007934839		2,498439126	-3,949145715
32		27			2,498439126	-3,949145715
33			Экссесс		2,498439126	-3,949145715

25				
26		n		
27		60		
28				
29				
30	Интервалы	Число регионов	Накопленные частоты	Накопленные частоты
31	До 60	10	0	0
32	60-70	29	10	0,166666667
33	70-80	2	39	0,65
34	80-90	13	41	0,683333333
35	90-100	0	54	0,9
36	Свыше 100	6	54	0,9
37				
38				
39				
40				
41				
42			Медиана	Мода
43			70,34482759	64,13043478

Задание № 3.

Задание № 3

Таблица 2

Интервалы	Число регионов
До 60	10
60-70	29
70-80	2
80-90	13
90-100	-
Свыше 100	6

Определите начало первого интервала и правую границу последнего интервала.

Решение:

	А	В
1	Интервалы	Число регионов
2	До 60	10
3	60-70	29
4	70-80	2
5	80-90	13
6	90-100	0
7	Свыше 100	6
8		
9	Начало первого интервала	Правая граница последнего
10	0	∞
11		

3.

3. Для данных Задания № 1 определите оптимальную величину интервала и представьте ряд из этого задания в виде интервального ряда.

Решение:

Интервальный ряд		Частоты	Накоп. Частоты	Накоп. Частости
8,03207	11,13207	1	0	0
11,13207	14,23207	13	1	0,032258065
14,23207	17,33207	12	14	0,451612903
17,33207	20,43207	1	26	0,838709677
20,43207	23,53207	3	27	0,870967742
23,53207	26,63207	0	30	0,967741935
26,63207	29,73207	1	30	0,967741935
29,73207	32,83207	0	31	1

4.

4. Интервальные ряды бывают с равными и неравными интервалами. Иногда при группировке с равными интервалами сначала определяют число интервалов (групп) z при заданном объеме совокупности, используя формулу:

$$L = 2 \ln n,$$

и тогда k в формуле Стелжера вычисляется по формуле

$$k = \frac{x_{\max} - x_{\min}}{L}$$

Для Задания 1 вычислите оптимальную величину интервала по данной формуле и сравните его с интервалом, вычисленным по формуле Стелжера.

Решение:

21	Ряд	n		8,03207	11,13207
22		9	31	11,13207	14,23207
23		12		14,23207	17,33207
24		12	k	17,33207	20,43207
25		13	5,807582086	20,43207	23,53207
26		13	6	23,53207	26,63207
27		13		26,63207	29,73207
28		14	Δ	29,73207	32,83207
29		14	3,09939657		
30		14	3,1		
31		14			
32		14	$x_{изч}$		
33		14	8,032069652		
34		14			
35		14			
36		15			
37		15			
38		15	L		
39		15	6,867974409		
40		15			
41		15	$\Delta 2$		
42		16	2,620860086		
43		16			
44		16	Интервал Стеджерса = 3,09939657 Интервал по формуле = 2,620860086 Интервал Стеджерса > Интервал по формуле		
45		16			
46		17			
47		17			
48		19			
49		21			
50		21			
51		23			
52		27			

5.

5. Для данных таблицы 1 накопленные частоты и расположите их в таблице в восходящем порядке и в нисходящем порядке. На что они указывают? Поясните.

Решение:

Варианты	Накопленные частоты	Накопленные частоты
9	0	31
12	1	30
13	3	29
14	6	27
15	14	26
16	20	24
17	24	20
19	26	14
21	27	6
23	29	3
27	30	1
28	31	0

Накопленные частоты показывают, сколько единиц совокупности имеют значение признака не больше, чем рассматриваемое. Они определяются путем последовательного суммирования частот (частостей) предшествующих интервалов и используются для построения куммулянты

6.

6. Постройте полигон распределения для Задания 1

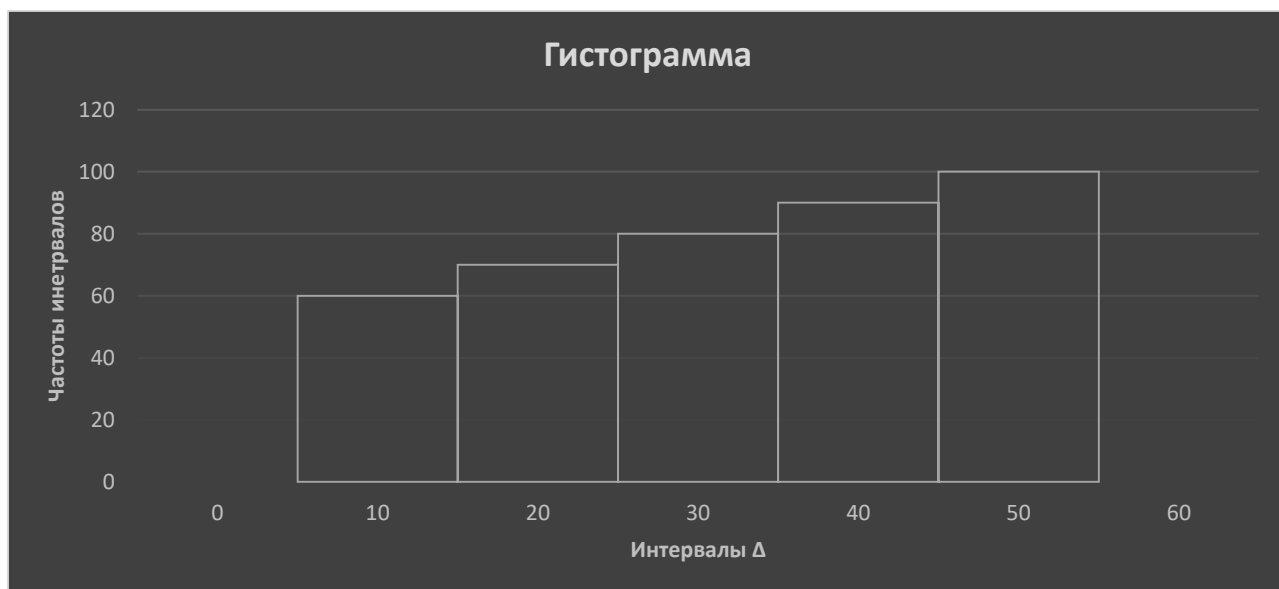
Решение:



7.

7. Постройте гистограмму и кумулянту для данных таблицы 2

Решение:

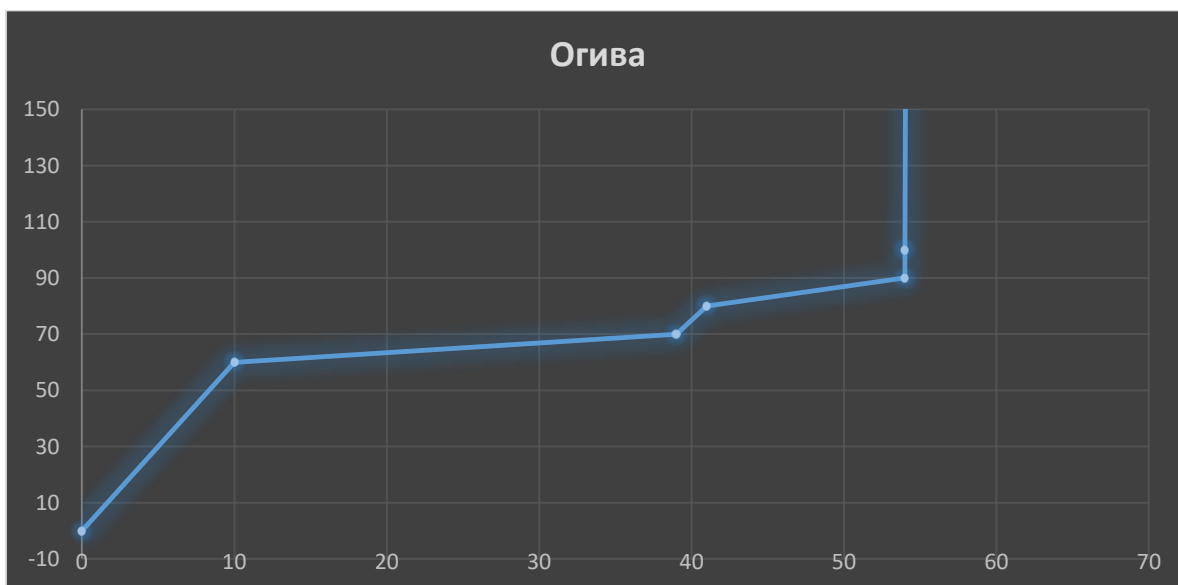




8.

8. Постройте огиву для данных таблицы 2
Огива строится аналогично кумулянте (но по оси абсцисс откладываются накопленные частоты (частости), а по оси ординат – значения признака).

Решение:



9.

9. Для данных Задания № 1 вычислить:
 1) 25-й, 50-й и 90-й перцентили в вариационном ряду.

Решение:

125	25 Персентиль
126	72,5
127	
128	50 Персентиль
129	85
130	
131	90 Персентиль
132	5050

10.

10. По данным таблицы 2 вычислить медиану (используйте формулу для нахождения медианы внутри медианного интервала).

Решение:

135	Медиана
136	70,34482759
137	

11.

11. По данным таблицы 2 вычислить моду (используйте формулу для нахождения моды внутри модального интервала).

Решение:

138	Мода
139	64,13043478

12.

12. Вычислите среднюю арифметическую для данных Задания № 1 по формулам:

- Средней арифметической;
- Средней арифметической взвешанной.

Решение:

143	Варианты	Частоты	Взвешенная средняя арифметическая
144	9	1	9
145	12	2	24
146	13	3	39
147	14	8	112
148	15	6	90
149	16	4	64
150	17	2	34
151	19	1	19
152	21	2	42
153	23	1	23
154	27	1	27
155	28	0	Результат
156			15,58064516
157			
158			Средняя арифметическая
159			15,58064516

13.

13. Проверить насколько медиана и средняя арифметическая чувствительна к положению крайних значений ряда значений (для данных Задания № 1).

Решение:

Среднее арифметическое является очень чувствительным к изменениям крайних значений ряда значений, так как в формуле

$$x = \frac{\sum_{i=1}^n x_i}{n}$$

присутствует сумма, а она в свою очередь будет меняться при каждом изменении крайнего значения, что и приведет к изменению среднего арифметического.

“Медиана - это такое значение признака, которое делит упорядоченное множество пополам, так что одна половина всех значений становится меньше медианы, а другая больше” - из определения медианы мы можем сделать вывод что значение медианы не будет чувствительным к изменениям крайних значений ряда, так как от этого количество значений не изменится, а, следовательно, и медиана тоже останется неизменной.

14.

14. Может ли быть в одном ряду несколько мод? Обоснуйте.

Решение:

Иногда в совокупности встречается более чем одна мода. В этом случае можно сказать, что совокупность мультимодальна. Из структурных средних величин только мода обладает таким уникальным свойством. Как правило, мультимодальность указывает на то, что набор данных не подчиняется нормальному распределению.

Вывод:

В итоге, используя табличный процессор Excel и его встроенные функции, нам удалось вычислить математические характеристики вариационного ряда и выполнить поставленные задачи.