

Образовательная организация высшего образования
(Частное учреждение)
«Международная Академия Бизнеса и Новых Технологий»
(МУБиНТ)
Кафедра_Менеджмент_
Специальность_Интернет-маркетинг и управление маркетинговой
деятельностью

ОТЧЕТ

По самостоятельной работе по предмету Статистика

Выполнил студент 1 курса

Группы № 22-МБ568

Зимин Дмитрий Николаевич

Руководитель лабораторной работы

Гультяев Вячеслав Евгеньевич

Отчет защищен с оценкой _____

Ярославль, 2022

Задание 1 – Решение примеров по образцу

Пример 13

B46 fx =СУММ(B40:B45)			
	A	B	C
38			
39	Число поездок в день	Количество граждан, совершающих в день данное количество поездок	Накопленные частоты
40	0	221	221
41	1	146	367
42	2	434	801
43	3	405	1206
44	4	302	1508
45	5	112	1620
46	Итого:	1620	
47			

Рисунок 1

Для представленного ряда наиболее часто встречающимся вариантом является 2 поездки в день, то есть мода равна 2. При четном числе элементов вариационного ряда за медиану принимают среднюю арифметическую из двух центральных вариантов упорядоченного ряда. Так по данным, представленным на Рисунке из 1620 элементов (общей сумме поездок) 810 и 811 будут являться центральными: они оба равны 3. Расчет медианы проведен следующим образом: $(3+3)/2 = 3$, то есть медиана равна 3.

Пример 14

C67 fx =G61/B61/C65^4-3							
	A	B	C	D	E	F	G
52	Зарботная плата, руб. в час $x_i - x_{i+1}$	Кол-во рабочих, чел. f_i	Центр интервала (средняя з/п за час), руб. за час x'_i	Средняя з/п за час по всем сотрудникам, руб. за час $x'_i \cdot f_i$	$(x'_i - \bar{x})^2 \cdot f_i$	$(x'_i - \bar{x})^3 \cdot f_i$	$(x'_i - \bar{x})^4 \cdot f_i$
53							
54	300-400	3	350	1050	388103,538	-139592078,984	50208118731,193
55	400-500	5	450	2250	337161,811	-87553308,885	22735617307,295
56	500-600	6	550	3300	152981,270	-24427654,325	3900544803,529
57	600-700	9	650	5850	32052,549	-1912813,434	114151769,424
58	700-800	18	750	13500	29266,389	1180096,338	47584529,751
59	800-900	19	850	16150	374118,106	52497218,120	7366545123,229
60	900-1000	2	950	1900	115509,886	27759633,782	6671266828,258
61	Сумма	62		44000	1429193,548	-172048907,388	91043829092,679
62							
63		Среднее значение Средней з/п за час по всем сотрудникам, руб. за час	709,677				
64							
65		σ	151,827				
66		As	-0,793				
67		Ex	-0,236				
68							

Рисунок 2

Данные представлены в виде интервального ряда (шаг 100 руб.). Для определения показателей асимметрии и крутизны необходимо найти среднее и среднее квадратическое отклонение распределения. Среднее значение вычислим по формуле средней арифметической взвешенной, для этого необходимо найти отношение средней заработной платы за час по всем сотрудникам к числу сотрудников.

$x_{cp} = 44000/62 = 709,677$ руб. за час. (средний уровень заработной платы 62 сотрудников за час составляет 709,68 руб.).

Далее найдем среднее квадратическое отклонение (σ).

$$\sigma = \sqrt{1429193,548/62} = 151,827.$$

Коэффициент асимметрии рассчитан следующим образом:

$$As = \frac{-172048907,388}{62 \cdot 152,827^3} = -0,793$$

В данном случае As меньше 0, то есть в вариационном ряду преобладают варианты, которые меньше средней (x_{cp}) и распределение имеет более длинную ветвь слева (скошено влево): левостороннее асимметричное распределение.

Далее найдем эксцесс распределения:

$$Ex = \frac{91043829092,679}{62 \cdot 152,827^4} = -0,236$$

В примере показатель Ex меньше 0, значит крутизна распределения плосковершинная.

Пример 15

B78		fx = =(A72*B72+A73*B73+A74*B74+A75*B75)/(B72+B73+B74+B75)		
	A	B	C	D
70				
71	Признак	Частота признака		
72		1	7	
73		2	21	
74		3	14	
75		4	14	
76				
77				
78	Среднее значение	2,625		
79				

Рисунок 3

В представленных данных чаще всего встречается значение признака 2 (его частота 21 является максимальной из всех представленных в задаче). Так, Мода = 2.

Среднее значение (x_{cp}) рассчитано по формуле средней взвешенной (рисунок 3). Таким образом, $x_{cp} = 2,625 > Mo = 2$, таким образом, в данном случае мы имеем правостороннюю асимметрию, то есть показатель $As > 0$ (Рисунок 4).

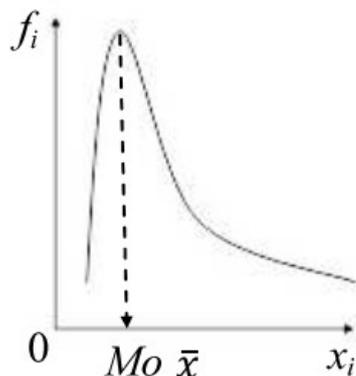


Рисунок 4

Пример 16

B78		fx = $=(A72*B72+A73*B73+A74*B74+A75*B75)/(B72+B73+B74+B75)$		
	A	B	C	D
70				
71	Признак	Частота признака		
72		1	5	
73		2	7	
74		3	10	
75		4	2	
76				
77				
78	Среднее значение	2,375		
79				
80				
81				

Рисунок 5

В представленных данных чаще всего встречается значение признака 3 (его частота 10 является максимальной из всех представленных в задаче). Так, Мода = 3.

Среднее значение (x_{cp}) рассчитано по формуле средней взвешенной (рисунок 5). Таким образом, $x_{cp} = 2,375 < Mo = 3$, таким образом, в данном случае мы имеем левостороннюю асимметрию, то есть показатель $As < 0$ (Рисунок 6).

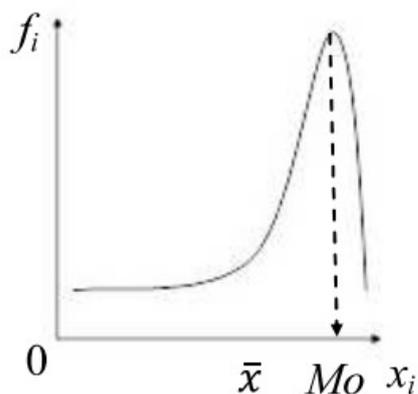


Рисунок 6

Пример 17

	A	B	C	D
70				
71	Признак	Частота признака		
72	1	2		
73	2	14		
74	3	10		
75	4	6		
76				
77				
78	Среднее значение	2,625		
79				

Рисунок 7

В представленных данных чаще всего встречается значение признака 2 (его частота 14 является максимальной из всех представленных в задаче). Так, Мода = 2.

Среднее значение (x_{cp}) рассчитано по формуле средней взвешенной (рисунок 7). Таким образом, $x_{cp} = 2,625 > Mo = 2$, таким образом, в данном случае мы имеем правостороннюю асимметрию, то есть показатель $As > 0$ (Рисунок 8).

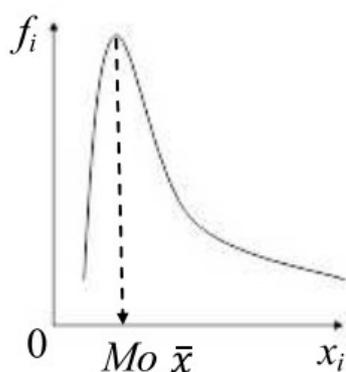


Рисунок 8

Пример 18

G94		fx	
	A	B	C
80			
	Среднедушевой совокупный доход, \$ в мес.	Количество граждан в соответствующей группе дохода, млн. чел.	Накопленные частоты
81			
82	до 5 000	10	10
83	5 000 - 25 000	26	36
84	25 000 - 45 000	31	67
85	45 000 - 100 000	17	84
86	свыше 100 000	16	100
87	Итого	100	-
88			

Рисунок 9

В задаче требуется найти децильный коэффициент дифференциации. Население упорядочено по среднедушевому совокупному доходу (по возрастанию, рисунок 9). В совокупности население составляет 100 млн. чел. (100%).

Необходимо рассчитать первый и девятый дециль.

Возьмем первую 1/10 часть наименее обеспеченного населения (от 0% до 10%: здесь максимальная граница интервала 10%, поэтому гражданин соответствующий 10% будет иметь максимальный доход из этой группы.

$$\text{Первый дециль} = D_1 = x_{D_1} + \frac{h_{D_1} * 0,1 * \sum f_i - S_{D_1-1}}{f_{D_1}}$$

Для нахождения интервала, содержащего первый дециль воспользуемся накопленными частотами, 10% попадает в первый интервал, так как 10% населения получает доход до 5 000 \$ в мес. Ширина данного интервала определяется по ближайшему к нему интервалу 5 000 - 25 000 \$ в мес. Ширина = 20 000 \$ в мес. и интервал имел бы вид -15 000 – 5 000 \$ в мес., но так как доход не может быть отрицательным, интервал принимает вид 0 – 5000 \$ в мес. Ширина интервала равна 5 000 \$ в мес. ($x_{D_1}=0$, $h_{D_1}=5000$). Частота в первом децильном интервале равна $f_{D_1} = 10$, накопленная частота в предыдущем децильном интервале $S_{D_1-1}=0$.

$D_1 = 0 + 5\,000 * (0,1*100-0)/10 = 5\,000$ \$ в мес., то есть максимальный доход 10% наименее обеспеченного населения составляет 5 000 \$ в мес.

Далее произведем расчет минимального дохода 10% наиболее обеспеченного населения (девятый дециль). 90%-й децильный интервал – это интервал с

доходом более 100 000 \$ в мес. (84-100% населения получает доход в размере более 101 000 \$ в мес.). Ширина данного интервала определяется по ближайшему к нему интервалу 45 000 - 100 000 \$ в мес. Ширина = 55 000 \$ в мес., интервал принимает вид 100 000 – 155 000 \$ в мес. ($x_{D_9}=100000$, $h_{D_9}=55000$). Частота в девятом дециальном интервале равна $f_{D_9} = 16$, накопленная частота в предыдущем децильном интервале $S_{D_9-1} = 84$.

$D_9 = 100\ 000 + 55\ 000 \cdot (0,9 \cdot 100 - 84) / 16 = 120\ 625$ \$ в мес., то есть минимальный доход 10% наиболее обеспеченного населения составляет 120 625 \$ в мес.

Децильный коэффициент дифференциации рассчитан по формуле:

$$K_d = D_9 / D_1 = 120\ 625 / 5000 = 24,125$$

Таким образом, минимальные доходы 10% наиболее обеспеченных граждан превышают в 24,125 раз максимальные доходы 10% наименее обеспеченного уровня населения.

Пример 19

F90		fx	
	A	B	C
80			
81	Среднедушевой совокупный доход, \$ в мес.	Количество граждан в соответствующей группе дохода, млн. чел.	Накопленные частоты
82	до 5 000	7	7
83	5 000 - 21 000	30	37
84	21 000 - 41 000	37	74
85	41 000 - 101 000	20	94
86	свыше 101 000	6	100
87	Итого	100	-
88			

Рисунок 10

В задаче требуется найти децильный коэффициент дифференциации. Население упорядочено по среднедушевому совокупному доходу (по возрастанию, рисунок 10). В совокупности население составляет 100 млн. чел. (100%).

Необходимо рассчитать первый и девятый дециль.

Возьмем первую 1/10 часть наименее обеспеченного населения (от 0% до 10%: здесь максимальная граница интервала 10%, поэтому гражданин соответствующий 10% будет иметь максимальный доход из этой группы.

$$\text{Первый дециль} = D_1 = x_{D_1} + \frac{h_{D_1} * 0,1 * \sum f_i - S_{D_1-1}}{f_{D_1}}$$

Для нахождения интервала, содержащего первый дециль воспользуемся накопленными частотами, 10% попадает во второй интервал, так как 37% населения получает доход до 21 000 \$ в мес. Ширина данного интервала 5 000 - 21 000 \$ в мес. Ширина = 16 000 \$ в мес. ($x_{D_1}=5000$, $h_{D_1}=16000$). Частота в первом дециальном интервале равна $f_{D_1} = 30$, накопленная частота в предыдущем дециальном интервале $S_{D_1-1}=7$.

$D_1 = 5000 + 16000 * (0,1*100-7)/30 = 6600$ \$ в мес., то есть максимальный доход 10% наименее обеспеченного населения составляет 6600 \$ в мес.

Далее произведем расчет минимального дохода 10% наиболее обеспеченного населения (девятый дециль). 90%-й децильный интервал – это интервал с доходом 41 000 - 101 000 \$ в мес. (74-94% населения получает доход в размере от 41 000 до 101 000 \$ в мес.). Ширина данного интервала 60 000 \$ в мес. ($x_{D_9}=41000$, $h_{D_9}=60000$). Частота в девятом дециальном интервале равна $f_{D_9} = 20$, накопленная частота в предыдущем дециальном интервале $S_{D_9-1}=74$.

$D_9 = 41000 + 60000*(0,9*100-74)/20 = 89000$ \$ в мес., то есть минимальный доход 10% наиболее обеспеченного населения составляет 89 000 \$ в мес.

Децильный коэффициент дифференциации рассчитан по формуле:

$$K_d = D_9/D_1 = 89000 / 6600 = 13,485$$

Таким образом, минимальные доходы 10% наиболее обеспеченных граждан превышают в 13,485 раз максимальные доходы 10% наименее обеспеченного уровня населения.

Пример 20

	А	В	С
80			
	Численность населения, чел.	Количество населенных пунктов с заданной численностью населения (шт)	Накопленные частоты
81			
82	до 50 000	8	8
83	50 000 - 150 000	10	18
84	150 000 - 250 000	14	32
85	250 000 - 450 000	18	50
86	450 000 - 750 000	25	75
87	свыше 750 000	25	100
88	Итого	100	-

Рисунок 11

По данным задачи необходимо найти квантильный коэффициент дифференциации. Все населенные пункты упорядочены по численности (по возрастанию, рисунок 11). Всего в выборке 100 населенных пунктов, которые составляют 100%.

Для вычисления квантильного коэффициента необходимо найти первый (20%) и четвертый (80%) квантили. Первый квантиль рассчитывается по формуле:

$$\text{Первый квантиль} = K_1 = x_{K_1} + \frac{h_{K_1} * 0,2 * \sum f_i - S_{K_1-1}}{f_{K_1}}$$

Первый квантиль показывает максимальную численность населения в 1/20 части (0-20%) самых малочисленных городов страны.

Для нахождения интервала, содержащего первый квантиль, используются накопленные частоты. В данном случае, первый квантиль содержится в третьем интервале (18-32% населенных пунктов имеет численность 150 000 – 250 000 чел.). Ширина данного интервала (h_{K_1}) 100 000 чел., нижняя граница (x_{K_1}) 150 000 чел., частота в первом квантильном интервале (f_{K_1}) 14, накопленная частота в предшествующем интервале составила 18 ($S_{K_1-1}=18$). Так, первый квантиль равен:

$$K_1 = 150\,000 + 100\,000 * (0,2 * 100 - 18) / 14 = 164\,285 \text{ чел.}$$

Так, максимальная численность населения 20% наименее населенных городов составляет 164 285 чел.

Четвертый квантиль (80%) показывает минимальную численность населения в (80-100%) самых крупнонаселенных городах.

$$\text{Четвертый квантиль} = K_4 = x_{K_4} + \frac{h_{K_4} * 0,8 * \sum f_i - S_{K_4-1}}{f_{K_4}}$$

Для нахождения интервала, содержащего четвертый квантиль, используются накопленные частоты. В данном случае, четвертый квантиль содержится в шестом интервале (75-100% населенных пунктов имеет численность более 750 000 чел.). Ширина данного интервала определяется по ближайшему к нему интервалу 450 000 – 750 000 чел., ширина (h_{K_4}) = 300 000 чел. и интервал принимает вид 750 000 – 1 050 000 чел., нижняя граница (x_{K_4}) 750 000, частота в четвертом квантильном интервале (f_{K_4}) 25, накопленная частота в предшествующем интервале 75 ($S_{K_4-1}=75$). Так, четвертый квантиль равен:

$$K_4 = 750\,000 + 300\,000 * (0,8 * 100 - 75) / 25 = 810\,000 \text{ чел.}$$

Так, минимальная численность населения 20% самых густонаселенных городов составляет 810 000 чел.

Квантильный коэффициент рассчитан следующим образом:

$$K_k = K_4/K_1 = 810\,000 / 164\,285 = 4,9305$$

Таким образом, минимальная численность населения 20% наиболее крупнонаселенных городов в 4,9305 раз превышает максимальную численность населения 20% самых малонаселенных городов.

Пример 21

	А	В	С
80			
81	Численность населения, чел.	Количество населенных пунктов с заданной численностью населения (шт)	Накопленные частоты
82	до 100 000	25	25
83	100 000 - 200 000	22	47
84	200 000 - 300 000	19	66
85	300 000 - 500 000	16	82
86	500 000 - 800 000	12	94
87	свыше 800 000	6	100
88	Итого	100	-

Рисунок 12

По данным задачи необходимо найти квантильный коэффициент дифференциации. Все населенные пункты упорядочены по численности (по возрастанию, рисунок 12). Всего в выборке 100 населенных пунктов, которые составляют 100%.

Для вычисления квантильного коэффициента необходимо найти первый (20%) и четвертый (80%) квантили. Первый квантиль рассчитывается по формуле:

$$\text{Первый квантиль} = K_1 = x_{K_1} + \frac{h_{K_1} * 0,2 * \sum f_i - S_{K_1-1}}{f_{K_1}}$$

Первый квантиль показывает максимальную численность населения в 1/20 части (0-20%) самых малочисленных городов страны.

Для нахождения интервала, содержащего первый квантиль, используются накопленные частоты. В данном случае, первый квантиль содержится в первом интервале (0-25% населенных пунктов имеет численность до 100 000 чел.). Ширина данного интервала (h_{K_1}) 100 000 чел., нижняя граница (x_{K_1}) 0, частота в первом квантильном интервале (f_{K_1}) 25, накопленная частота в предшествующем интервале отсутствует ($S_{K_1-1}=0$). Так, первый квантиль равен:

$$K_1 = 0 + 100\,000 * (0,2*100-0)/25 = 80\,000 \text{ чел.}$$

Так, максимальная численность населения 20% наименее населенных городов составляет 80 000 чел.

Четвертый квантиль (80%) показывает минимальную численность населения в (80-100%) самых крупнонаселенных городах.

$$\text{Четвертый квантиль} = K_4 = x_{K_4} + \frac{h_{K_4} * 0,8 * \sum f_i - S_{K_4-1}}{f_{K_4}}$$

Для нахождения интервала, содержащего четвертый квантиль, используются накопленные частоты. В данном случае, четвертый квантиль содержится в четвертом интервале (66-82% населенных пунктов имеет численность 300 000 – 500 000 чел.). Ширина данного интервала (h_{K_4}) 200 000 чел., нижняя граница (x_{K_4}) 300 000, частота в четвертом квантильном интервале (f_{K_4}) 16, накопленная частота в предшествующем интервале 66 ($S_{K_4-1}=66$). Так, четвертый квантиль равен:

$$K_4 = 300\,000 + 200\,000 * (0,8*100-66)/16 = 475\,000 \text{ чел.}$$

Так, минимальная численность населения 20% самых густонаселенных городов составляет 475 000 чел.

Квантильный коэффициент рассчитан следующим образом:

$$K_k = K_4/K_1 = 475\,000 / 80\,000 = 5,9375$$

Таким образом, минимальная численность населения 20% наиболее крупнонаселенных городов в 5,9375 раз превышает максимальную численность населения 20% самых малонаселенных городов.