

## Отчёт по задаче\_1

57.17.	Тип рекламы		
Повто- рения	A	B	C
1	35,6	31,4	33,5
2	34,5	28,5	30,3
3	36,8	26,0	35,7
4	34,1	36,2	21,6
5	38,1	32,7	20,4
6	36,9	35,5	30,9
7	33,8	36,0	27,3
8	34,4	40,7	35,8

57. Изучается влияние типа рекламного буклета на прибыль компании-производителя. Буклет каждого из трех типов разослан восьми компаниям, потребляющим рекламируемую продукцию. Результаты (прибыль от заказов потребителей) приведены в таблице. Использовать метод дисперсионного анализа при уровне значимости 0,05 для проверки гипотезы о том, что средняя прибыль не зависит от типа рекламы.

### 1. Данные заносим в Excel

	A1	A2	A3
1	35,6	31,4	33,5
2	34,5	28,5	30,3
3	36,8	26	35,7
4	34,1	36,2	21,6
5	38,1	32,7	20,4
6	36,9	35,5	30,9
7	33,8	36	27,3
8	34,4	40,7	35,8

### 2. Имеем таблицу с дисперсиями и проверяем рав-во дисперсий критерием Бартлета

Однофакторный дисперсионный анализ

ИТОГИ

Группы	Счет	Сумма	Средне е	Дисперс ия
Столбе ц 1	8	284,2	35,525	2,49642 9
Столбе ц 2	8	267	33,375	22,2221 4
Столбе ц 3	8	235,5	29,4375	35,2512 5

критерия Бартлетта.

Число наблюдений на каждом уровне  $n_i$  («Счет»), несмещенные оценки групповых средних  $\bar{y}^{(i)} = \sum_{k=1}^{n_i} y_k^{(i)} / n_i$  («Средние») и групповых дисперсий

$s_i^2 = \sum_{k=1}^{n_i} (y_k^{(i)} - \bar{y}^{(i)})^2 / (n_i - 1)$  («Дисперсии») рассчитаны программой ( $i = 1, 2, 3$ ).

Для расчета числового значения статистики Бартлетта

$$\Phi_{v-1} = q \left( (n_1 - 1) \ln \frac{s_{\text{ост}}^2}{s_1^2} + (n_2 - 1) \ln \frac{s_{\text{ост}}^2}{s_2^2} + (n_3 - 1) \ln \frac{s_{\text{ост}}^2}{s_3^2} \right),$$

343

где  $v = 3$  — число уровней фактора,

$$q = \frac{1}{1 + \frac{1}{3(v-1)} \left( \frac{1}{n_1-1} + \frac{1}{n_2-1} + \frac{1}{n_3-1} - \frac{1}{(n_1-1) + (n_2-1) + (n_3-1)} \right)},$$

$$s_{\text{ост}}^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2 + (n_3-1)s_3^2}{(n_1-1) + (n_2-1) + (n_3-1)},$$

вычислим последовательно:

$$q = \frac{1}{1 + \frac{1}{3(3-1)} \left( \frac{1}{4-1} + \frac{1}{6-1} + \frac{1}{5-1} - \frac{1}{(4-1) + (6-1) + (5-1)} \right)} = 0,896;$$

$$s_{\text{ост}}^2 = \frac{(4-1)1633,33 + (6-1)1466,67 + (5-1)1470}{(4-1) + (6-1) + (5-1)} = 1509,44,$$

тогда числовое значение статистики Бартлетта  $\Phi_{v-1}$  будет равно

$$0,896 \left( (4-1) \ln \frac{1509,44}{16,33,33} + (6-1) \ln \frac{1509,44}{1466,67} + (5-1) \ln \frac{1509,44}{1470} \right) = 0,012.$$

Статистика Бартлетта  $\Phi_{v-1}$  в предположении справедливости гипотезы о равенстве групповых генеральных дисперсий имеет распределение, близкое к  $\chi_{v-1}^2$ . При  $\alpha = 0,05$  критическая точка  $\chi_{v-1; \alpha}^2 = \chi_{2; 0,05}^2 = 5,991$ . Поскольку наблюдаемое числовое значение статистики Бартлетта оказалось меньше критической точки, нет оснований отвергнуть проверяемую гипотезу. Таким образом, второе требование, предъявляемое к результатам наблюдений, выполнено.

H0- групповые генеральные дисперсии количества голосов равны

H1- не равны

$v$  = число уровней фактора =

3

$q$  = 0,940299

$S^2_{\text{ост}}$  = 19,98994

Стат Бартлетта  $\Phi$  = 9,262554

$\chi^2$  = 9,21034

Стат Бартлетта  $\Phi > \chi^2$ , отвергаем проверяемую гипотезу. Таким образом, второе требование, предъявляемое к результатам наблюдений, не выполнено, но все равно продолжаем расчёты.

Так как гипотеза о равенстве дисперсий не подтверждается, следовательно применение дисперсионного анализа не обоснованно

### 3) .Выводы по таблице дисп. анализ

Источник вариации	SS	df	MS	F	P- Значение	F критическое
Между группами	152,4908	2	76,24542	3,814189	0,038631	3,4668
Внутри групп	419,7888	21	19,98994			
Итого	572,2796	23				

Оценка Дисперсионного анализа по столбцам					Оценка дисперсионного анализа по строкам				
H0 рекламный букет не влияет на прибыль					H0 не зависят				
H1 рекламный букет влияет на прибыль					H1 зависят				
P-значение =	0,03863	, больше 0,01, а значит принимаем гипотезу H0 с вер. 96%			Принимаем гипотезу H0, так как	0,03863	> 0,01		

### 4). Можем ещё проверить по строкам

H0: рекламный букет не влияет на  
прибыль

H1: рекламный букет влияет на прибыль

При запуске процедуры – выбираем по  
строкам

Однофакторный дисперсионный анализ

ИТОГИ

Группы	Счет	Сум ма	Средн ее	Дисперс ия
		100,		
Строка 1	3	5	33,5	4,41
Строка 2	3	93,3	31,1	9,48
			32,833	35,3233
Строка 3	3	98,5	33	3
			30,633	62,3033
Строка 4	3	91,9	33	3
Строка 5	3	91,2	30,4	82,29
		103,	34,433	9,85333
Строка 6	3	3	33	3

			32,366	20,4633
Строка 7	3	97,1	67	3
		110,	36,966	10,9433
Строка 8	3	9	67	3

#### Дисперсионный анализ

Источник вариации	SS	df	MS	F	P- Значени е	F критическо е
	102,14		14,592	0,49661		
Между группами	63	7	32	9	0,823526	2,657197
	470,13		29,383			
Внутри групп	33	16	33			
	572,27					
Итого	96	23				

Р-значение = **0,823526** > **0,1** => принимаем гипотезу H<sub>0</sub> –рекламный букет не влияет на прибыль – здесь рав-во дисперсий не проверяем

## Отчёт по задаче\_2

## Тема Двухфакторный дисп. анализ с повторениями

### Постановка задачи

424. При исследовании зависимости размера запрашиваемого кредита  $Y$  от профессии ( $A^{(1)}$  — менеджеры,  $A^{(2)}$  — продавцы,  $A^{(3)}$  — инженеры), пола ( $B^{(1)}$  — женщины,  $B^{(2)}$  — мужчины) и их взаимодействия было выделено случайным образом 18 групп, которые приписывались в равных количествах шести комбинациям профессий и пола. Известны данные о средних размерах запрашиваемых кредитов в группах:

$A \backslash B$	$B^{(1)}$			$B^{(2)}$		
$A^{(1)}$	650 000	600 000	650 000	620 000	600 000	590 000
$A^{(2)}$	620 000	630 000	600 000	580 000	580 000	580 000
$A^{(3)}$	521 000	520 000	$(50+i) \cdot 10^4$	510 000	520 000	500 000

$q =$  0.9411764705882353

$s^2 =$  2,39E+09

$F1 =$  0.3910826641108622

$q =$  0.9183673469387754

$s^2 =$  1768500000.1333334

$F2 =$  6.234472291748413

$\chi^2_{обр} =$  3,841459

$\chi^2_{обр} =$  6,437752

Так как  $F1$  не входит в крит. область,  $F2$  не входит в крит. область (каждый в свою), значит использование дисперсионного анализа обосновано.

ИТОГИ	B1	B2	Итого
<b>A1</b>			
Счет	3	3	6
Сумма	1900000	1810000	3710000
Среднее	633333,3333	603333,3333	618333,3333
Дисперсия	83333333,3	23333333,3	69666666,7
<b>A2</b>			
Счет	3	3	6
Сумма	1850000	1740000	3590000
Среднее	616666,6667	580000	598333,3333
Дисперсия	23333333,3	0	49666666,7
<b>A3</b>			
Счет	3	3	6
Сумма	1711000	1530000	3241000
Среднее	570333,3333	510000	540166,6667
Дисперсия	745033333,3	100000000	411216666,7
<b>Итого</b>			
Счет	9	9	
Сумма	5461000	5080000	
Среднее	606777,7778	564444,4444	
Дисперсия	292844444,4	185277777,8	

$q = 0.9411764705882353$   
 $s^2 = 2,39E+09$   
 $F = 0.3910826641108622$   
 $q = 0.9183673469387754$   
 $s^2 = 1768500000.1333334$   
 $F = 6.234472291748413$

## 1). Данные занесены след. образом

	B1	B2
A1	650000	620000
	600000	600000
	650000	590000
A2	620000	580000
	630000	580000
	600000	580000
A3	521000	510000
	520000	520000
	670000	500000

## 2). Мои рез-ты

Дисперсионный анализ

Источник вариации	SS	df	MS	F	P-Знач
Выборка	19786777778	2	9893388889	6,70713	0,01108
Столбцы	8064500000	1	8064500000	5,467252	0,0375
Взаимодействие	762333333,3	2	381166666,7	0,258408	0,77
Внутри	17700666667	12	1475055556		
Итого	46314277778	17			

## 3). Выводы

H0: средний кредит не зависит от профессии

H1: средний кредит зависит от профессии

**P-значение = 0,01108212 < 0,02** следовательно принимаем H1 с вер. 98%

средний кредит зависит от профессии 98%

H0: средняя оценка не зависит от пола

H1: средняя оценка зависит от пола

**P-значение** = 0,037507 < 0,05 следовательно принимаем H1 с вер.

95%- средняя оценка зависит от направления подготовки на 95%

## 5) Оценка силы влияния

Вывод: сумма кредита на **42%** объясняется профессией, на **17%** объясняется направлением обучения . **1.6 %** взаимодействием этих факторов и **38%** идёт на неучтённые факторы

Профессия	0,427228	42,72285	%
Пол	0,174126	17,41256	%
Взаимодействие	0,01646	1,646001	%
Остальное	0,382186	38,2186	%

