



ГЛАВНАЯ / КАБИНЕТ /

## Краткий отчет

ПОЛУЧИТЬ ПОЛНЫЙ ОТЧЕТ

[ВЕРСИЯ ДЛЯ ПЕЧАТИ](#)
[ЭКСПОРТ](#)
[ИСТОРИЯ ОТЧЕТОВ](#)
[РУКОВОДСТВО](#)
[ВЫЙТИ В КАБИНЕТ](#)
[ЕЩЁ...](#)

### КУРСОВАЯ осн исправленная

ПРОВЕРЕНО: 04.09.2018 17:14:32

№	Доля в отчете	Доля в тексте	Источник	Актуальна на	Модуль поиска	Блоков в отчете	Блоков в тексте
[01]	0%	16,88%	Применение методов математическо...	раньше 2011	Модуль поиска Интернет	0	17
[02]	16,78%	16,88%	Применение методов математическо...	раньше 2011	Модуль поиска Интернет	16	17
[03]	9,2%	9,2%	Дисперсионный анализ при помощи ...	22 Ноя 2012	Модуль поиска Интернет	24	24

ЗАИМСТВОВАНИЯ

48,37%

ЦИТИРОВАНИЯ

0%

ОРИГИНАЛЬНОСТЬ

51,63%

ИСТОЧНИКОВ: 10

ЕЩЕ НАЙДЕНО

ИСТОЧНИКОВ: 7

ЗАИМСТВОВАНИЯ: 22,39%

## Содержание

<b>Введение</b>	3
<b>1 Теоретические аспекты использования дисперсионного анализа в управленческой практике</b>	5
1.1 Понятие и значение дисперсионного анализа	5
1.2 Методы дисперсионного анализа	7
<b>2 Анализ управленческих данных средствами дисперсионного анализа с использованием пакета MS Excel</b>	18
2.1 Инструменты пакета дисперсионного анализа в Microsoft Excel	18
2.2 Проект решения управленческой задачи средствами дисперсионного анализа в пакете Microsoft Excel	22
<b>Заключение</b>	27
<b>Библиографический список</b>	28

## Введение

Когда генеральные совокупности распределены нормально и имеют одинаковую, хотя и неизвестную, дисперсию; математические ожидания также неизвестны и могут быть различными, тогда требуется при заданном уровне значимости по выборочным средним проверить нулевую гипотезу о равенстве всех математических ожиданий. Другими словами, требуется установить, значимо или незначимо различаются выборочные средние. Казалось бы, для сравнения нескольких средних можно сравнить их попарно. Однако с возрастанием числа средних возрастает и наибольшее различие между ними: среднее новой выборки может оказаться больше наибольшего или меньше наименьшего из средних, полученных до нового опыта. По этой причине для сравнения нескольких средних пользуются другим методом, который основан на сравнении дисперсий и поэтому назван дисперсионным анализом.

Актуальность данной работы обусловлена тем, что в современном обществе методы дисперсионного анализа применяются повсеместно, принося существенную экономическую выгоду и предупреждая финансовые крахи. Они позволяют принимать разнообразные управленческие решения в условиях риска и неопределенности.

Цель данной работы заключается в исследовании возможностей использования дисперсионного анализа в управленческой практике с помощью табличного процессора Microsoft Excel.

Электронные таблицы Microsoft Excel имеют огромный набор возможностей. Однако обычно пользователь в своей работе применяет только их небольшую часть. Происходит это не только от незнания, но и от отсутствия в этом острой необходимости. И действительно, все функции, заложенные в Excel, можно условно разделить на те, знание которых просто необходимо знать, и те, изучать которые можно по мере надобности.

В соответствии с целью исследования были поставлены следующие задачи:

- исследовать теоретические аспекты использования дисперсионного анализа в управленческой практике;

- рассмотреть процесс анализа управленческих данных средствами дисперсионного анализа с использованием пакета MS Excel.

Объект исследования - дисперсионный анализ.

Предмет исследования – способы использования дисперсионного анализа в управленческой практике.

Основным методом исследования является изучение и анализ специальной литературы по исследуемому вопросу.

В работе были использованы труды следующих авторов: С. Бондаренко, В.И. Борисова, И.К. Васильева, С.В. Глушакова, А.Леоненкова и др.

# 1 Теоретические аспекты использования дисперсионного анализа в управленческой практике

## 1.1 Понятие и значение дисперсионного анализа

На практике часто возникает необходимость проверки существенности различия выборочных средних  $t$  совокупностей ( $t > 2$ ). Например, требуется оценить влияние различных свойств сырья на показатели качества продукции, плавок на механические свойства металла, количества вносимых удобрений на урожайность, и т.п.

Для эффективного решения такой задачи нужен новый подход, который реализуется в дисперсионном анализе.

Основной целью дисперсионного анализа является исследование значимости различия между средними. Если просто сравниваются средние, в двух выборках, дисперсионный анализ даст тот же результат, что, и обычный  $t$ -критерий. Для независимых выборок (если сравниваются две независимые группы объектов или наблюдений) или  $t$ -критерий. Для зависимых выборок (если сравниваются две переменные на одном. И том же множестве объектов или наблюдений) [7, С. 72].

Фундаментальная концепция дисперсионного анализа предложена английским математиком - статистиком Р.А. Фишером. Для обработки результатов агрономических опытов по выявлению условий получения максимального урожая различных сортов сельскохозяйственных культур в 1918 году. Возможно, более естественным был бы термин анализ суммы квадратов или анализ вариации. Но, в силу традиции употребляется термин дисперсионный анализ. Сам термин «дисперсионный анализ» Фишер употребил позднее [7, С. 73]. По числу факторов, влияние которых исследуется, различают однофакторный. И многофакторный дисперсионный анализ.

Фактор (factors) - категориальная независимая переменная. Чтобы использовать дисперсионный анализ, независимые переменные должны все

быть категориальными (неметрическими). Конкретная комбинация уровней факторов называется факторным экспериментом (условиями испытаний) (treatment).

Факторный эксперимент (условия испытаний) (treatment). В дисперсионном анализе конкретная комбинация категорий (уровней) факторов.

Однофакторный дисперсионный анализ (one-way analysis of variance) – метод дисперсионного анализа, при котором используется только один фактор. Однофакторный дисперсионный анализ включает только одну категориальную переменную или единственный фактор. Если существует два или больше факторов, то анализ называют многофакторным дисперсионным анализом. Многофакторный дисперсионный анализ (n-way analysis of variance) – модель дисперсионного анализа. Которая включает два или больше факторов. Взаимосвязь дисперсионного анализа. С t-критерием, другими методами анализа, такими как регрессионный анализ, показана на рис. 1.1 [13, С. 80].



Рисунок 1.1 - Взаимосвязь между t-критерием, дисперсионным и ковариационным анализом и регрессией

Во всех представленных методах анализа (рис. 1.1) используется метрическая зависимая переменная. Дисперсионный и ковариационный анализ может включать несколько независимых переменных (степень использования продукта, лояльность к торговой марке, отношение, важность). Более того, одна из независимых переменных должна быть категориальной и категориальные переменные могут иметь больше двух уровней. С другой стороны, t-критерий предназначен для использования в случае с единственной бинарной независимой переменной. Например, различие в предпочтениях товара у лояльных и нелояльных респондентов можно узнать, выполнив проверку с помощью t-критерия. Регрессионный анализ, подобный дисперсионному и ковариационному, также может включать несколько независимых переменных. Однако все независимые переменные, в основном, измеряются интервальной шкалой, хотя бинарные или категориальные переменные могут приспособиваться к анализу за счет введения фиктивных (dummy) переменных.

Итак, в настоящее время дисперсионный анализ определяется как статистический метод, предназначенный для оценки влияния различных факторов на результат эксперимента, а также для последующего планирования аналогичных экспериментов.

## **1.2 Методы дисперсионного анализа**

При проведении экономического анализа часто необходимо оценить влияние на целевую функцию  $y$  качественного фактора  $x$ . Таким фактором могут быть, например, партии сырья, отрасли промышленности, регионы и т.д.

На современном этапе выделяются следующие основные методы дисперсионного анализа

1. Метод по Фишеру (Fisher) - критерий F. Данный метод применяется в однофакторном дисперсионном анализе, когда совокупная дисперсия всех

наблюдаемых значений раскладывается на дисперсию внутри отдельных групп и дисперсию между группами.

2. Метод «общей линейной модели», в его основе лежит корреляционный или регрессионный анализ, применяемый в многофакторном анализе.

Однофакторная дисперсионная модель имеет вид:

$$x_{ij} = \mu + F_j + \varepsilon_{ij}, \quad (1.1)$$

где  $x_{ij}$  – значение исследуемой переменной, полученной на  $i$ -м уровне фактора ( $i=1,2,\dots,t$ ) с  $j$ -м порядковым номером ( $j=1,2,\dots,n$ );

$F_i$  – эффект, обусловленный влиянием  $i$ -го уровня фактора;

$\varepsilon_{ij}$  – случайная компонента, или возмущение, вызванное влиянием неконтролируемых факторов, т.е. вариацией переменной внутри отдельного уровня.

Простейшим случаем дисперсионного анализа является одномерный однофакторный анализ для двух или нескольких независимых групп, когда все группы объединены по одному признаку.

В ходе анализа проверяется нулевая гипотеза о равенстве средних. При анализе двух групп дисперсионный анализ тождественен двухвыборочному  $t$ -критерию Стьюдента для независимых выборок, и величина  $F$ -статистики равна квадрату соответствующей  $t$ -статистики.

Для подтверждения положения о равенстве дисперсий обычно применяется критерий Ливена (*Levene's test*). В случае отвержения гипотезы о равенстве дисперсий основной анализ неприменим. Если дисперсии равны, то для оценки соотношения межгрупповой и внутригрупповой изменчивости применяется  $F$ -критерий Фишера. Если  $F$ -статистика превышает критическое значение, то нулевая гипотеза отвергается и делается вывод о неравенстве средних. При анализе средних двух групп результаты могут быть интерпретированы непосредственно после применения критерия Фишера.



Пусть данные о влиянии некоторого качественного фактора на количественный в форме таблицы 1.1 [13, С. 81].

Таблица 1.1 - Влияние качественного фактора на исследуемый показатель

$x_1$	$x_2$	...	$x_u$
$y_{11}$	$y_{12}$	....	$y_{1u}$
$y_{21}$	$y_{22}$	...	$y_{2u}$
...	...	...	...
$y_{m1}$	$y_{m2}$		$y_{mu}$

Модель зависимости значений  $y_{ij}$  от фактора столбцов можно представить в следующем виде:

$$y_{jk} = \mu + \alpha_j + \varepsilon_{jk}, (j=1, 2, \dots, u; k=1, 2, \dots, m_j) \quad (1.1)$$

где  $\mu$  - общее среднее,

$\alpha_j$  - отклонение от общего среднего для j-го уровня фактора,

$\varepsilon_{jk}$  - случайная составляющая.

По выборочным данным можно вычислить:

1) среднее  $\bar{y}_j$  для каждого уровня фактора (среднее по столбцам)  $x_j$  ( $j=1, 2, \dots, u$ ), по  $m_j$  параллельным опытам, где  $m_j$  – число данных в столбце j:

$$\bar{y}_j = \frac{1}{m_j} \sum_{k=1}^{m_j} y_{jk} \quad (1.2)$$

2) общее среднее  $\bar{y}$  по всем N опытам, т.е. по всем  $m_j$  параллельным

опытам на всех уровнях фактора  $x_j$  ( $N = \sum_{j=1}^u m_j$ ):

$$\bar{y} = \frac{1}{N} \sum_{j=1}^u \sum_{k=1}^{m_j} y_{jk}; \quad (1.3)$$

3) общую сумму квадратов отклонений  $Q_0$ :

$$Q_0 = \sum_{j=1}^u \sum_{k=1}^{m_j} (y_{jk} - \bar{y})^2; \quad (1.4)$$

4) сумму квадратов, характеризующую влияние фактора  $x$  (отклонения между группами)

$$Q_x = \sum_{j=1}^u (\bar{y}_j - \bar{y})^2 \cdot m_j; \quad (1.5)$$

5) остаточную сумму квадратов, зависящую от ошибки  $\varepsilon$  (отклонения внутри групп)

$$Q_\varepsilon = \sum_{j=1}^u \sum_{k=1}^m (y_{jk} - \bar{y}_j)^2 \quad (1.6)$$

Тождество дисперсионного анализа имеет вид:

$$Q_0 = Q_x + Q_\varepsilon. \quad (1.7)$$

На основании вычисленных сумм квадратов вычисляются:

1) оценка дисперсии относительно общего среднего  $\bar{Y}$ :

$$s^2_0\{y\} = \frac{1}{N-1} \cdot Q_0 \quad (1.8)$$

где  $f_0 = N - 1$  - число степеней свободы;

2) оценка дисперсии «между группами», определяемыми уровнями  $x_j$ :

$$s^2_x\{y\} = \frac{1}{u-1} \cdot Q_x, \quad (1.9)$$

где число степеней свободы  $f_x = u - 1$ .

3) выборочная оценка дисперсии «внутри групп», вычисляемая как средняя оценка по всем  $u$  группам:

$$s^2_\varepsilon\{y\} = \frac{1}{N-u} \cdot Q_\varepsilon, \quad (1.10)$$

с числом степеней свободы  $f_\varepsilon = N - u$ .

Числа степеней свободы должны удовлетворять соотношению

$$f_0 = f_{x_1} + f_{x_2} + f_{x_3} + f_\varepsilon \quad (1.11)$$

Для того, чтобы сделать вывод о том, влияет ли на исследуемые показатели качественный фактор, сопоставляют дисперсию между группами с общей дисперсией. При этом выдвигают следующие гипотезы [18, С. 315]:

$H_0$ :  $\bar{y}_1 = \bar{y}_2 = \bar{y}_3 = \dots = \bar{y}_u = \bar{y}$ , т.е. средние значения по всем столбцам равны и равны общему среднему, откуда следует, что среднеквадратическое отклонение по факторам равно среднеквадратическому отклонению по всем данным и

равно нулю. Т.е. качественный фактор не оказывает влияния на исследуемый показатель.

$H_1: \bar{y}_1 \neq \bar{y}_2 \neq \bar{y}_3 \neq \dots \neq \bar{y}_u \neq \bar{y}$  , , т.е. средние значения по всем столбцам не равны между собой и не равны общему среднему, откуда следует, что среднеквадратическое отклонение по факторам не совпадает со среднеквадратическим отклонением по всем данным. Т.е. качественный фактор оказывает существенное влияние на исследуемый показатель.

Оценивание значимости влияния фактора  $x$  выполняется по F-критерию Фишера, для чего формируется следующее F-отношение:

$$F = \frac{s^2_x\{y\}}{s^2_\varepsilon\{y\}}. \quad (1.12)$$

Фактор  $x$  признается незначимым, если соответствующее F-отношение оказывается меньше критического, выбранного из таблиц для принятого уровня значимости  $\alpha$  и числа степеней свободы сравниваемых дисперсий  $f_x$  и  $f_\varepsilon$ .

Табличное значение критерия Фишера определяется для числа степеней свободы  $u-1$  и  $N-1$  и вероятности ошибки  $\alpha$ .

Т.е. если  $|F| < F_{\alpha, \hat{u}\hat{u}}$ , то принимается нулевая гипотеза при соответствующем уровне значимости о том, что исследуемый фактор не оказывает существенного влияния на количественные данные.

Если  $|F| > F_{\alpha, \hat{u}\hat{u}}$ , то нулевая гипотеза отвергается и принимается альтернативная при соответствующем уровне значимости. Исходя из этого, можно сделать вывод о том, что исследуемый фактор оказывает существенное влияние на количественные данные.

Результаты дисперсионного анализа сводятся в таблицу 1.2 [18, С. 316].

Таблица 1.2 - Однофакторный дисперсионный анализ

Источник изменчивости	Сумма квадратов отклонений	Число степеней свободы	Оценка дисперсии	F – отношение
Между группами	$Q_x = \sum_{j=1}^u (\bar{y}_j - \bar{y})^2 \cdot m_j$	$f_x = u - 1$	$s^2_x\{y\} = \frac{1}{u-1} \cdot Q_x$	$F = \frac{s^2_x\{y\}}{s^2_\varepsilon\{y\}}$
Внутри групп (ошибка $\varepsilon$ )	$Q_\varepsilon = \sum_{j=1}^u \sum_{k=1}^{m_j} (y_{jk} - \bar{y}_j)^2$	$f_\varepsilon = N - u$	$s^2_\varepsilon\{y\} = \frac{1}{N-u} \cdot Q_\varepsilon$	
Общая сумма	$Q_0 = \sum_{j=1}^u \sum_{k=1}^{m_j} (y_{jk} - \bar{y})^2$	$f_0 = N - 1$	$s^2_0\{y\} = \frac{1}{N-1} \cdot Q_0$	

$m_j$  - число данных в столбце,  $u$ - число столбцов,  $m$  – число строк.

Часто необходимо качественно оценить значимость или незначимость влияния на целевую функцию  $y$  двух одновременно действующих факторов  $x_1$  и  $x_2$ . Такими факторами могут быть, например, форма собственности предприятия  $x_1$  и вид экономической деятельности  $x_2$ .

Модель двухфакторного дисперсионного анализа имеет вид:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}, (i=1, 2, \dots, u_1; j=1, 2, \dots, u_2; k=1, 2, \dots, m), \quad (1.13)$$

где  $\mu$  - общее среднее,  $\alpha_i$  - отклонение от общего среднего для фактора  $x_1$ ,

$\beta_j$  - отклонение от общего среднего для фактора  $x_2$ ,

$\gamma_{ij}$  - отклонение от общего среднего для взаимодействия двух факторов,

$\varepsilon_{ijk}$  - случайная составляющая.

В этом случае общую сумму квадратов отклонений  $Q_0$  можно разбить на четыре суммы:

- 1)  $Q_{x_1}$ -по фактору  $x_1$ ,
- 2)  $Q_{x_2}$ -по фактору  $x_2$ ,
- 3)  $Q_\varepsilon$ -остаточную сумму квадратов, зависящую от ошибки  $\varepsilon$ ,
- 4)  $Q_{x_1 \times x_2}$ -зависящую от взаимодействия (произведения)  $x_1 \times x_2$  двух факторов.

В этом случае по выборочным значениям вычисляются:

- 1) среднее  $\bar{y}_i$  для каждого уровня фактора  $x_1$ :

$$\bar{y}_i = \frac{1}{u_2 \cdot m} \sum_{j=1}^{u_2} \sum_{k=1}^m y_{ijk} \quad (1.14)$$

2) среднее  $\bar{y}_j$  для каждого уровня фактора  $x_2$ :

$$\bar{y}_j = \frac{1}{u_1 \cdot m} \sum_{i=1}^{u_1} \sum_{k=1}^m y_{ijk} \quad (1.15)$$

3) общее среднее  $\bar{y}$  по всем  $N$  опытам, т.е. по всем  $m$  параллельным опытам на всех сочетаниях уровней факторов  $x_1$  и  $x_2$  ( $N = u_1 \cdot u_2 \cdot m$ ):

$$\bar{y} = \frac{1}{N} \sum_{i=1}^{u_1} \sum_{j=1}^{u_2} \sum_{k=1}^m y_{ijk} \quad (1.16)$$

4) среднее  $\bar{y}_{ij}$  по  $m$  параллельным опытам для каждого сочетания уровней факторов  $x_1$  и  $x_2$ :

$$\bar{y}_{ij} = \frac{1}{m} \sum_{k=1}^m y_{ijk} \quad (1.17)$$

В таблице 1.3 показаны данные полного факторного эксперимента с одинаковым числом наблюдений в ячейках [18, С. 320].

Таблица 1.3 - Данные эксперимента и расчёты средних при двухфакторном дисперсионном анализе

j =		1	2	...	$u_2$	$\bar{y}_i = \frac{1}{u_2 \cdot m} \sum_{j=1}^{u_2} \sum_{k=1}^m y_{ijk}$
i =	k	$y_{i_1 k}$	$y_{i_2 k}$	...	$y_{i u_2 k}$	
1	1		$y_{121}$			
	2		$y_{122}$			
	...		...			
	m		$y_{12m}$			
.	1					
	2					
	...					
	m					
$u_1$	1					
	2					
	...					
	m					
$\bar{y}_j = \frac{1}{u_1 \cdot m} \sum_{i=1}^{u_1} \sum_{k=1}^m y_{ijk}$						$\bar{y} = \frac{1}{N} \sum_{i=1}^{u_1} \sum_{j=1}^{u_2} \sum_{k=1}^m y_{ijk}$

В табл. 1.3  $\overline{y_{ij}}$  вычисляется по выделенной части столбца, содержащей  $m$  параллельных опытов.

Общая сумма квадратов отклонений  $Q_0$  рассчитывается по формуле:

$$Q_0 = \sum_{i=1}^{u_1} \sum_{j=1}^{u_2} \sum_{k=1}^m (y_{ijk} - \overline{y})^2 \quad (1.17)$$

Эту сумму можно разложить на 4 составляющие:

- 1) сумму, характеризующую влияние фактора  $x_1$ :

$$Q_{x_1} = u_2 \cdot m \sum_{i=1}^{u_1} (\overline{y}_i - \overline{y})^2 \quad (1.18)$$

- 2) сумму, характеризующую влияние фактора  $x_2$ :

$$Q_{x_2} = u_1 \cdot m \sum_{j=1}^{u_2} (\overline{y}_j - \overline{y})^2 \quad (1.19)$$

- 3) сумму, характеризующую результат влияния взаимодействия  $x_1 x_2$ :

$$Q_{x_1 x_2} = m \sum_{i=1}^{u_1} \sum_{j=1}^{u_2} (\overline{y}_{ij} - \overline{y}_i - \overline{y}_j + \overline{y})^2 \quad (1.20)$$

- 4) сумму, характеризующую влияние ошибки  $\varepsilon$ :

$$Q_\varepsilon = \sum_{i=1}^{u_1} \sum_{j=1}^{u_2} \sum_{k=1}^m (y_{ijk} - \overline{y}_{ij})^2 \quad (1.21)$$

Указанные пять сумм, поделенные на соответствующее число степеней свободы, дают пять различных оценок дисперсии, если влияние факторов  $x_1$  и  $x_2$  незначимо. Для проведения дисперсионного анализа вычисляются следующие дисперсии:

- 1) оценка дисперсии относительно общего среднего  $\overline{y}$ :

$$s^2_0\{y\} = \frac{1}{N-1} \cdot Q_0 \quad (1.22)$$

где  $N = u_1 \cdot u_2 \cdot m$  - общее число наблюдений, а число степеней свободы

$$f_0 = N - 1 = u_1 \cdot u_2 \cdot m - 1 \quad (1.23)$$

2) оценка дисперсии «между строками», определяемыми уровнями  $x_{1j}$ :

$$s^2_{x_1}\{y\} = \frac{1}{u_1 - 1} \cdot Q_{x_1} \quad (1.24)$$

где  $f_{x_1} = u_1 - 1$  - число степеней свободы.

3) оценка дисперсии «между столбцами», соответствующими уровням фактора  $x_2$ :

$$s^2_{x_2}\{y\} = \frac{1}{u_2 - 1} \cdot Q_{x_2} \quad (1.25)$$

где  $f_{x_2} = u_2 - 1$  - число степеней свободы;

4) оценка дисперсии «между сериями» по  $m$  параллельным опытам каждая

$$s^2_{x_1 x_2}\{y\} = \frac{1}{(u_1 - 1) \cdot (u_2 - 1)} \cdot Q_{x_1 x_2} \quad (1.26)$$

с числом степеней свободы  $f_{x_1 x_2} = (u_1 - 1) \cdot (u_2 - 1)$ ;

5) оценка дисперсии «внутри серий» по  $m$  параллельным опытам, вычисляемая как средняя оценка по всем  $u_1 u_2$  сериям:

$$s^2_{\varepsilon}\{y\} = \frac{1}{u_1 \cdot u_2 \cdot \sum_{j=1}^{u_1} \sum_{g=1}^{u_2} \frac{1}{m-1} \sum_{k=1}^m (v_{jgk} - \bar{v}_{jg})^2} = \frac{1}{u_1 \cdot u_2 \cdot (m-1)} \cdot Q_{\varepsilon} \quad (1.27)$$

с числом степеней свободы  $f_{\varepsilon} = u_1 \cdot u_2 \cdot (m - 1) = N - u_1 \cdot u_2$ .

Числа степеней свободы должны удовлетворять соотношению

$$f_0 = f_{x_1} + f_{x_2} + f_{x_1 x_2} + f_{\varepsilon} \quad (1.28)$$

Статистическое оценивание значимости влияния факторов  $x_1$ ,  $x_2$  и взаимодействия  $x_1 x_2$  выполняются по F-критерию Фишера, для чего формируются следующие F-отношения:

$$F_1 = \frac{s^2_{x_1}\{y\}}{s^2_{\varepsilon}\{y\}}, \quad F_2 = \frac{s^2_{x_2}\{y\}}{s^2_{\varepsilon}\{y\}}, \quad F_{1,2} = \frac{s^2_{x_1 x_2}\{y\}}{s^2_{\varepsilon}\{y\}} \quad (1.29)$$

Фактор  $x_1$  или  $x_2$ , или взаимодействие  $x_1 x_2$  признаются незначимым, если соответствующее F-отношение оказывается меньше критического, выбранного

из таблиц для принятого уровня значимости  $\alpha$  и числа степеней свободы сравниваемых дисперсий.

Для того, чтобы сделать вывод о том, влияют ли на исследуемые показатели качественные факторы, выдвигают следующие гипотезы:

$H_0$ :  $\bar{y}_1 = \bar{y}_2 = \bar{y}_3 = \dots = \bar{y}_u = \bar{y}$  , т.е средние значения по всем столбцам равны фактор столбца не оказывает влияния на исследуемый показатель.

$H_1$ :  $\bar{y}_1 \neq \bar{y}_2 \neq \bar{y}_3 \neq \dots \neq \bar{y}_u \neq \bar{y}$  , т.е средние значения по всем столбцам не равны фактор столбца оказывает существенное влияние на исследуемый показатель.

$H_0$ :  $\bar{y}_1 = \bar{y}_2 = \bar{y}_3 = \dots = \bar{y}_u = \bar{y}$  , т.е средние значения по всем строкам равны фактор строки не оказывает влияния на исследуемый показатель.

$H_1$ :  $\bar{y}_1 \neq \bar{y}_2 \neq \bar{y}_3 \neq \dots \neq \bar{y}_u \neq \bar{y}$  , , т.е средние значения по всем строкам не равны фактор строки оказывает существенное влияние на исследуемый показатель.

$H_0$ :  $Q_{x_1x_2} = m \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} (\bar{y}_{ij} - \bar{y}_i - \bar{y}_j + \bar{y})^2 = 0$  , т.е отклонение взаимодействия факторов равно нулю и взаимодействие не значимо.

$H_1$ :  $Q_{x_1x_2} = m \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} (\bar{y}_{ij} - \bar{y}_i - \bar{y}_j + \bar{y})^2 \neq 0$  , фактор взаимодействия значим..

Если  $|F| < F_{\alpha \hat{a}\hat{a}}$  , то принимается нулевая гипотеза при соответствующем уровне значимости о том, что исследуемый фактор не оказывает существенного влияния на количественные данные.

Если  $|F| > F_{\alpha \hat{a}\hat{a}}$  , то нулевая гипотеза отвергается и принимается альтернативная при соответствующем уровне значимости. Исходя из этого, можно сделать вывод о том, что исследуемый фактор оказывает существенное влияние на количественные данные.

Результаты двухфакторного дисперсионного анализа представляются в виде таблице 1.4 [18, С. 321].



Таблица 1.4 - Двухфакторный дисперсионный анализ при равном числе наблюдений в ячейках

Вид изменчивости	Сумма квадратов отклонений	Число степеней свободы	Оценка дисперсии	F – отношение
От фактора x1	$Q_{x_1} = u_2 \cdot m \sum_{i=1}^{u_1} (\bar{y}_i - \bar{y})^2$	$u_1 - 1$	$s^2_{x_1} \{y\} = \frac{1}{u_1 - 1} \cdot Q_{x_1}$	$F_1 = \frac{s^2_{x_1} \{y\}}{s^2_{\varepsilon} \{y\}}$
От фактора x2	$Q_{x_2} = u_1 \cdot m \sum_{j=1}^{u_2} (\bar{y}_j - \bar{y})^2$	$u_2 - 1$	$s^2_{x_2} \{y\} = \frac{1}{u_2 - 1} \cdot Q_{x_2}$	$F_2 = \frac{s^2_{x_2} \{y\}}{s^2_{\varepsilon} \{y\}}$
От взаимодействия x1x2	$Q_{x_1x_2} = m \sum_{i=1}^{u_1} \sum_{j=1}^{u_2} (\bar{y}_{ij} - \bar{y}_i - \bar{y}_j + \bar{y})^2$	$(u_1 - 1) \cdot (u_2 - 1)$	$s^2_{x_1x_2} \{y\} = \frac{Q_{x_1x_2}}{(u_1 - 1) \cdot (u_2 - 1)}$	$F_{1,2} = \frac{s^2_{x_1x_2} \{y\}}{s^2_{\varepsilon} \{y\}}$
Остаточная (от ε)	$Q_{\varepsilon} = \sum_{i=1}^{u_1} \sum_{j=1}^{u_2} \sum_{k=1}^m (y_{ijk} - \bar{y}_{ij})^2$	$N - u_1 \cdot u_2$	$s^2_{\varepsilon} \{y\} = \frac{1}{u_1 \cdot u_2 \cdot (m - 1)}$	
Общая	$Q_0 = \sum_{i=1}^{u_1} \sum_{j=1}^{u_2} \sum_{k=1}^m (y_{ijk} - \bar{y})^2$	$N - 1$	$s^2_0 \{y\} = \frac{1}{N - 1} \cdot Q_0$	

где  $m$  – число данных в строке (число повторов в ячейке),  $u_2$  - число столбцов,  $u_1$  - число строк.

В заключении необходимо отметить, что существуют условия применения дисперсионного анализа:

1. Задачей исследования является определение силы влияния одного (до 3) факторов на результат или определение силы совместного влияния различных факторов (пол и возраст, физическая активность и питание и т.д.).

2. Изучаемые факторы должны быть независимые (несвязанные) между собой. Например, нельзя изучать совместное влияние стажа работы и возраста, роста и веса детей и т.д. на заболеваемость населения.

3. Подбор групп для исследования проводится рандомизированно (случайный отбор). Организация дисперсионного комплекса с выполнением принципа случайности отбора вариантов называется рандомизацией (перев. с англ. - random), т.е. выбранные наугад.

4. Можно применять как количественные, так и качественные (атрибутивные) признаки.

При проведении однофакторного дисперсионного анализа рекомендуется (необходимое условие применения):

1. Нормальность распределения анализируемых групп или соответствие выборочных групп генеральным совокупностям с нормальным распределением.
2. Независимость (не связанность) распределения наблюдений в группах.
3. Наличие частоты (повторность) наблюдений.

Нормальность распределения определяется кривой Гаусса (Де Мавура), которую можно описать функцией  $y=f(x)$ , так как она относится к числу законов распределения, используемых для приближенного описания явлений, которые носят случайный, вероятностный характер. Предмет медико-биологических исследований - явления вероятностного характера, нормальное распределение в таких исследованиях встречается весьма часто.

## **2 Анализ управленческих данных средствами дисперсионного анализа с использованием пакета MS Excel**

### **2.1 Инструменты пакета дисперсионного анализа в Microsoft Excel**

При решении реальных задач методом дисперсионного анализа используются различные статистические программные пакеты.

Пакет анализа включает в себя три средства дисперсионного анализа. Выбор конкретного инструмента определяется числом факторов и числом выборок в исследуемой совокупности данных.

В рамках данной проведен анализ данных средствами дисперсионного анализа с использованием пакета MS Excel.

В системе электронных таблиц Microsoft Excel имеется набор инструментов для анализа данных, называемый пакет анализа, который может

быть использован для решения сложных статистических задач. Для использования одного из этих инструментов указать входные данные и выбрать параметры; анализ будет проведен с помощью подходящей статистической макрофункции, и результаты будут представлены в выходном диапазоне.

В меню Сервис выбрать команду Анализ данных. Если такая команда отсутствует в меню Сервис, то необходимо установить в Microsoft Excel пакет анализа данных [16, С. 105].

Установка производится следующим образом. В меню Сервис выберите команду Надстройки. Если в списке надстроек нет пакета анализа данных, то нажмите кнопку “Обзор” и задайте диск, каталог и имя файла для надстройки “Пакет анализа”, или запустите программу установки Microsoft Excel. Установите флажок “Пакет анализа” (надстройки, установленные в Microsoft Excel, остаются доступными, пока не будут удалены).

Выберите необходимую строку в списке “Инструменты анализа” (2.1).

Введите входной и выходной диапазоны, затем выберите необходимые параметры. Для использования инструментов анализа исследуемые данные следует представить в виде строк или столбцов на листе. Совокупность ячеек, содержащих анализируемые данные, называется входным диапазоном.

В MS Excel для проведения однофакторного дисперсионного анализа используется процедура «Однофакторный дисперсионный анализ» (рис. 2.2) [16, С. 106].

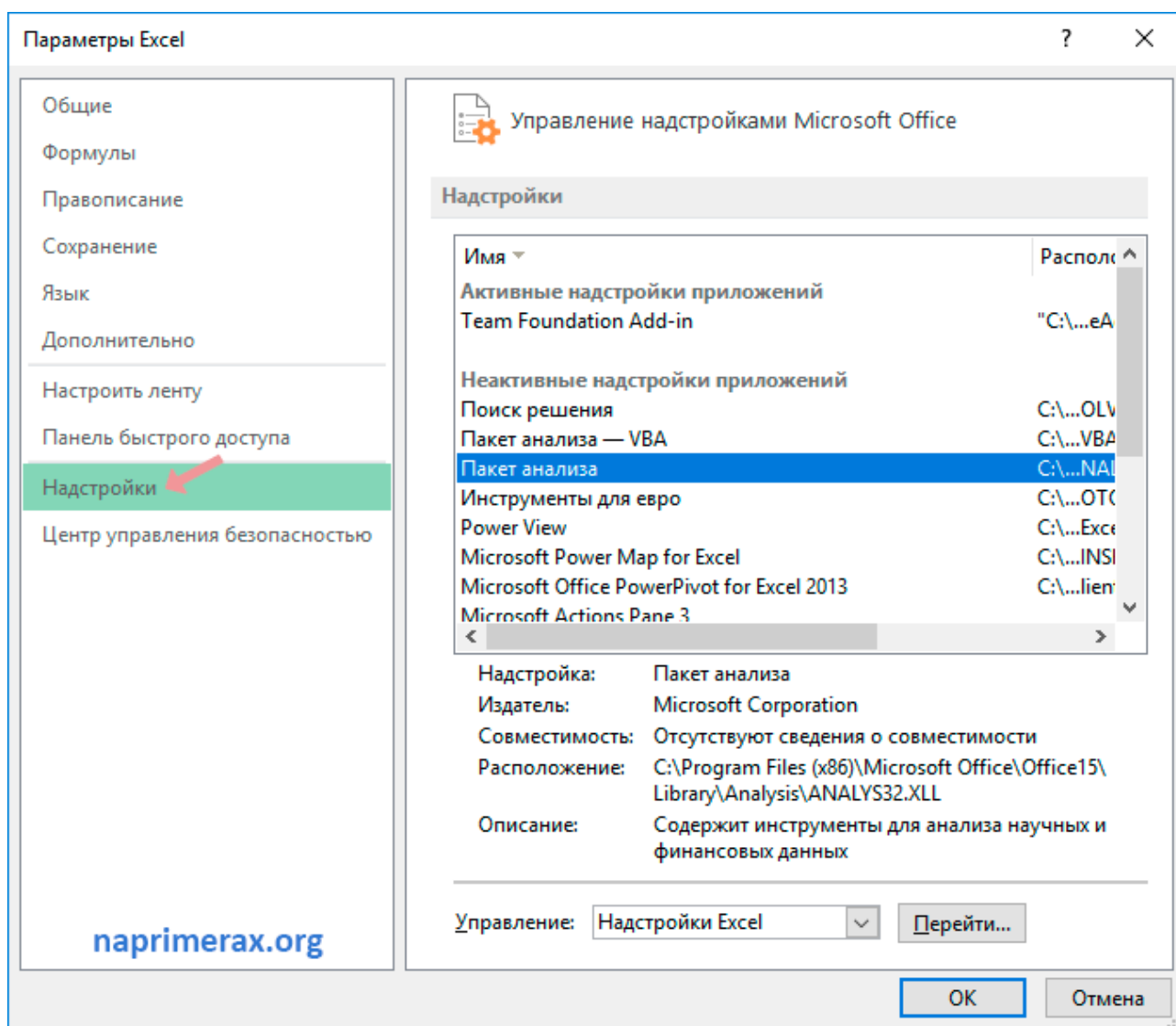


Рисунок 2.1 – Выбор инструмента анализа

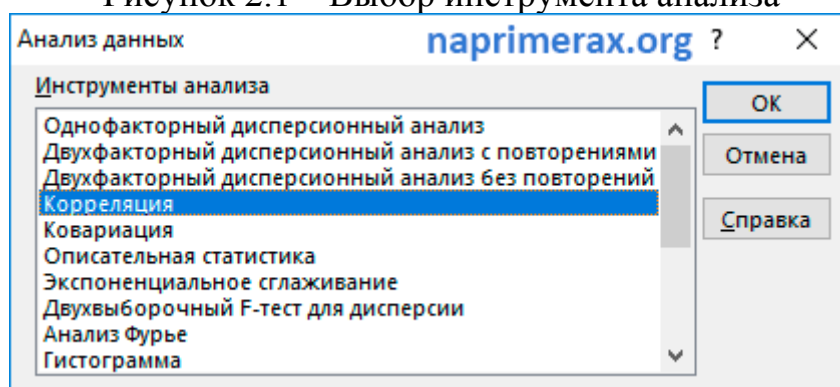


Рисунок 2.2 – Выбор диалогового окна дисперсионного анализа

Для проведения дисперсионного анализа необходимо:

- ввести данные в таблицу, так чтобы в каждом столбце оказались данные, соответствующие одному значению исследуемого фактора, а столбцы располагались в порядке возрастания (убывания) величины исследуемого фактора,

- выполнить команду Сервис > Анализ данных;
- в появившемся диалоговом окне Анализ данных в списке Инструменты анализа выбрать процедуру Однофакторный дисперсионный анализ (рис. 2.2), указав курсором мыши и щелкнув левой кнопкой мыши. Затем нажать кнопку ОК;
- в появившемся диалоговом окне задать Входной интервал, то есть ввести ссылку на диапазон анализируемых данных, содержащий все столбцы данных. Для этого следует навести указатель мыши на верхнюю левую ячейку диапазона данных, нажать левую кнопку мыши и, не отпуская ее, протянуть указатель мыши к нижней правой ячейке, содержащей анализируемые данные, затем отпустить левую кнопку мыши;
- в разделе Группировка переключатель установить в положение по столбцам;
- указать выходной диапазон, то есть ввести ссылку на ячейки, в которые будут выведены результаты анализа. Для этого следует поставить переключатель в положение Выходной интервал (навести указатель мыши и щелкнуть левой кнопкой), далее навести указатель мыши на правое поле ввода Выходной интервал и щелкнуть левой кнопкой мыши, затем указатель мыши навести на левую верхнюю ячейку выходного диапазона и щелкнуть левой кнопкой мыши. Размер выходного диапазона будет определен автоматически, и на экран будет выведено сообщение в случае возможного наложения выходного диапазона на исходные данные.
- нажать кнопку ОК [16, С. 107]. Результаты анализа. Выходной диапазон будет включать в себя результаты дисперсионного анализа: средние, дисперсии, критерий Фишера и другие показатели.

Интерпретация результатов. Влияние исследуемого фактора определяется по величине значимости критерия Фишера, которая находится в таблице Дисперсионный анализ на пересечении строки Между группами и столбца Р-Значение. В случаях, когда Р-Значение < 0,05, критерий Фишера значим, и влияние исследуемого фактора можно считать доказанным.

В пакете анализа программы MS Excel не существует инструмента многофакторного дисперсионного анализа с повторениями или без повторений. Но зато в рамках многофакторного регрессионного анализа (инструмент анализа «Регрессия») для оценки достоверности получаемой регрессионной модели производится многофакторный дисперсионный анализ.

## 2.2 Проект решения управленческой задачи средствами дисперсионного анализа в пакете Microsoft Excel

В рамках данной работы исследован пример решения задачи средствами факторного дисперсионного анализа в пакете Microsoft Excel.

Условие задачи: у условного предприятия были известны продажи группы товаров за последние 4 месяца (рис. 2.3). Поставлена задача провести анализ спроса на товары.

	A	B	C	D	E	F
1	Наименование	1 мес.	2 мес.	3 мес.	4 мес	Итого
2	Товар 1	310	421	513	646	1890
3	Товар 2	183	235	272	317	1007
4	Товар 3	136	110	428	319	993
5	Товар 4	112	231	349	473	1165
6	Товар 5	57	175	234	400	866
7	Товар 6	68	179	305	425	977
8	Товар 7	38	10	87	133	268
9	Товар 8	47	85	132	143	407
10	Итого:	951	1446	2320	2856	7573

Рисунок 2.3 – Условие задачи

В пакете Microsoft Excel исследовано, за счет, каких наименований произошел основной рост по итогам второго месяца. Если продажи какого-то товара выросли, положительная дельта – в столбец «Рост». Отрицательная – «Снижение». Формула в Excel для «роста»: =ЕСЛИ((C2-B2)>0;C2-B2;0), где C2-B2 – разница между 2 и 1 месяцем. Формула для «снижения»: =ЕСЛИ(J3=0;B2-C2;0), где J3 – ссылка на ячейку слева («Рост»). Во втором

столбце – сумма предыдущего значения и предыдущего роста за вычетом текущего снижения (рисунок 2.4).

Н	И	Ж	К
Факторы		Рост	Снижение
Итог 1 мес.	951	0	0
Товар 1	951	111	0
Товар 2	1062	52	0
Товар 3	1088	0	26
Товар 4	1088	119	0
Товар 5	1207	118	0
Товар 6	1325	111	0
Товар 7	1408	0	28
Товар 8	1408	38	0
Итог 2 мес.	1446		
	=I3+J3-K4		

Рисунок 2.4 – Введение формулы динамики спроса на товары в Microsoft Excel

Далее рассчитан процент роста по каждому наименованию товара. Формула: =ЕСЛИ(J3/\$I\$11=0;-K3/\$I\$11;J3/\$I\$11). Где J3/\$I\$11 – отношение «роста» к итогу за 2 месяц, -K3/\$I\$11 – отношение «снижения» к итогу за 2 месяц.

Г	Н	И	Ж	К
% роста	Факторы		Рост	Снижение
	Итог 1 мес.	951	0	0
8%	Товар 1	951	111	0
4%	Товар 2	1062	52	0
-2%	Товар 3	1088	0	26
8%	Товар 4	1088	119	0
8%	Товар 5	1207	118	0
8%	Товар 6	1325	111	0
-2%	Товар 7	1408	0	28
3%	Товар 8	1408	38	0
	Итог 2 мес.	1446		

Рисунок 2.5 – Расчет факторов роста (снижения) спроса

Далее выделена область данных для построения диаграммы, на вкладке «Вставка» - «Гистограмма».

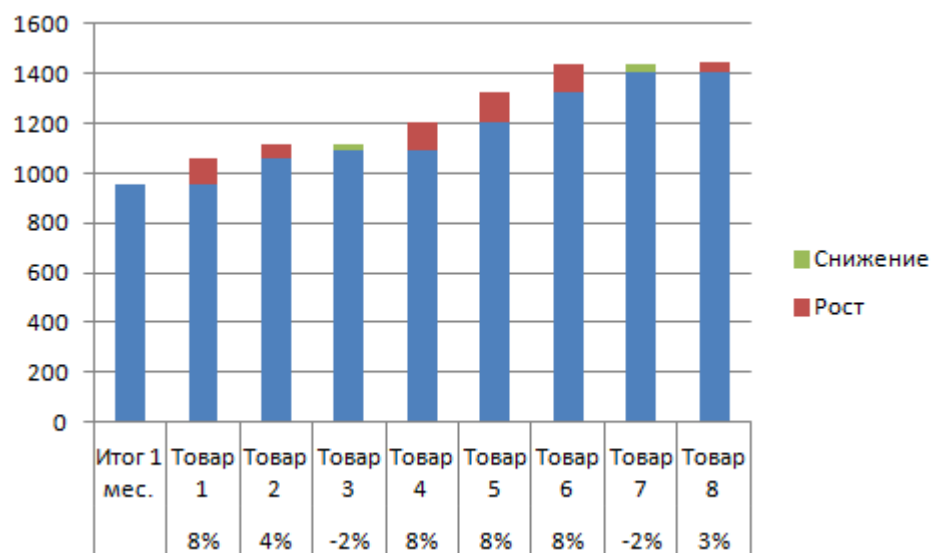


Рисунок 2.6 – Построение диаграммы динамики спроса на товары

Далее через «Формат ряда данных» - «Заливка» («Нет заливки»), изменен цвет для «снижения» и «роста» (таблица 2.7).

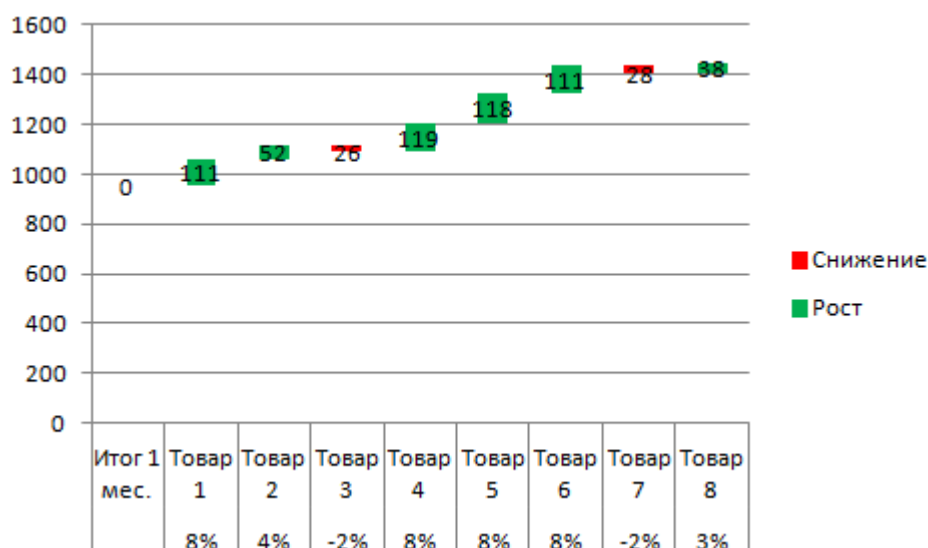


Таблица 2.7 – Итоговая диаграмма динамики спроса на товары

Теперь наглядно видно, продажи какого товара дают основной рост. Проект решения задачи представлен в Microsoft Excel - Приложение 1.

Далее с помощью факторного дисперсионного анализа на примере условной организации приведен пример анализа стратегии поведения сотрудников в конфликтной ситуации. Специалист по кадрам предполагает, что на поведение влияет уровень образования (1 – среднее, 2 – среднее специальное, 3 – высшее). Данные для анализа в Excel представлен на рисунке 2.8.



	A	B	C	D	E	F	G
1	№	Образ.	Соперн.	Сотрудн.	Избег.	Присп.	Компром.
2	1	1	10	8	6	5	5
3	2	1	5	7	9	4	6
4	3	1	5	8	8	5	7
5	4	1	6	6	7	5	8
6	5	1	9	8	4	5	5
7	6	2	5	6	5	7	4
8	7	2	8	6	4	6	7
9	8	2	7	5	4	6	6
10	9	2	5	7	6	7	7
11	10	2	6	5	7	4	8
12	11	3	5	5	10	6	5
13	12	3	4	4	8	7	7
14	13	3	6	5	6	8	6
15	14	3	10	5	4	6	5
16	15	3	8	4	6	9	4

Рисунок 2.8 – Исходные данные для факторного дисперсионного анализа

Для решения задачи в Excel открывается диалоговое окно аналитического инструмента, в раскрывшемся списке выбрано «Однофакторный дисперсионный анализ» и ОК.

В поле «Входной интервал» введена ссылка на диапазон ячеек, содержащихся во всех столбцах таблицы (рисунок 2.9).

	A	B	C	D	E	F	G
1	№	Образ.	Соперн.	Сотрудн.	Избег.	Присп.	Компром.
2	1	1	10	8	6	5	5
3	2	1	5	7	9	4	6
4	3	1	5	8	8	5	7
5	4	1	6	6	7	5	8
6	5	1	9	8	4	5	5
7	6	2	5	6	5	7	4
8	7	2	8	6	4	6	7
9	8	2	7	5	4	6	6
10	9	2	5	7	6	7	7
11	10	2	6	5	7	4	8
12	11	3	5	5	10	6	5
13	12	3	4	4	8	7	7
14	13	3	6	5	6	8	6
15	14	3	10	5	4	6	5
16	15	3	8	4	6	9	4

Рисунок 2.9 – Определения интервала факторного дисперсионного анализа

«Группирование» назначено по столбцам. Результаты анализа выводятся на отдельный лист (рисунок 2.10).

Однофакторный дисперсионный анализ						
ИТОГИ						
Группы	Счет	Сумма	Среднее	Дисперсия		
Столбец 1	15	30	2	0,7142857		
Столбец 2	15	99	6,6	3,8285714		
Столбец 3	15	89	5,9333333	1,9238095		
Столбец 4	15	94	6,2666667	3,6380952		
Столбец 5	15	90	6	2		
Столбец 6	15	90	6	1,7142857		
Дисперсионный анализ						
источник вариации	SS	df	MS	F	P-Значение	F критическое
Между группами	220,9333	5	44,18667	19,185114	1,0955E-12	2,323126496
Внутри групп	193,4667	84	2,303175			
Итого	414,4	89				

Рисунок 2.10 – Результаты анализа

Значимый параметр залив желтым цветом. Так как P-Значение между группами больше 1, критерий Фишера нельзя считать значимым. Следовательно, в исследуемой ситуации был сделан вывод, что поведение сотрудников условной организации в конфликтной ситуации не зависит от уровня образования.

Проект решения задачи представлен в Microsoft Excel - Приложение 2.

## Заключение

По результатам исследования проведенного в данной работе можно сделать следующие основные выводы:

Дисперсионный анализ используется в управленческой практике для изучения различий средних значений зависимых переменных, вызванных влиянием контролируемых независимых переменных, при условии, что учтено влияние неконтролируемых независимых переменных. По сути, дисперсионный анализ применяют как проверку статистической значимости различий выборочных средних для двух или больше совокупностей. Обычно нулевая гипотеза утверждает, что все выборочные средние равны. Проверку нулевой гипотезы можно выполнить, используя дисперсионный анализ.

Дисперсионный анализ (analysis of variance - ANOVA) - статистический метод изучения различий между выборочными средними для двух или больше совокупностей.

В своей простейшей форме дисперсионный анализ должен иметь зависимую переменную, которая является метрической. Кроме того, должна быть одна или больше независимых переменных. Все независимые переменные должны быть категориальными (неметрическими), их еще называют факторами (factors).

Условно цель дисперсионного метода можно сформулировать так: вычленив из общей вариативности параметра 3 частные вариативности:

- 1 – определенную действием каждого из изучаемых значений;
- 2 – продиктованную взаимосвязью между исследуемыми значениями;
- 3 – случайную, продиктованную всеми неучтенными обстоятельствами.

В программе Microsoft Excel дисперсионный анализ можно выполнить с помощью инструмента «Анализ данных» (вкладка «Данные» - «Анализ»). Это надстройка табличного процессора. В рамках данной работы были

представлены варианты решения управленческих задач средствами дисперсионного анализа в пакете Microsoft Excel.

Несмотря на то, что в пакете анализа программы MS Excel не существует инструмента многофакторного дисперсионного анализа с повторениями или без повторений. Но зато в рамках многофакторного регрессионного анализа (инструмент анализа «Регрессия») для оценки достоверности получаемой регрессионной модели производится многофакторный дисперсионный анализ.

Итак, MS Excel позволяет подбирать оптимальное решение к определенным задачам, сокращая при этом время и делая результат более точным при сложных вычислениях, так же табличный процессор помогает выполнять функцию анализа более точно и содержит ряд необходимых функций для работы с дисперсионным анализом.

## Библиографический список

1. Банди, Б. Методы оптимизации. Вводный курс / Б. Банди. - М.: ИНФРА, 2015. – 325 с.
2. Бондаренко, С. Популярный самоучитель Excel / Бондаренко С. – СПб, 2016. - 320 с
3. Борисов, В.И. Основы автоматизации / В.И. Борисов. – М.: Наука, 2015. - 318 с.
4. Бугорский, В.Н. Экономика и проектирование информационных систем / В.Н. Бугорский. - СПб.: Питер, 2016. -320с.
5. Васильев, И.К. Автоматизированные системы управления / И.К. Васильев. – М.: Бином, 2015. – 305 с.
6. Глушаков, С.В. Microsoft Office. Лучший самоучитель / С.В. Глушаков, А.С. - М.: АСТ МОСКВА; Владимир: ВКТ, 2015. - 446. с.
7. Дулькейт, В.И. Решение управленческих задач дисперсионным анализом / В.М. Дулькейт // Вычислительные методы в дискретной математике. - 2016. - № 1. - С. 72-78.
8. Заславский, Ю.Л. Сборник задач по линейному программированию / Ю.Л. Заславский. - М.: ИНФРА, 2015. – 255 с.
9. Карпова, Т.С. Современные средства и этапы разработки и баз данных / Т.С.Карпова. - СПб.: Питер, 2015. - 304 с.
10. Когаловский, М.Р. Перспективные технологии информационных систем / М.Р. Когаловский. - М.: ДМК Пресс, 2015. - 288 с.
11. Когаловский, М.Р. Энциклопедия оптимизации в среде MS Excel / М.Р. Когаловский. - М.: Финансы и статистика, 2016. - 800 с.
12. Корнеев, В.В. Интеллектуальная обработка информации / В.В. Корнеев. - М.: ИНФРА, 2016. - 352 с.
13. Кузнецов, Ю.Н. Математическое программирование / Ю.Н. Кузнецов. -

- М.: ЮНИТИ, 2015. - 311 с.
14. Лавренко, С.М. Excel: Сборник примеров и задач / С.М. Лавренко. - М.: Финансы и статистика, 2014. – 336 с.
  15. Леоненков, А. Решение задач оптимизации в сфере MS / Леоненков А. – СПб, 2016. - 704 с.
  16. Сдвинков, О.А. Математика в MS Excel / О.А. Сдвинков. – М.: Солон-Пресс, 2016. – 295 с.
  17. Тиори, Т. Теория вероятностей и математическая статистика / Т.Тиори. Пер. с англ. – М.: Мир, 2015. - 384 с.
  18. Хансен, Г. Математическая статистика / Г.Хансен. - М.: Бином, 2014. - 704 с.
  19. Шапкин, А.С. Математические методы и модели исследования операций / А.С. Шапкин. - М.. Дашков и К, 2016. – 305 с.