

**Министерство образования и науки Российской Федерации  
ФГАОУ ВО «Севастопольский государственный университет»**

**Институт информационных технологий  
и управления в технических системах**

### **Лабораторная работа №3**

## **«Задача дисперсионного анализа. Методы дисперсионного анализа. Однофакторный дисперсионный анализ»**

по дисциплине «Интеллектуальный анализ данных»  
для студентов всех форм обучения направления подготовки  
09.03.02 «Информационные системы и технологии»



Севастополь  
2017

**Задача дисперсионного анализа. Методы дисперсионного анализа. Однофакторный дисперсионный анализ.** Методические указания к лабораторным занятиям по дисциплине «Интеллектуальный анализ данных» / Сост.: О.А. Сырых – Севастополь: Изд-во СевГУ, 2017 – 22 с.

Методические указания предназначены для проведения лабораторных работ по дисциплине «Интеллектуальный анализ данных». Целью методических указаний является помощь студентам в изучении возможностей системы RStudio. Излагаются практические сведения необходимые для выполнения лабораторной работы, требования к содержанию отчета.

Методические указания рассмотрены и утверждены на методическом семинаре и заседании кафедры «Информационные системы» (протокол № 1 от 29 августа 2016 г.)

### Лабораторная работа №3

#### Задача дисперсионного анализа. Методы дисперсионного анализа. Однофакторный дисперсионный анализ.

##### Цель:

- приобрести практические навыки в проведении дисперсионного анализа по экспериментальным данным
- исследовать возможности языка R для проведения дисперсионного анализа.

**Время:** 4 часа

**Лабораторное оборудование:** персональные компьютеры, выход в сеть Internet, RStudio.

#### Краткие теоретические сведения

Дисперсионный анализ разработан в 20-х годах XX века английским математиком и генетиком Рональдом Фишером для обработки результатов агрономических опытов по выявлению условий получения максимального урожая различных сортов сельскохозяйственных культур. Сам термин «дисперсионный анализ» Фишер употребил позднее.

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.

**Дисперсионный анализ** (от латинского Dispersio – рассеивание / на английском Analysis Of Variance – ANOVA) применяется для исследования влияния одной или нескольких качественных переменных (факторов) на одну зависимую количественную переменную (отклик).

В основе дисперсионного анализа лежит предположение о том, что одни переменные могут рассматриваться как причины (факторы, независимые переменные):  $f_1, \dots, f_k$ , а другие как следствия (зависимые переменные). Независимые переменные называют иногда регулируемыми факторами именно потому, что в эксперименте исследователь имеет возможность варьировать ими и анализировать получающийся результат.

Дисперсионный анализ особенно эффективен при изучении нескольких факторов. При классическом методе исследования варьируют только один фактор, а остальные оставляют постоянными. При этом для каждого фактора проводится серия наблюдений, не используемая при изучении других факторов. Кроме того, при таком методе исследования не удастся определить взаимодействие факторов при одновременном их изменении. При дисперсионном анализе каждое наблюдение служит для одновременной оценки всех факторов и их взаимодействий.

Дисперсионный анализ состоит в выделении и оценке отдельных факторов, вызывающих изменчивость изучаемой случайной величины. Для этого производится разложение суммарной выборочной дисперсии на составляющие, обусловленные независимыми факторами. Каждая из этих составляющих представляет собой оценку дисперсии генеральной совокупности. Чтобы решить, значимо ли влияние данного фактора, необходимо оценить значимость соответствующей выборочной дисперсии в сравнении с дисперсией воспроизводимости, обусловленной случайными факторами. Проверка значимости оценок дисперсий проводится по критерию Фишера. Если рассчитанное значение критерия Фишера окажется меньше табличного, то влияние рассматриваемого фактора нет оснований считать значимым. Если же рассчитанное значение критерия Фишера окажется больше табличного, то рассматриваемый фактор влияет на изменчивость средних.

В дальнейшем будем полагать, что выполняются следующие допущения:

- случайные ошибки наблюдений имеют нормальное распределение;
- факторы влияют только на изменение средних значений, а дисперсия наблюдений остается постоянной;

- эксперименты равноточны.

Требование нормального распределения определяет выбор основных факторов при исследовании процесса методом дисперсионного анализа. Если нужно получить нормальное распределение выходной величины, к случайным желательным относятся только те факторы, влияние которых на выходную величину очень мало. Исключение можно делать лишь для тех факторов, которые сами по себе (из каких-либо других соображений) дают нормальное распределение результатов.

Факторы рассматриваемые в дисперсионном анализе, бывают двух родов:

- со случайными уровнями;
- с фиксированными.

В первом случае предполагается, что выбор уровней производится из бесконечной совокупности возможных уровней и сопровождаются рандомизацией. При этом результаты эксперимента имеют большее значение, поскольку выводы по эксперименту можно распространить на всю генеральную совокупность. Если все уровни выбираются случайным образом, математическая модель эксперимента называется моделью со случайными уровнями факторов (случайная модель). Когда все уровни фиксированы, модель называется моделью с фиксированными уровнями. Когда часть факторов рассматривается на фиксированных уровнях, а уровни остальных выбираются случайным образом, модель называется моделью смешанного типа. Иногда отсутствие различия в критериях, применяемых для разных моделей, и единственное различие состоит в общности выводов, в других случаях существует различие в критериях.

Дисперсионный анализ может применяться в различных формах в зависимости от структуры исследуемого процесса; выбор соответствующей формы является обычно одной из главных трудностей в практическом применении анализа.

По числу факторов различают однофакторный и многофакторный дисперсионный анализ

### **Однофакторный дисперсионный анализ**

Задачей дисперсионного анализа является изучение влияния одного или нескольких факторов на рассматриваемый признак.

Однофакторный дисперсионный анализ используется в тех случаях, когда есть в распоряжении три или более независимые выборки, полученные из одной генеральной совокупности путем изменения какого-либо независимого фактора, для которого по каким-либо причинам нет количественных измерений.

Для этих выборок предполагают, что они имеют разные выборочные средние и одинаковые выборочные дисперсии. Поэтому необходимо ответить на вопрос, оказал ли этот фактор существенное влияние на разброс выборочных средних или разброс является следствием случайностей, вызванных небольшими объемами выборок. Другими словами, если выборки принадлежат одной и той же генеральной совокупности, то разброс данных между выборками (между группами) должен быть не больше, чем разброс данных внутри этих выборок (внутри групп).

Пусть  $x_{ik}$  –  $i$  – элемент ( $i = \overline{1, n_k}$ )  $k$  -выборки ( $k = \overline{1, m}$ ), где  $m$  – число выборок,  $n_k$  – число данных в  $k$  -выборке. Тогда  $\overline{x_{ik}}$  – выборочное среднее  $k$  -выборки определяется по формуле

$$\overline{x_k} = \frac{1}{n_k} \sum_{i=1}^{n_k} x_{ik}$$

Общее среднее вычисляется по формуле

$$\bar{x} = \frac{1}{n} \sum_{k=1}^m \sum_{i=1}^{n_k} x_{ik}$$

$$n = \sum_{k=1}^m n_k .$$

где

Основное тождество дисперсионного анализа имеет следующий вид:

$$Q = Q_1 + Q_2 ,$$

где  $Q_1$  – сумма квадратов отклонений выборочных средних  $\bar{x}_k$  от общего среднего  $\bar{x}$  (сумма квадратов отклонений между группами);

$Q_2$  – сумма квадратов отклонений наблюдаемых значений  $x_{ik}$  от выборочной средней  $\bar{x}_k$  (сумма квадратов отклонений внутри групп);

$Q$  – общая сумма квадратов отклонений наблюдаемых значений  $x_{ik}$  от общего среднего  $\bar{x}$ .

Расчет этих сумм квадратов отклонений осуществляется по следующим формулам:

$$Q_1 = \sum_{k=1}^m n_k (\bar{x}_k - \bar{x})^2 = \sum_{k=1}^m n_k \bar{x}_k^2 - n \bar{x}^2 ,$$

$$Q_2 = \sum_{k=1}^m \sum_{i=1}^{n_k} (x_{ik} - \bar{x}_k)^2 = \sum_{k=1}^m \sum_{i=1}^{n_k} x_{ik}^2 - \sum_{k=1}^m n_k \bar{x}_k^2 .$$

В качестве критерия необходимо воспользоваться критерием Фишера:

$$F = \frac{Q_1 / (m - 1)}{Q_2 / (n - m)} .$$

Если расчетное значение критерия Фишера будет меньше, чем табличное значение нет оснований считать, что независимый фактор оказывает влияние на разброс средних значений, в противном случае, независимый фактор оказывает существенное влияние на разброс средних значений ( $\alpha$ – уровень значимости, уровень риска, обычно для экономических задач  $\alpha=0,05$ ).

Недостаток однофакторного анализа: невозможно выделить те выборки, которые отличаются от других. Для этой цели необходимо использовать метод Шеффе или проводить парные сравнения выборок.

### Задание и порядок выполнения лабораторной работы №3

#### Дисперсионный анализ в MS EXCEL

1. Создать файл с исходными данными (варианты заданий представлены в Приложении А).

2. Запустить “Пакет анализа”.

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 макروفункции, и результаты будут представлены в выходном диапазоне.

В меню Сервис (Данные) выберите команду Анализ данных. Если такая команда отсутствует в меню Сервис (Данные), то необходимо установить в Microsoft Excel пакет анализа данных.

Установка производится следующим образом. В меню Сервис (Файл → Параметры) 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”.

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.

3. Провести однофакторный дисперсионный анализ.

В меню Сервис выбираем команду Анализ данных.

В списке инструментов статистического анализа выбираем Однофакторный дисперсионный анализ (Рисунок 1).

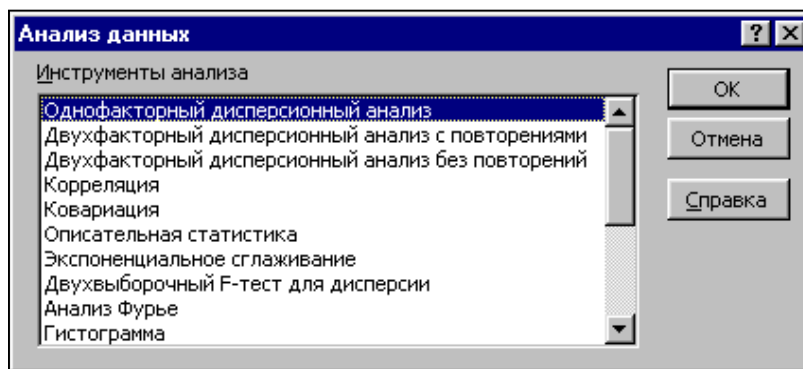


Рисунок 1 – Выбор инструмента анализа

В диалоговом окне режима (Рисунок 2) указываем входной интервал, способ группирования, выходной интервал, метки в первой строке/ Метки в первом столбце, альфа (уровень значимости).

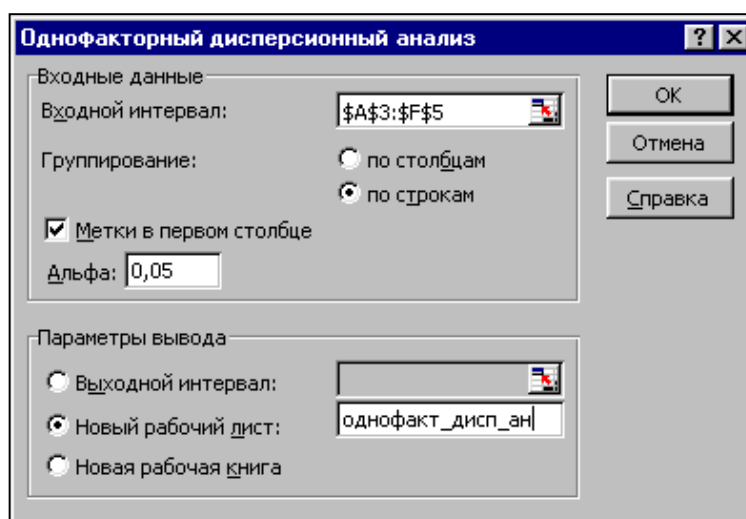


Рисунок 2 – Диалоговое окно однофакторного дисперсионного анализа

Входной диапазон – это ссылка на ячейки, содержащие анализируемые данные. Ссылка должна состоять как минимум из двух смежных диапазонов данных, организованных в виде столбцов или строк. Входной интервал можно задать при помощи мыши, или набрать на клавиатуре.

Группирование. Установите переключатель в положение “по столбцам” или “по строкам” в зависимости от расположения данных во входном диапазоне.

Метки в первой строке/ Метки в первом столбце. Установите переключатель в положение “Метки в первой строке”, если первая строка во входном диапазоне содержит названия столбцов. Установите переключатель в положение “Метки в первом столбце”, если названия строк находятся в первом столбце входного диапазона. Если входной диапазон не содержит меток, то необходимые заголовки в выходном диапазоне будут созданы автоматически.

Выходной диапазон. Введите ссылку на ячейку, расположенную в левом верхнем углу выходного диапазона. Размеры выходной области будут рассчитаны автоматически, и соответствующее сообщение появится на экране в том случае, если выходной диапазон занимает место существующих данных или его размеры превышают размеры листа.

Новый лист. Установите переключатель, чтобы открыть новый лист в книге и вставить результаты анализа, начиная с ячейки A1. Если в этом есть необходимость, введите имя нового листа в поле, расположенном напротив соответствующего положения переключателя.

Новая книга. Установите переключатель, чтобы открыть новую книгу и вставить результаты анализа в ячейку A1 на первом листе в этой книге.

В результате обработки данных получили следующее:

	A	B	C	D	E	F	G
1	Однофакторный дисперсионный анализ						
2							
3	ИТОГИ						
4	<i>Группы</i>	<i>Счет</i>	<i>Сумма</i>	<i>Среднее</i>	<i>Дисперсия</i>		
5	I группа (контр.)	5	1673	334,6	56,8		
6	II группа	5	1812	362,4	220,8		
7	III группа	5	1885	377	276,5		
8							
9	ANOVA						
10	<i>Источник вариации</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-Значение</i>	<i>F критическое</i>
11	Между группами	4640	2	2319,8	12,55983	0,0011415	3,885290312
12	Внутри групп	2216	12	184,7			
13							
14	Итого	6856	14				
15							
16							
17							

Рисунок 3 – Результаты однофакторного дисперсионного анализа

Таблица ИТОГИ:

“Счет” – число повторностей. “Сумма” – сумма значений показателя по строкам. “Дисперсия” – частная дисперсия показателя.

Таблица ANOVA представляет результаты дисперсионного анализа однофакторного комплекса, в котором первая колонка “Источник вариации” содержит наименование дисперсий. Графа “SS” – это сумма квадратов отклонений, “df” – степень свободы, графа “MS” – средний квадрат, “F” – критерий фактического F – распределения. “P – значение” – вероятность того, что дисперсия, воспроизводимая уравнением, равна дисперсии остатков. Определяет вероятность того, что полученная количественная определенность взаимосвязи между факторами и результатом может считаться случайной. “F - критическое” – это значение F – теоретического, которое впоследствии сравнивается с F – фактическим.

#### 4. Сформулировать выводы.

##### Пример:

**Задание 0 варианта.** Для проверки влияния методики обучения производственным навыкам на качество подготовки отбираются случайным образом из выпускников ПТУ четыре группы учеников, которые после окончания обучения показали следующие производственные результаты (Таблица 1).

Группа (методика)	Выработка, шт./день						Число учеников
1	60	80	75	80	85	70	6
2	75	66	85	80	70	80	90
3	60	80	65	60	86	75	6
4	95	85	100	80			4

Проверить существенность влияния методики обучения на производственные результаты учеников

##### Выполнение.

1. Создать файл с исходными данными (рис 4)

	Методика 1	Методика 2	Методика 3	Методика 4
1	60	75	60	95
2	80	66	80	85
3	75	85	65	100
4	80	80	60	80
5	85	70	86	
6	70	80	75	
7		90		

Рисунок 4 – Исходные данные.

2. Запустить “Пакет анализа” (рис 5)

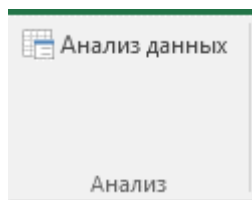


Рисунок 5 – Пакет анализа

3. Провести однофакторный дисперсионный анализ

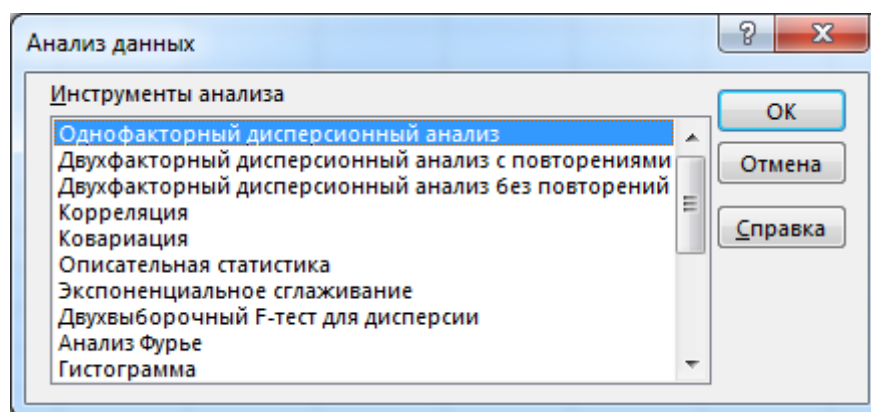


Рисунок 6 – Запуск процедуры «Однофакторный дисперсионный анализ»



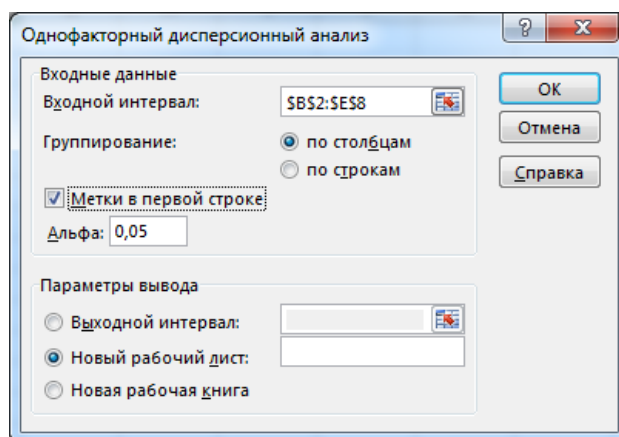


Рисунок 7 – Параметры однофакторного дисперсионного анализа

Однофакторный дисперсионный анализ						
ИТОГИ						
Группы	Счет	Сумма	Среднее	Дисперсия		
Столбец 1	6	450	75	80		
Столбец 2	7	546	78	69,66666667		
Столбец 3	6	426	71	120		
Столбец 4	4	360	90	83,33333333		
Дисперсионный анализ						
Источник вариации	SS	df	MS	F	P-Значение	F критическое
Между группами	917,7391	3	305,913	3,484620999	0,036142647	3,127350005
Внутри групп	1668	19	87,78947			
Итого	2585,739	22				

Рисунок 8 – Результирующая таблица выполненного дисперсионного анализа.

#### 4. Сформулировать выводы

Таблица ИТОГИ:

Сравнение средних значений показывает, что Методика 4 (столбец 4) позволяет добиться лучшего результата, а Методика 3 наименее эффективна.

Назначение итоговой таблицы дисперсионного анализа проверить нулевую гипотезу  $H_0$ : об отсутствии значимого влияния уровней факторов на исследуемый отклик.

Сравнение  $F$  и  $F_{\text{критическое}}$  показывает, что  $F > F_{\text{критическое}}$ , следовательно отвергнута гипотеза  $H_0$  и принята гипотеза  $H_1$  и с вероятностью ошибки  $\alpha = 0,05$  можно утверждать, что влияние рассматриваемого фактора на результативный признак существенно.

### Дисперсионный анализ средствами языка R

1. Создать набор данных согласно варианту  
Набор данных создается или с использованием среды Rcmdr или импортируется.
2. Провести однофакторный дисперсионный анализ в среде Rcmdr  
Статистика  
Средние  
Одномерный дисперсионный анализ
3. По результатам дисперсионного анализа сформулировать выводы.
4. Построить диаграмму, отображающую средние значения и их доверительные интервалы для каждой группы  
График  
График средних

**Пример:** Задание с предыдущего примера.

1. Создать набор данных

	Vyr	Metod
1	60	Методика1
2	80	Методика1
3	75	Методика1
4	80	Методика1
5	85	Методика1
6	70	Методика1
7	75	Методика2
8	66	Методика2
9	85	Методика2
10	80	Методика2
11	70	Методика2
12	80	Методика2
13	90	Методика2
14	60	Методика3
15	80	Методика3
16	65	Методика3
17	60	Методика3
18	86	Методика3
19	75	Методика3
20	95	Методика4
21	85	Методика4
22	100	Методика4
23	80	Методика4

Рисунок 9 – Набор данных в среде Rcmdr

2. Провести однофакторный дисперсионный анализ в среде Rcmdr

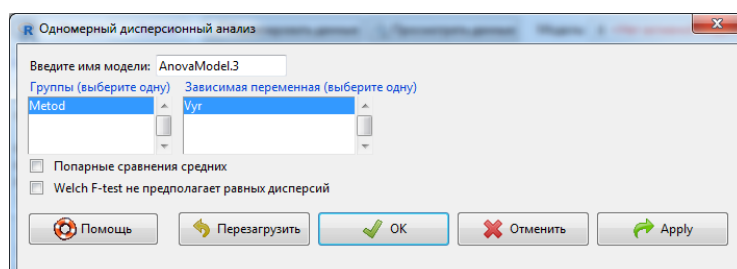


Рисунок 10 - Одномерный дисперсионный анализ в среде Rcmdr

```
> AnovaModel.3 <- aov(Vyr ~ Metod, data=Dataset)

> summary(AnovaModel.3)
              Df Sum Sq Mean Sq F value Pr(>F)
Metod           3   917.7   305.91    3.485 0.0361 *
Residuals      19  1668.0    87.79
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> with(Dataset, numSummary(Vyr, groups=Metod, statistics=c("mean", "sd")))
      mean      sd data:n
Методика1  75  8.944272     6
Методика2  78  8.346656     7
Методика3  71 10.954451     6
Методика4  90  9.128709     4
```

Рисунок 11 – Результирующая таблица

3. По результатам дисперсионного анализа сформулировать выводы

Строка, обозначенная как *Metod*, соответствует источнику дисперсии в данных, связанному с действием изучаемого экспериментального фактора/

Строка, обозначенная как *Residuals*, характеризует внутригрупповую дисперсию.

В столбце *F value* представлено рассчитанное по имеющимся данным значение F-критерия, он равен 3,485. В столбце *Pr(>F)* представлена вероятность получить F-значение, равное или превышающее то значение, которое в действительности рассчитали по имеющимся выборочным данным (при условии, что нулевая гипотеза верна). Как видно, эта вероятность не

высокая, равна 0,0361 (3,61%). Не превышает 5%-ный уровень значимости, в связи с чем мы заключаем, что нулевая гипотеза не верна. Таким образом, можно утверждать, что экспериментальные условия оказали существенное влияние на результативный признак.

Сравнение средних значений показывает, что Методика 4 позволяет добиться лучшего результата, а Методика 3 наименее эффективна.

4. Построить диаграмму, отображающую средние значения и их доверительные интервалы для каждой группы

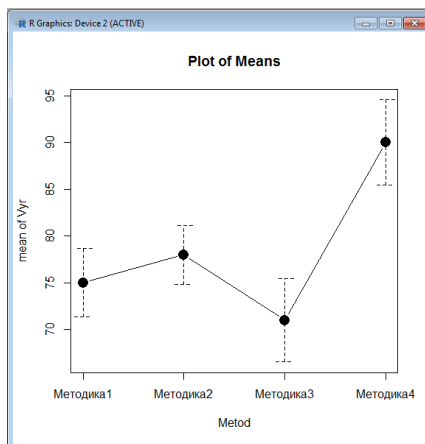


Рисунок 12 – Диаграмма, отображающая средние значения и их доверительные интервалы

Данная диаграмма визуально подтверждает, что Методика 4 позволяет добиться лучшего результата, а Методика 3 наименее эффективна

### Контрольные вопросы

1. Дисперсионный анализ.
2. Виды дисперсионного анализа.
3. В чем заключается основная идея дисперсионного анализа?
4. Однофакторный дисперсионный анализ.
5. Функции для выполнения дисперсионного анализа на языке R.

### Библиография

1. Алексей Шипунов и др. Наглядная статистика. Используем R! – М.: ДМК Пресс, 2014. – 298 с. [Электронный ресурс]. Режим доступа: <http://ashipunov.info/shipunov/school/books/rbook.pdf>.
2. Зарядов И.С. Введение в статистический пакет R: типы переменных, структуры данных, чтение и запись информации, графика. М.: Издательство Российского университета дружбы народов, 2010. – 207 с.
3. Роберт И. Кабаков R в действии. Анализ и визуализация данных в программе R / пер. с англ. Полины А. Волковой. – М.: ДМК Пресс, 2014. – 588 с.: ил.
4. Официальный сайт RStudio. Режим доступа: <https://www.rstudio.com>.
5. Профессиональный информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных [Электронный ресурс]. Режим доступа: <http://machinelearning.ru>.
6. Мاستицкий С.Э., Шитиков В.К. (2014) Статистический анализ и визуализация данных с помощью R. – Электронная книга. Режим доступа: <http://r-analytics.blogspot.com>

## Приложение А Варианты заданий

### Задача 1

При исследовании влияния стажа работы на производительность труда (количество деталей в день) в одном из цехов завода получен следующий однофакторный дисперсионный комплекс (таблица А1):

Таблица А.1.

Номер наблюдения	Стаж работы рабочих (лет)			
	До 5	5-10	10-15	15-20
Вариант 1				
1	155	154	153	164
2	153	158	162	162
3	149	157	164	163
4	150	161	163	
5			167	
Вариант 2				
1	147	149	155	160
2	149	150	149	163
3	153	152	156	166
4		148	161	
5			160	
Вариант 3				
1	155	159	158	168
2	158	161	166	167
3	156	164	163	164
4	154	158	165	166
5		163		
Вариант 4				
1	172	175	177	183
2	170	178	183	176
3	169	171	181	182
4		169	180	179
5		174		

### Задача 2

В процессе исследования влияния цены за единицу продукции на объем продаж (шт.) в месяц были получены следующие результаты (таблица А2):

Таблица А2

Номер наблюдения	Цена за единицу продукции (руб.)			
	1000-1100	1100-1200	1200-1300	1300-1500
1	215	218	214	211
2	221	214	217	210
3	222	220	210	208
4	219	221		209
5		213		
Вариант 5				
1	267	266	262	264
2	270	271	265	265
3	275	272	267	260

4		265	268	259
5				261
Вариант 6				
1	310	311	308	299
2	314	309	307	287
3	311	305	300	301
4		307		300
Вариант 7				
1	56	55	49	44
2	58	52	51	43
3	55	53	45	39
4	59	48	41	
5		50	46	
Вариант 8				
1	97	85	89	79
2	93	88	83	81
3	96	90	85	80
4		94		82

Задача 3. Дана однофакторная таблица зависимости объема выручки (млн. руб.) от расходов на рекламу (тыс. руб.) (таблица А3):

Таблица А3

Номер исследования	Расходы на рекламу (тыс. руб.)			
	150-200	200-250	250-300	300-400
Вариант 9				
1	6,3	6,9	6,8	6,7
2	5,6	5,7	7,1	7,3
3	7,2	6,8	7,0	6,9
4	4,7		6,5	
Вариант 10				
1	7,1	7,4	7,5	7,7
2	7,3	7,9	7,7	7,9
3	7,7	8,4	7,2	7,8
4		7,6	7,8	8,3
5				8,0
Вариант 11				
1	6,6	6,0	8,4	8,7
2	5,9	6,8	7,5	7,8
3	6,4	7,4	6,9	7,1
4	7,1	8,1	7,3	7,6
5			7,7	
Вариант 12				
1	5,2	5,8	5,7	6,4
2	5,9	5,7	6,7	5,9
3	5,6	6,2	6,4	6,8
4		6,4	6,8	

Задача 4. Туристическими агентствами исследовалось влияние расстояния от пляжа (км) на наполняемость гостиниц (%). Были получены следующие данные (таблица А4):

Таблица А4

Номер исследования	Расстояние до пляжа (км)			
	До 1	1-2	2-4	4-6
Вариант 13				
1	99	98	96	89
2	98	97	94	90
3	100	99	95	93
4		97	94	92
5		96		91
Вариант 14				
1	98	97	98	90
2	97	96	95	87
3	99	94	94	94
4		99	96	88
5			97	
Вариант 15				
1	100	98	94	94
2	97	99	100	95
3	98	96	93	89
4		98	92	
5			90	
Вариант 16				
1	98	99	89	88
2	99	96	94	89
3	97	93	93	95
4		94	91	96
5			87	

Задача 5. Исследовалось влияние диеты на количество сброшенных килограммов за один месяц. Были получены следующие данные (таблица А5):

Таблица А5

Номер исследования	Номер диеты			
	1	2	3	4
Вариант 17				
1	3,2	4,5	3,3	4,1
2	1,6	3,4	5,4	1,7
3	2,3	1,8	2,9	3,9
4		1,9	3,0	
Вариант 18				
1	3,6	2,0	4,2	6,2
2	5,1	4,3	6,0	5,7
3	4,7	5,0	3,7	3,3
4	1,9	3,8	5,3	4,1
5			4,6	
Вариант 19				
1	2,2	4,2	5,5	5,1
2	4,1	4,3	4,7	4,3
3	2,6	5,6	3,9	6,0
4		3,9	5,0	4,2
5				3,8
Вариант 20				

1	5,2	6,2	4,8	7,1
2	4,6	6,6	4,9	7,6
3	5,1	5,1	5,7	4,5
4	5,5		4,3	5,8
5				5,9

Задача 6. На предприятии были проведены исследования влияния периода реализации продукции а объем выручки (млн. руб.) (таблица А6):

Таблица А6

Номер исследования	Период проведения исследования			
	I квартал	II квартал	III квартал	IV квартал
Вариант 21				
1	150	156	149	155
2	154	154	152	166
3	155	161	148	162
4		155		
Вариант 22				
1	168	166	154	164
2	167	159	159	163
3	159	163	151	157
4	156	158		160
5		167		
Вариант 23				
1	191	200	197	187
2	198	210	211	195
3	201	199	204	188
4		189	208	
5			194	
Вариант 24				
1	179	191	189	189
2	181	189	182	197
3	183	196	176	199
4	189	182		201
5		188		

Задача 7. Исследовалось влияние количества осадков за год на урожайность пшеницы (т/га). Получены следующие данные (таблица А7):

Таблица А7

Номер исследования	Количество осадков (мм)			
	250-260	260-270	270-280	280-290
Вариант 25				
1	33	29	34	34
2	31	30	37	36
3	32	28	33	38
4		33		
Вариант 26				
1	28	29	33	37
2	29	32	35	38
3	31	33	31	34
4		27	36	
5			32	

Вариант 27				
1	39	28	31	32
2	30	26	33	37
3	27	31	28	35
4		32	35	31
5			30	30
Вариант 28				
1	28	30	29	33
2	26	29	34	34
3	29	34	31	37
4			33	36
5			30	