Лабораторная работа №3

«Задача дисперсионного анализа. Методы дисперсионного

анализа. Однофакторный дисперсионный анализ»

1 ЦЕЛЬ РАБОТЫ

– приобрести практические навыки в проведении дисперсионного анализа по экспериментальным данным;

– исследовать возможности языка R для проведения дисперсионного анализа.

2 ХОД РАБОТЫ

1. Произведём однофакторный дисперсионный анализ данных 9 варианта, для заданной однофакторной таблицы зависимости объёма выручки (млн. руб.) от расходов на рекламу (тыс. руб.) (таблица 1).

Таблица 1:

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

С помощью «Пакета анализа» программы Excel произведём анализ, результаты которого продемонстрированы на рисунке 1:

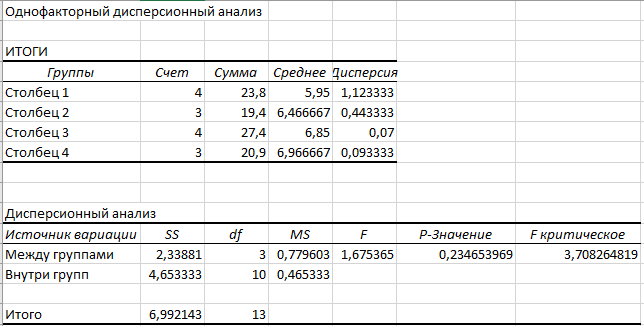


Рисунок 1 – Результирующая таблица выполненного дисперсионного анализа

Сформулируем выводы полученного анализа.

Исходя из таблицы «ИТОГИ», сравнение средних значений показывает, что при расходах на рекламу в 300-400 тыс. руб. объёмы выручки достигают лучшего результата, а при расходах 150-200 тыс. руб. наименьшая выручка.

Проанализируем таблицу «Дисперсионный анализ». Сравнение F и Fкритическое показывает, что F < Fкритическое, следовательно, гипотеза Н0 об отсутствии значимого влияния уровней факторов на исследуемый отклик не отвергнута и можно утверждать, что влияние рассматриваемого фактора на результативный признак не существенно.

Далее проведём дисперсионный анализ средствами языка R пакета Rcmdr (см. рисунок 2).

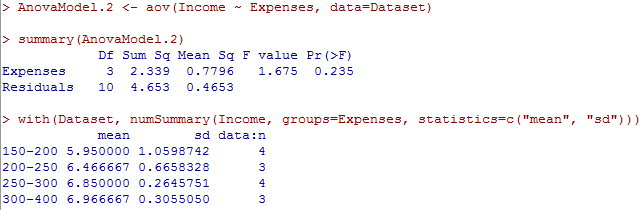


Рисунок 2 – Результат одномерного дисперсионного анализа в Rcmdr

В столбце F value представлено рассчитанное по имеющимся данным значение F-критерия, равное 1,675. 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 равна 0,235 (23,5%), что превышает 5%-ный уровень значимости. В связи с этим мы заключаем, что нулевая гипотеза верна. Таким образом, экспериментальные условия оказали несущественное влияние на результативный признак.

Сравнение средних значений показывает, что расход на рекламу на сумму 300-400 тыс. руб. позволяет добиться лучшего результата, а вливание на рекламу средств на сумму 150-200 тыс. руб. наименее эффективно.

Построим диаграмму, отображающую средние значения и их доверительные интервалы для каждой группы (см. рисунок 3).



Рисунок 3 – График средних

Данная диаграмма на рисунке 3 визуально подтверждает, что реклама на 300-400 тыс. руб. позволяет добиться большей прибыли, нежели с рекламой за 150-200 тыс. руб.

1. Произведём дисперсионный анализ экспериментальных данных с помощью «Пакета анализа» Excel.

Проанализируем зависимость употребления алкоголя сверх установленной нормы от уровня безработицы населения. Создадим файл с исходными данными в Excel (см. рисунок 4).

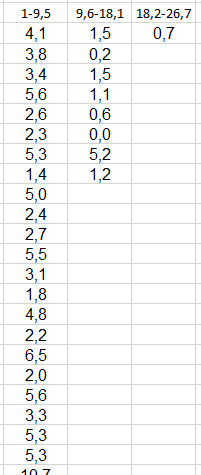


Рисунок 4 – Экспериментальные данные в Excel

Результаты произведённого однофакторного дисперсионного анализа приведены на рисунке 5.

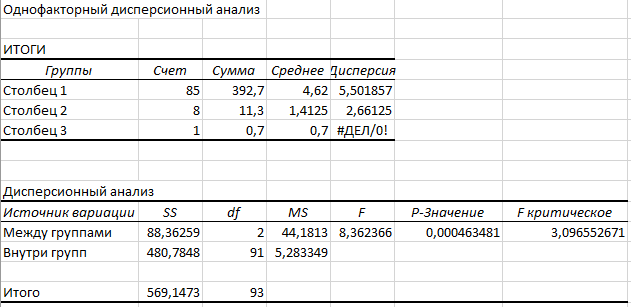


Рисунок 5 – Результирующая таблица выполненного дисперсионного анализа

Сформулируем выводы по данному анализу.

Исходя из таблицы «ИТОГИ», сравнение средних значений показывает, что при уровне безработицы от 1 до 9,5 употребление алкоголя сверх нормы максимально, а при уровне от 18,2 до 26,7 минимально.

Сравнение F и Fкритическое показывает, что F > Fкритическое, следовательно отвергнута гипотеза Н0 и принята гипотеза Н1, и с вероятностью ошибки α = 0,05 можно утверждать, что влияние рассматриваемого фактора на результативный признак существенно.

Далее проведём дисперсионный анализ средствами языка R.

Создадим набор данных (см. рисунок 6).

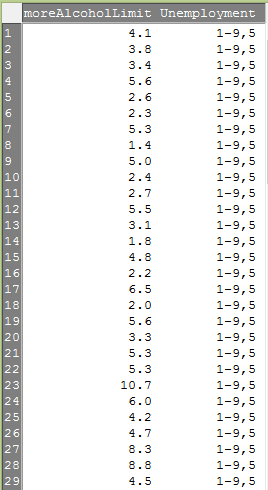


Рисунок 6 – Набор данных в среде Rcmdr

Результаты произведённого однофакторного анализа приведены на рисунке 7.

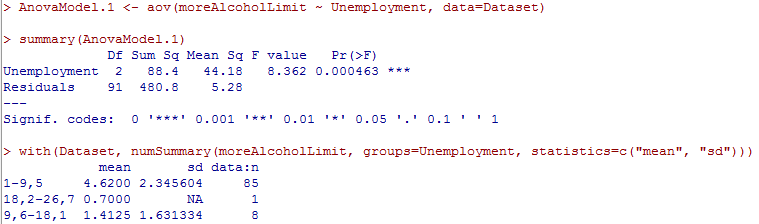


Рисунок 7 – Результат одномерного дисперсионного анализа в Rcmdr

По результирующей таблице можно сделать следующие выводы: значение F-критерия равно 8,362, а вероятность получить F-значение, равное или превышающее то значение, которое в действительности рассчитали по имеющимся данным при условии, что нулевая гипотеза верна, принимает очень маленькое значение 0,000463 (0,04%), что не превышает 5%-уровень значимости, в связи с чем можно заключить, что нулевая гипотеза не верна. Следовательно, экспериментальные условия оказали существенное влияние на результативный признак.

Другими словами, уровень безработицы с 1 до 9,5 приводит к употреблению алкоголя в большей степени, нежели уровень безработицы с 18,2 до 26,7.

Построим диаграмму, отображающую средние значения и их доверительные интервалы для каждой группы (см. рисунок 8):



Рисунок 8 – График средних

Данная диаграмма визуально подтверждает, что уровень безработицы с 1 до 9,5 приводит к употреблению алкоголя в большей степени, нежели уровень безработицы с 18,2 до 26,7.

1. Ответим на контрольные вопросы.

Дисперсионный анализ – это статистический метод, предназначенный для оценки влияния различных факторов на результат эксперимента, а также для последующего планирования аналогичных экспериментов.

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

* 1. Виды дисперсионного анализа. Однофакторный и многофакторный ДА.
  2. В чем заключается основная идея дисперсионного анализа? Дисперсионный анализ состоит в выделении и оценке отдельных факторов, вызывающих изменчивость изучаемой случайной величины. 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.
  3. Однофакторный дисперсионный анализ.

ВЫВОДЫ

В ходе данной лабораторной работы были приобретены практические навыки в проведении дисперсионного анализа по экспериментальным данным, а также исследованы возможности языка R и программы Excel для проведения дисперсионного анализа.