МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ  
«Харківський Політехнічний Інститут»  
  
Кафедра Стратегічного Управління

ЗВІТ

з лабораторної роботи № 13

з дисципліни Математична статистика

на тему

«Двухфакторний дисперсіонний аналіз без повторень»

Перевірила: старший викладач  
Мошко Є. О.  
Виконав: ст. гр. КН-27

Харків, 2019

Цель: Провести двухфакторный дисперсионный анализ без повторений.

Задание: Можно ли утверждать, что фильмы снятые одним режиссером более успешны, чем фильмы других режиссеров. Существует ли разница между успехом фильмов, снятых в разных жанрах разными режиссерами? Определить какой режиссер и какой жанр в комбинации дают самые успешные фильмы.

Ход работы.

Теоретическая часть.

Также как и в однофакторном дисперсионном анализе, в двухфакторном присутствуют факторы. В этом случае их два (A и B). Уровням фактора соответствуют различные значения этого фактора. Обозначим уровни фактора A буквой i, а уровни фактора B j. Измеренные значения зависимой переменной Y будем обозначать yij.

Если фактор действительно влияет на зависимую переменную, то при различных уровнях фактора мы должны в среднем получать различные средние значения выборок.

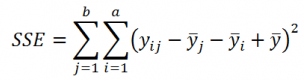
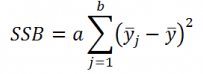
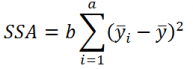
Среднее значение при различных уровнях фактора A будет отличаться:

Среднее значение при различных уровнях фактора B будет отличаться:

Общее среднее вычисляется по формуле:

Общая изменчивость по аналогии с однофакторным дисперсионным анализом рассчитывается как комбинация нескольких компонентов.

SST=SSA+SSB+SSE



В случае двухфакторного дисперсионного анализа формируются две нулевых гипотезы. Первая гипотеза H01 заключается в том, что уровень фактора A не влияет на измеренные значения. Соответственно вторая гипотеза H01 заключается в том, что уровень фактора B не влияет на измеренные значения.

В двухфакторном дисперсионном анализе участвуют две статистики:

Полученные статистики сравниваются с F критическим, вычисленное для заданного значения уровня альфа. Если , то нулевую гипотезу об отсутствии влияния уровней фактора А отклоняют. Соответственно если , то нулевую гипотезу об отсутствии влияния уровней фактора В отклоняют.

Практическая часть.

Дан список оценок фильмов на сайте для разных режиссеров и разных жанров – рисунок 1.



Рисунок 1 – Исходные данные задачи

Необходимо, воспользовавшись инструментом «двухфакторный дисперсионный анализ» пакета анализа, установить зависимость между различными факторами (режиссер и жанр) и оценками фильма.

Выполнение.

Воспользуемся пакетом анализа для проведения дисперсионного анализа – рисунок 2.

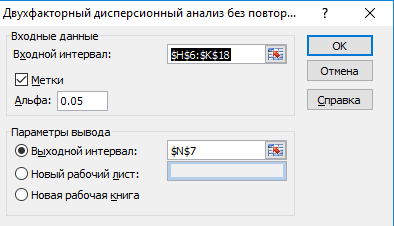


Рисунок 2 – Задание параметров анализа

Получим следующие данные – рисунок 3.



Рисунок 3 – Полученные данные

- Как видно для каждого уровня двух факторов рассчитаны Счет, Сумма, Среднее и Дисперсия. Далее проводится дисперсионный анализ.

- В столбце SS данные соответствуют параметрам SSA, SSB, SSE.

- df – число степеней свободы, необходимо для нахождения квантилей доверительных интервалов.

- В столбце MS данные соответствуют параметрам MSA, MSB, MSE.

Выводы.

Анализируя полученные данные можем утверждать, что нет оснований для отклонения обеих нулевых гипотез, так как

То есть мы сделали вывод о том, что выбор режиссера влияет на оценку фильма, а выбор жанра оказывает минимальное влияние.