# ЛАБОРАТОРНАЯ РАБОТА 3

# Проверка статистических гипотез. Гипотеза однородности

**Цель работы:**

Проверка статистических гипотез о однородности выборочных совокупностей, используя критерии Хи-квадрат и Колмогорова.

**Задание.**

В файле “Данные” на листе “Котировки” приведены цены закрытия акций на ММВБ за период с 05.01.2015 по 02.09.2015 с периодичностью 1 день (торговый). Для указанной в варианте задания ценной бумаги, используя критерии однородности Колмогорова и Хи-квадрат, на уровне значимости проверить гипотезу о том, что распределения ежедневных относительных приращений котировок за периоды с 05.01.2015 по 04.30.2015 и с 05.05.2015 по 02.09.2015 не различаются. Указать достигнутый уровень значимости (для критерия Хи-квадрат).

***Вариант 18 – TATN***

**Ход работы:**

Предварительно получим выборки и относительных приращений цен акций объемами соответственно и (чтобы получить относительное приращение цены, из последующей цены вычитаем предыдущую и делим на предыдущую). Заметим, что для второй выборки, на основе имеющихся данных, за указанный период можно получить в принципе и 85 значений относительных приращений цен, но будем полагать, что у нас имеются как бы две несвязанные выборки цен.

Для каждой выборки определяем наименьшее и наибольшее значения:

Для : -0,06645, 0,071038.

Для : -0,06848, 0,061357.

Число интервалов группирования для статистических рядов определяем по формуле Стерджесса по выборке наименьшего объема: . В качестве нижней границы первого интервала возьмем значение -0,08, а в качестве верхней границы последнего интервала значение , тогда длина каждого интервала группирования (при равных длинах всех интервалов) будет равна: . Подсчитывая частоты, получаем следующие два ряда:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Интервал | (-0,08; -0,06) | (-0,06; -0,04) | (-0,04; -0.02) | (-0;02,0) | (0;0,02) | (0,02  ;0,04) | (0,04;0,06) | (0,06;0,08) |
|  | 1 | 7 | 14 | 18 | 11 | 17 | 8 | 4 |
|  | 1 | 3 | 6 | 30 | 35 | 7 | 1 | 1 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Интервал | (-0,08; -0,03) | (-0,03; -0,015) | (-0,015; -0,005) | (-0,005;0,005) | (0,005;0,014) | (0,014  ;0,028) | (0,028;0,035) | (0,035;0,08) | ni |
|  | 15 | 11 | 10 | 7 | 7 | 6 | 9 | 15 | 80 |
|  | 6 | 10 | 15 | 18 | 16 | 3 | 3 | 3 | 84 |
|  | 21 | 21 | 25 | 25 | 23 | 19 | 12 | 18 | 164 |

Видим, что частоты распределены по интервалам крайне неравномерно (для корректного применения критерия хи-квадрат необходимо, чтобы количество значений попавших в каждую ячейку было не меньше 5), поэтому делаем перегруппировку данных, добиваясь более равномерного распределения частот по интервалам. В результате получаем следующие ряды:

Для проверки гипотезы однородности распределений двух выборок вычисляем значение статистики критерия Хи-квадрат:

26,76381.

Зададим уровень значимости . Для заданного уровня значимости и числа степеней свободы найдем критическое значение статистики, как критическую точку распределения уровня (или что тоже самое – квантиль уровня 0,95): 14,06714. Так как , то гипотеза о однородности двух выборок (равенстве распределений двух совокупностей) не принимается.

Найдем также достигнутый уровень значимости, то есть такое значение , для которого при истинности нашей гипотезы Для 7 степеней свободы и 26,76381 получим 0,000368. Поскольку данная вероятность меньше 0,05, гипотеза о однородности выборочных совокупностей не принимается.

Для проверки гипотезы однородности по критерию Колмогорова, строим для одной из выборок (например, для второй) вариационный ряд , упорядочивая выборочные значения по возрастанию. Далее для каждого значения  вычисляем величину , где  - значение эмпирической функции распределения для первой выборки в точке . Из полученных значений отбираем наибольшее и получаем значение статистики:

1,714612

Для уровня значимости  критическое значение статистики есть квантиль распределения Колмогорова уровня 0,95:  (берем из Приложения 3 ниже).

Так как , то гипотеза о однородности двух выборок (равенстве распределений двух совокупностей) не принимается.

На основе имеющейся таблицы можно утверждать, что достигнутый уровень значимости .

Таким образом гипотеза не принимается.