# ЛАБОРАТОРНАЯ РАБОТА 4

# Проверка статистических гипотез. Гипотеза независимости

**Цель работы:**

Проверка статистических гипотез о независимости выборочных совокупностей, используя критерий Хи-квадрат. Оценка значимости коэффициента корреляции Пирсона.

**Задание.**

В файле “Данные” на листе “Котировки” приведены цены закрытия акций на ММВБ за период с 05.01.2015 по 01.09.2015 с периодичностью 1 день. Для указанных в варианте задания ценных бумаг, проверить гипотезу о независимости ежедневных относительных приращений котировок, используя критерий независимости Хи-квадрат. Оценить также наличие, степень и характер связи, используя выборочный коэффициент корреляции Пирсона.

|  |  |
| --- | --- |
| ВАРИАНТ | 18 |
| Акция 1 | VTBR |
| Акция 2 | MTLR |

***Вариант 18.***

**Ход работы:**

По формуле Стерджесса получаем, при объеме выборки , рекомендуемое число ячеек для таблицы сопряженности равным . Если мы для каждой из переменных определим 3 интервала группирования, то для таблицы сопряженности получим: ячеек, то есть, значение, близкое к рекомендуемому.

Для определения границ интервалов построим статистические ряды для относительных приращений котировок акций VTBR и MTLR, содержащие по 3 интервала:

Интервальный ряд для относительных приращений VTBR:

|  |  |  |  |
| --- | --- | --- | --- |
| Интервал | (-¥; -0,007) | (-0,007;0,007) | (0,007; +¥) |
| Код интервала | 1 | 2 | 3 |
| Частота | 55 | 69 | 41 |

Интервальный ряд для относительных приращений MTLR:

|  |  |  |  |
| --- | --- | --- | --- |
| Интервал | (-¥; -0,007) | (-0,007;0,007) | (0,007; +¥) |
| Код интервала | 1 | 2 | 3 |
| Частота | 63 | 36 | 66 |

Чтобы воспользоваться средствами построения таблиц сопряженности EXCEL, строим следующие ряды: два ряда кодов исходных переменных и ряд, состоящий из единиц, который указывает количество пар таких значений. Код переменной присваивается в соответствии с тем, к какому интервалу она относится. Фрагмент исходных рядов относительных приращений и полученных рядов кодов переменных приведен ниже:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| VTBR | MTLR | Код VTBR | Код MTLR | Кол-во |
| -0,01244 | 0,013216 | 1 | 3 | 1 |
| 0,011251 | 0,043478 | 3 | 3 | 1 |
| -0,03115 | 0,003788 | 1 | 2 | 1 |
| -0,02725 | -0,00377 | 1 | 2 | 1 |
| -0,04769 | -0,05303 | 1 | 1 | 1 |
| -0,00661 | 0,01 | 1 | 3 | 1 |
| 0,04376 | 0,034851 | 3 | 3 | 1 |
| 0,039694 | 0,005741 | 3 | 2 | 1 |
| -0,01855 | 0,019787 | 1 | 3 | 1 |
| -0,02984 | 0,016045 | 1 | 3 | 1 |
| 0,022544 | 0,180683 | 3 | 3 | 1 |
| 0,00189 | 0,105754 | 2 | 3 | 1 |
| 0,002829 | -0,03235 | 2 | 1 | 1 |

Строим таблицу сопряженности для переменных Код VTBR и Код MTLR:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Частота** | **MTLR** |  |  |  |
| **VTBR** | (-∞; -0,007) | (-0,007;0,007) | (0,007; +∞) | **Общий итог** |
| (-∞; -0,007) | 31 | 8 | 16 | **55** |
| (-0,007;0,007) | 20 | 21 | 28 | **69** |
| (0,007; +∞) | 12 | 7 | 22 | **41** |
| **Общий итог** | **63** | **36** | **66** | **165** |

Строим таблицу ожидаемых частот, вычисляя ожидаемые частоты для каждой ячейки по формуле: , где и - маргинальные частоты по строкам и столбцам соответственно:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Частота** | **MTLR** |  |  |  |
| **VTBR** | (-∞; -0,007) | (-0,007;0,007) | (0,007; +∞) | **Общий итог** |
| (-∞; -0,007) | 21,00 | 12,00 | 22,00 | **55** |
| (-0,007;0,007) | 26,34 | 15,05 | 27,60 | **69** |
| (0,007; +∞) | 15,65 | 8,94 | 16,40 | **41** |
| **Общий итог** | **63** | **36** | **66** | **165** |

Отмечаем, что все ожидаемые частоты больше 5 (что важно для корректного применения критерия Хи-квадрат).

Вычисляем значение статистики критерия Хи-квадрат:

14,80.

Для числа степеней свободы статистики находим наблюдаемый уровень значимости критерия при условии истинности гипотезы : «относительные приращения независимы» 0,005. Согласно данному значению, следует отвергнуть нулевую гипотезу о независимости ежедневных относительных приращений котировок акций VTBR и MTLR и принять гипотезу о зависимости данных величин.

Исследуем также наличие связи между относительными приращениями котировок, используя коэффициент корреляции Пирсона, который характеризует степень линейной связи между величинами. Выборочный коэффициент корреляции Пирсона между величинами и вычисляется по формуле:

.

В нашем случае 0,269.

Проверяем значимость . Вычисляем статистику Стьюдента: 3,57. Находим уровень значимости по формуле: , где случайная величина, имеющая распределение Стьюдента, с числом степеней свободы . Получаем 0,00046 , то есть, значение является статистически высоко значимым, следовательно, следует признать наличие положительной слабо значимой корреляции между относительными приращениями.