

ОП «Политология», 2023-24**Введение в ТВиМС****Концепция p-value. Коэффициент корреляции Пирсона. (15 мая)***А. А. Макаров, А. А. Тамбовцева*

Задача 1. Студент проверял гипотезу о равенстве доли сторонников консервативной партии в стране N значению 0.56 против двусторонней альтернативы. Наблюдаемое значение статистики критерия оказалось равным 1.72.

- (a) Запишите нулевую и альтернативную гипотезы. Постройте критическую область, считая уровень значимости равным 5%. Сделайте статистический и содержательный вывод относительно нулевой гипотезы.
- (b) Посчитайте p-value. Объясните, какой содержательный смысл имеет это значение, и каким должно оно быть, чтобы гипотеза о равенстве доли сторонников консерваторов НЕ отвергалась на 5%-ном уровне значимости.

Задача 2. По результатам некоторого социологического исследования, проведенного среди 225 офисных сотрудников компании «Vita activa», известно, что 70% сотрудников не успевают утром завтракать дома, и уезжают на работу голодными. Проверьте на 1% уровне значимости гипотезу о том, что доля сотрудников, успевающих завтракать дома, равна 0.25, против односторонней альтернативы.

- (a) Запишите нулевую и альтернативную гипотезы. Вычислите наблюдаемое значение статистики критерия.
- (b) Вычислите p-value. Сделайте статистический и содержательный вывод относительно нулевой гипотезы.

Задача 3. По результатам выборочного обследования (выборка из 169 сотрудников компании «Vita activa») было зафиксировано следующее: средняя заработная плата равна 48 тысячам рублей, стандартное отклонение заработной платы равно 10 тысячам рублей. Можно ли считать среднюю заработную плату всех сотрудников равной 50 тысячам рублей?

- (a) Запишите нулевую гипотезу и одностороннюю альтернативную гипотезу. Постройте критическую область, считая уровень значимости равным 10%. Сделайте статистический и содержательный вывод относительно нулевой гипотезы.
- (b) Посчитайте p-value. Сделайте статистический и содержательный вывод относительно нулевой гипотезы, считая уровень значимости равным 10%.

Задача 4. В ходе учебного исследования 30 студентам были заданы следующие вопросы: «Сколько времени вы обычно расходуете на сон в будни (укажите с точностью до получаса)?» и «Сколько обычно времени вы проводите в социальных сетях в течение дня (укажите с точностью до получаса)?» Известно, что коэффициент Пирсона, посчитанный на основе полученных ответов, оказался равным -0.4 . Можно ли считать, что между временем, проводимым в социальных сетях и временем, которое студенты тратят на сон, существует линейная связь?

- (a) Сформулируйте подходящую нулевую гипотезу, которую необходимо проверить для ответа на этот вопрос. Сформулируйте двустороннюю альтернативную гипотезу.
- (b) Какое распределение имеет статистика критерия? Найдите наблюдаемое значение статистики критерия.
- (c) Приняв уровень значимости равным 10%, постройте критическую область, проверьте гипотезу и сделайте статистический и содержательный выводы.
- (d) Вычислите (примерно) p -value. Сделайте статистический и содержательный выводы, приняв уровень значимости равным 5%.

Задача 5. Для 20 судей Верховного суда США посчитаны значения двух показателей: индекс подготовленности к судебному процессу и индекс способности справляться с потоком дел. Выборочная ковариация между этими показателями равна 0.78, выборочная дисперсия первого индекса равна 0.9, выборочная дисперсия второго индекса равна 0.7.

- (a) Вычислите коэффициент корреляции Пирсона. Проинтерпретируйте его: укажите направление и силу связи между показателями.
- (b) Вычислите наблюдаемое значение статистики критерия и примерное p -value. Проверьте на 5%-ном уровне значимости, есть ли связь между подготовленностью судьи к процессу и его способностью справляться с потоком дел.

Источники:

1. Задача 2. Идея основана на задаче №2 из раздела 3.4, Макаров А.А., Пашкевич А.В., Тамбовцева А.А. Задачник по математической статистике. 2018.
2. Задача 4. Идея основана на задаче №11 из раздела 10.2 того же задачника.