

Практика 6.

Задания по темам:

Общие принципы применения статистических критериев.

Параметрические и непараметрические критерии.

Уровень статистической значимости критерия.

Блок 1. Общие принципы применения статистических критериев

Задание 1 Гипотезы в исследовании

Для каждой ситуации сформулируйте нулевую (H_0) и альтернативную (H_1) гипотезы. Укажите, является ли альтернатива односторонней или двусторонней.

- (a) Психолог проверяет, влияет ли музыка на скорость решения теста (увеличивает или уменьшает).
- (b) Производитель утверждает, что среднее время работы батареек не менее 50 часов. Проверяющий хочет опровергнуть это утверждение.
- (c) Экономист полагает, что внедрение новой технологии увеличило производительность труда.

Задание 2 Ошибки первого и второго рода в реальной жизни

В аэропорту система безопасности проверяет багаж на наличие запрещённых предметов. H_0 : «багаж безопасен», H_1 : «в багаже есть запрещённый предмет».

- (a) Что является ошибкой первого рода (ложная тревога)?
- (b) Что является ошибкой второго рода (пропуск угрозы)?
- (c) Какая ошибка опаснее с точки зрения безопасности пассажиров? Почему?
- (d) Как изменив порог чувствительности детектора, можно уменьшить вероятность ошибки первого рода? Что при этом произойдёт с вероятностью ошибки второго рода?

Задание 3 Последовательность шагов

Расположите в правильном порядке (цифрами 1–6):

Вычисление наблюдаемого значения критерия
Формулировка статистических гипотез
Сравнение с критическим значением и вывод
Выбор уровня значимости α
Выбор подходящего критерия
Определение критической области

Задание 4 р-значение

При проверке гипотезы получено $p = 0.047$.

- (a) Отвергается ли H_0 на уровне $\alpha = 0.05$? Объясните.
- (b) А на уровне $\alpha = 0.01$? Почему?
- (c) Может ли быть такое, что $p = 0.047$, но исследователь говорит, что результат не значим? При каком условии?
- (d) Можно ли утверждать, что вероятность верности H_0 равна 4.7%? Почему нет?

Блок 2. Параметрические и непараметрические критерии

Задание 5 Заполните таблицу различий

Характеристика	Параметрический критерий	Непараметрический критерий
Предположение о распределении		
Тип данных (обычно)		
Примеры	t-критерий, _____	_____, критерий знаков
Мощность при выполнении условий		
Чувствительность к выбросам		

Задание 6 Выбор критерия

Для каждого пункта выберите: параметрический (П) или непараметрический (Н). Если можно использовать оба, укажите «оба», но кратко обоснуйте, какой предпочтительнее.

- (a) Сравнение среднего роста в двух выборках, распределение нормальное.
- (b) Сравнение оценок качества обслуживания (1–5) двух ресторанов, распределение не нормальное.

- (с) Анализ связи между полом и предпочтением бренда (5 категорий).
- (d) Сравнение дисперсий двух выборок.
- (е) Проверка нормальности распределения.

Задание 7 Робастность

Что значит, что непараметрический критерий является робастным (устойчивым) к выбросам? Приведите числовой пример: выборка $X = 2, 4, 6, 8, 100$. Сравните, как изменится среднее арифметическое и медиана при добавлении выброса (в исходной выборке без выброса значения 2, 4, 6, 8). Какой параметрический и какой непараметрический критерии пострадают сильнее?

Задание 8 Мощность критерия

Мощность критерия — это вероятность отвергнуть H_0 , когда она ложна. Что выше: мощность параметрического критерия при выполнении его предположений или мощность соответствующего непараметрического критерия? Почему? При каком условии непараметрический критерий может оказаться почти таким же мощным?

Блок 3. Уровень статистической значимости α

Задание 9 Выбор уровня значимости

В медицинском исследовании тестируется новое лекарство от редкого заболевания. Ошибка первого рода: признать лекарство эффективным, когда оно неэффективно (выпустить на рынок бесполезное лекарство). Ошибка второго рода: не признать эффективное лекарство (пациенты не получают лечение).

- (а) Какой уровень α вы выберете: 0.1, 0.05 или 0.01? Обоснуйте.
- (b) Что произойдёт с вероятностью ошибки второго рода при снижении α ?
- (с) Как можно одновременно снизить и α , и β ? (Подсказка: изменить объём выборки.)

Задание 10 Критическое значение в зависимости от α

Для тестовой статистики $Z \sim N(0, 1)$ (двусторонняя альтернатива). Критическое значение $z_{\alpha/2}$ определяется из условия $P(|Z| > z_{\alpha/2}) = \alpha$.

- (а) Найдите $z_{\alpha/2}$ для $\alpha = 0.10$, $\alpha = 0.05$, $\alpha = 0.01$.
- (b) Как меняется ширина критической области (значение $z_{\alpha/2}$) при уменьшении α ?
- (с) Нарисуйте схематично плотность $N(0, 1)$ и заштрихуйте критические области для $\alpha = 0.05$ и $\alpha = 0.01$.

Задание 11 Проблема множественных сравнений

Исследователь проверяет 20 независимых гипотез на уровне $\alpha = 0.05$. Все нулевые гипотезы верны.

- (а) Какова вероятность того, что хотя бы одна гипотеза будет отвергнута ошибочно? (Вычислите приближённо: $1 - (1 - 0.05)^{20}$.)

- (b) Назовите метод коррекции на множественные сравнения (например, Бонферрони). Как он работает?
- (c) Как изменится критический уровень для отдельного теста при поправке Бонферрони для 20 сравнений, чтобы общий α остался 0.05?

Комплексное задание

Задание 12 Применение на учебных данных

Педагог сравнивает две методики обучения (А и В) по результатам теста (баллы от 0 до 100). В каждой группе по 15 учеников. Распределение баллов не является нормальным, есть выбросы.

- (a) Какой параметрический критерий можно было бы использовать, если бы данные были нормальны?
- (b) Почему в данной ситуации лучше применить непараметрический критерий? Назовите конкретный непараметрический критерий для сравнения двух независимых выборок.
- (c) Если педагог получил $p = 0.04$ при использовании параметрического критерия и $p = 0.08$ при непараметрическом, какой вывод он должен сделать? Почему p -значения различаются?
- (d) Уровень значимости выбран $\alpha = 0.05$. Будет ли отвергнута H_0 в каждом из случаев? Какой критерий дал более надёжный результат?

Задание 13 Интерпретация результатов

Прочитайте фразу из научной статьи: «Различия между группами оказались статистически значимыми ($t(58)=2.34$, $p=0.023$)».

- (a) Что означают числа в скобках: $t(58)$?
- (b) Какой критерий использовался? Параметрический или нет?
- (c) Что можно сказать о вероятности ошибки первого рода в данном исследовании?
- (d) Можно ли утверждать, что вероятность того, что различия реальны, составляет 97.7%? Объясните.
- (e) Как бы изменилась интерпретация, если бы вместо t -критерия использовался U -критерий Манна–Уитни с тем же p ?

Задание 14 Проектирование эксперимента

Фирма тестирует два дизайна веб-сайта (старый и новый) для увеличения конверсии. Конверсия — бинарный признак (купил/не купил). Планируется показать каждый дизайн 1000 посетителей.

- (a) Сформулируйте H_0 и H_1 . Какой статистический критерий вы предложите (параметрический или непараметрический)? Почему?
- (b) Выберите уровень значимости α . Обоснуйте.
- (c) Если в результате теста $p = 0.03$, что вы скажете руководству фирмы? Какова практическая значимость результата, если конверсия увеличилась с 5% до 5.2%?

- (d) Какое дополнительное измерение (помимо p -значения) следует приводить для оценки размера эффекта?