ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ

Лекция 13

Проверка статистических гипотез

Оспан Ә.Ғ., PhD, доцент-исследователь

КазНУ, кафедра ИИ и Big Data

2026-2027 учебный год

Цель и содержание лекции

Цель и результаты обучения

- Формулировать H0, H1 и направление проверки до анализа
- Вычислять p-value и отличать его от вероятности H0
- Сообщать эффект, решение и ограничения дизайна

Содержание лекции

1. **План проверки**
2. **Статистика и хвосты**
3. **Ошибки и мощность**
4. **Средние и доли**
5. **Множественность и вывод**

Проверка начинается с вопроса и плана

1

Вопрос и единица

Что измеряем и какое наблюдение независимо?

2

H0, H1 и alpha

Направление и порог выбираем до просмотра результата.

3

Дизайн и критерий

Фиксируем отбор, метрику и правило остановки.

4

Данные и вывод

Вычисляем эффект, статистику и p; проверяем ограничения.

- Проверка оценивает совместимость данных с конкретной моделью.
- Поисковый и подтверждающий анализ выполняют разные роли.

Альтернатива определяет направление подозрения

Левосторонний вопрос

- H0: mu=14; H1: mu<14.
- Интересует возможное уменьшение срока службы.
- Экстремальные значения расположены в левом хвосте.

Двусторонний вопрос

- H0: mu=14; H1: mu!=14.
- Проверяется любое отличие от заявления.
- Экстремальность учитывается в обоих хвостах.

Выбор H1 определяется вопросом, а не знаком полученного эффекта.

Z-статистика измеряет отклонение от H0

$$z = (X_bar - mu0) / (sigma/sqrt(n))$$

Числитель

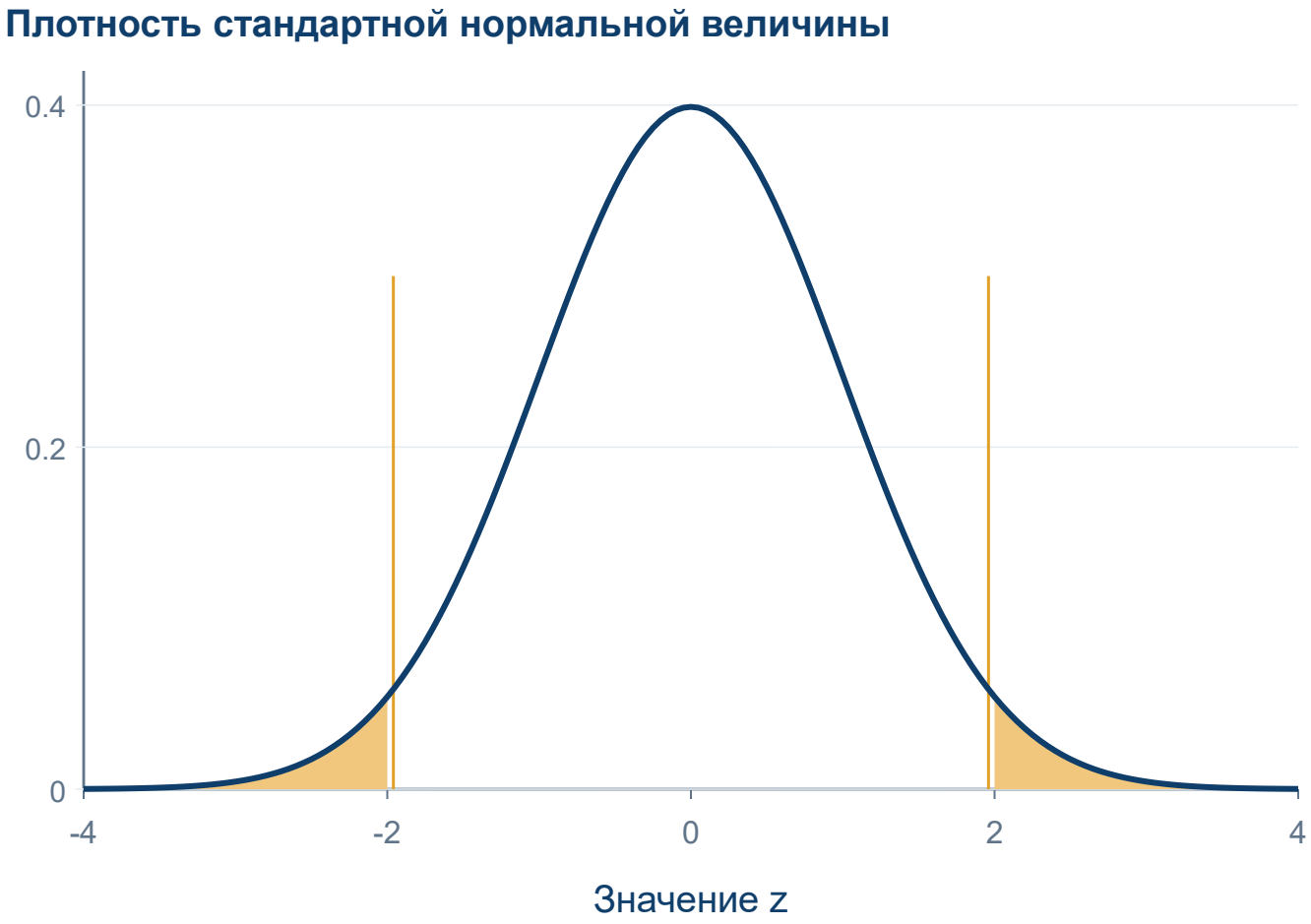
- Наблюдаемая разность между оценкой и H0.
- Знак показывает направление отличия.

Знаменатель

- Стандартная ошибка среднего при заданной модели.
- Требует известной sigma и обоснованной модели среднего.

X_bar=13.5, mu0=14, sigma=2, n=64: SE=0.25; z=-2.

P-value зависит от выбранного хвоста



Вероятность при нулевой модели

- Левая H1: p=norm.cdf(z).
- Правая H1: p=norm.sf(z).
- Двусторонняя H1: p=2*norm.sf(abs(z)).
- Площадь считается для не менее экстремальных результатов.

Условия H0 и выбранное направление должны предшествовать p-value.

Одна статистика - два заранее разных вопроса

Пример Python

```
import numpy as np
from scipy.stats import norm, ttest_1samp

z = (13.5 - 14) / (2 / np.sqrt(64))
p_left = norm.cdf(z)
p_two = 2 * norm.sf(abs(z))
print(z, round(p_left, 5), round(p_two, 5))
print('Левая H0 отклонена:', p_left < 0.05)
```

Результат: -2.0 0.02275 0.04550; левая H0 отклонена: True

Статистика

- Отклонение среднего равно -0.5 месяца.
- Стандартная ошибка 0.25 дает z=-2.

Решение

- p сравнивается с заранее выбранной alpha.
- Сравнение выполняется до округления.

Чего P-value не сообщает

Корректно

- Вероятность экстремальности данных при H0 и модели.
- p<alpha: H0 отклоняется указанным тестом.
- p>=alpha: оснований отклонить H0 недостаточно.

Некорректно

- «p - вероятность истинности H0».
- «Большое p доказывает точное равенство».
- «Малое p доказывает причинность или умысел».

Неотклонение H0 оставляет неопределенность о величине эффекта.

Ошибки двух родов и мощность

Реальное положение	H0 не отклонена	H0 отклонена
H0 верна	Корректное сохранение	Ошибка I рода
Конкретная H1 верна	Ошибка II рода	Обнаружение эффекта
Контроль процедуры	beta зависит от эффекта	alpha ограничивает ошибку I
Мощность	Не равна одному минус p	1-beta

Малое p, высокая мощность и большой эффект - разные понятия.

Неизвестная sigma: одновыборочная T-проверка

Пример Python

```
returns = np.array([
    6.1, 9.2, 11.5, 8.6, 12.1,
    3.9, 8.4, 10.1, 9.4, 8.9,
])
result = ttest_1samp(
    returns, popmean=10, alternative='less')
print('Среднее:', returns.mean())
print('t:', result.statistic, 'p:', result.pvalue)
print('Отклонить H0:', result.pvalue < 0.05)
```

Условие

- H0: mu=10; H1: mu<10.
- Независимая приближенно нормальная выборка.

Интерпретация

- Неизвестный разброс оценивается по выборке.
- Большое p не подтверждает точное равенство.

Содержательный результат включает эффект и неопределенность, а не только p.

Две группы: независимый или парный дизайн

Независимые группы

- Разные независимые объекты в A и B.
- Уэлч: equal_var=False; равные дисперсии не требуются.
- Эффект записывают как mean_A-mean_B.

Парные наблюдения

- Те же объекты измеряются до и после.
- Единица анализа - разность внутри пары.
- Повторные измерения не увеличивают число независимых людей.

Функция выбирается по структуре наблюдений, а не по удобству результата.

Проверка доли строит ошибку при H0

$$z = (p_hat - p0)/\sqrt{p0*(1-p0)/n}$$

Нулевая модель

- В SE используется p0, потому что предполагается H0.
- Проверяются ожидаемые n*p0 и n*(1-p0).
- Один объект дает один независимый бинарный исход.

Ограничения

- Редкие события: нормальное приближение может быть плохим.
- Повторы одного участника не независимы.
- Самовыбор ответивших не устраняется формулой.

k=87, n=125: p_hat=0.696; H0: p=0.75; H1: p<0.75.

Двадцать проверок меняют риск ложной находки

P(хотя бы одна ошибка) = 1-(1-alpha)m**

Независимые проверки при верных H0

- Семейный риск превышает уровень отдельного теста.
- Формула требует независимости тестов.

Контроль множественности

- Бонферрони: сравнивать p с alpha/m.
- m относится ко всем проверкам выбранного семейства.
- Предварительный план снижает риск удобного выбора.

Отчет об одной найденной значимости должен учитывать весь поиск.

Football: описание групп и условная модель

Что известно из файла

- Размеры групп GK и остальных строк.
- Средние Age и их наблюдаемая разность.
- Файл исторический: 12897 строк и 42 столбца.

Что требует дополнительных оснований

- Генеральная совокупность и механизм отбора.
- Независимость игроков с учетом клубов.
- Причинность и перенос на современные составы.

Учебное р при модели независимости не заменяет дизайн исследования.

Содержательный вывод шире статистической значимости

Статистика

- H0/H1, alpha, критерий, статистика и p.
- Решение сравнивают до округления.

Эффект

- Размер отличия и интервал в исходных единицах.
- Практический смысл задается предметной задачей.

Дизайн

- Отбор, независимость, остановка и множественность.
- Неотклонение не доказывает H0; значимость не доказывает причинность.

Главные результаты лекции

1. Направление H1, alpha и правило остановки задаются до проверки.
2. p-value описывает экстремальность статистики при H0 и модели.
3. p>=alpha означает недостаток оснований отклонить H0.
4. Критерий выбирают по дизайну: одна выборка, пары или независимые группы.
5. Эффект, интервал, множественность и ограничения входят в вывод.

Проверка понимания и самостоятельная работа

1. Как определить p-value без перестановки условия и события?
2. Почему p=0.2 не доказывает H0?
3. Когда вместо независимых групп анализируют парные разности?
4. Почему большое n не гарантирует практическую важность эффекта?
5. Что меняется при двадцати тестах вместо одного?

Самостоятельная работа

Для z=-1.92 найдите левое, правое и двустороннее p. До вычисления определите, какое отвечает заранее заданному вопросу о недовесе.

Документация и исходные материалы

Документация

1. SciPy: одновыборочная t-проверка

https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.ttest_1samp.html

2. SciPy: независимые группы и Уэлч

https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.ttest_ind.html

3. SciPy: точная биномиальная проверка

https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.binomtest.html

Основа лекции

- Лекция 13 Тестирование гипотез.pptx
- Лабораторная работа 13 Тестирование гипотез.docx
- Лабораторная работа 13 с ответами Тестирование гипотез.docx