

ОБЩИЕ ПРИНЦИПЫ ПРОВЕРКИ СТАТИСТИЧЕСКИХ ГИПОТЕЗ

Тема 4

РЕПОЗИТОРИЙ БГУ

**Статистическая гипотеза –
предположение о свойствах
случайных процессов и явлений**

РЕПОЗИТОРИЙ БГПУ

Принцип практической невозможности

- 📄 Случайное событие с очень малой вероятностью полагается практически невозможным.
- 📄 Если такое событие все-таки происходит, значит оно находится под влиянием некоторых неслучайных факторов

ОСНОВАНИЯ ДЛЯ ДОПУЩЕНИЯ ДЕЙСТВИЯ НЕСЛУЧАЙНОГО ФАКТОРА

- Если некоторое событие происходит случайно в 5 % случаев или еще реже, то предполагается, что это происходит благодаря действию некоторых неслучайных факторов
- Это значение называется 5 %-м уровнем статистической значимости или уровнем статистической значимости, равным $p=0,05$ ($\alpha=0,05$)

ОСНОВАНИЯ ДЛЯ ДОПУЩЕНИЯ ДЕЙСТВИЯ НЕСЛУЧАЙНОГО ФАКТОРА

- Если некоторое событие происходит случайно в 1 % случаев или еще реже, то предполагается, что это происходит благодаря действию некоторых неслучайных факторов
- Это значение называется 1 %-м уровнем статистической значимости или уровнем статистической значимости, равным $p=0,01$ ($\alpha=0,01$)

Статистические гипотезы

📄 **Нуль-гипотеза - это гипотеза об отсутствии различий.**

📄 **Она обычно обозначается H_0 и называется нулевой**

Например, девушки такие же умные, как и юноши; монетка правильная

Статистические гипотезы

- 📄 **Альтернативная гипотеза (гипотеза исследования, экспериментальная гипотеза, ...) - это гипотеза о значимости различий.**
- 📄 **Она обычно обозначается H_1 .**

Альтернативные гипотезы бывают

Направленные

девушки умнее, чем юноши;
орел выпадает чаще, чем решка

Ненаправленные

юноши и девушки отличаются по
интеллекту;
монетка неправильная

📄 Нельзя прямо доказать нуль-гипотезу

📄 Нельзя прямо доказать альтернативную гипотезу

📄 Однако, если мы можем отвергнуть нуль-гипотезу, то можем принять альтернативную ей

📄 Альтернативная гипотеза всегда подтверждается не прямо, а косвенно.

📄 Именно поэтому никогда не пишут, что «*гипотеза доказана*», а пишут «*гипотеза подтверждается*».

**Уровень статистической
значимости p представляет
собой вероятность
неправильного отвержения
нуль-гипотезы**

Уровень статистической
значимости p представляет
собой вероятность
допустимой ошибки при
проверке статистических
гипотез

Проверка гипотез осуществляется с помощью критериев статистической оценки различий

📄 Статистический критерий (критерий)
- это случайная величина, закон
распределения которой известен и
которая служит для проверки нуль-
гипотезы.

Статистический критерий можно рассматривать как инструмент, позволяющий определить вероятность того, что результаты получились случайно.

Если эта вероятность достаточно мала (например, $< 0,05$), то можно сделать вывод о том, что данные результаты получились неслучайно (т.е. отвергнуть нуль-гипотезу)

1. На уровне значимости $p=0,05$ нет оснований отвергнуть нуль-гипотезу. В этой ситуации нет смысла пытаться проверить статистические гипотезы на уровне значимости $p=0,01$, т.к. действия в условиях более низкой допустимой вероятности ошибки ничего не изменят.


2. На уровне значимости $p=0,05$ отвергается нуль-гипотеза и тем самым подтверждается в качестве верной альтернативная гипотеза для этого уровня значимости. Тогда имеет смысл проверить гипотезы на уровне значимости $p=0,01$. Если в результате такой проверки обнаруживается, что на уровне значимости $p=0,01$ нет оснований отвергнуть нуль-гипотезу, то в качестве окончательного принимается вывод: **отвергается нуль-гипотеза и подтверждается в качестве верной альтернативная гипотеза на уровне значимости $p=0,05$.**

3. Если и на уровне значимости $p=0,05$ и на уровне значимости $p=0,01$ отвергается нуль-гипотеза и подтверждается альтернативная гипотеза, то в качестве окончательного принимается вывод: **отвергается нуль-гипотеза и подтверждается в качестве верной альтернативная гипотеза на уровне значимости $p=0,01$.**

ПРОЦЕДУРА ПРОВЕРКИ СТАТИСТИЧЕСКИХ ГИПОТЕЗ

- С помощью статистического критерия определяется уровень статистической значимости p (вероятность того, что полученная в опыте разница случайна), он сравнивается с заранее выбранным порогом ошибки (например, $0,05$).
- Если $p > 0,05$, то нет оснований для отвержения нуль-гипотезы.
- Если $p \leq 0,05$, то можно отвергнуть нуль-гипотезу и сделать вывод о том, что альтернативная гипотеза подтвердилась

Ошибки I и II рода

 Пример: У нас в общежитии установлена противопожарная система, которая подает сигнал тревоги, когда концентрация дыма достигает определенного уровня.

Возможны 4 ситуации:

	Нет пожара	Пожар
Подает сигнал тревоги	Ошибка I рода	Нет ошибки
Нет сигнала тревоги	Нет ошибки	Ошибка II рода

Ошибка I рода

❏ Ошибка I рода -
сигнал без пожара,
например, когда вы
просто прижарили
вкусные тосты



Ошибка II рода

- ❏ Ошибка II рода - пожар без сигнала.
- ❏ Известно, как избежать ошибки I рода - отключить или сломать противопожарную сигнализацию. К несчастью, это приведет к увеличению возможности допустить ошибку II рода.

Точно так же и в статистике:

Решение	Гипотеза	
	Нуль-гипотеза верна	Исследовательская гипотеза верна
Отвержение нуль-гипотезы	Ошибка I рода	Нет ошибки
Принятие нуль- гипотезы	Нет ошибки	Ошибка II рода

Ошибка первого рода

Ошибка, состоящая в том, что мы отклонили нуль-гипотезу, в то время как она верна, называется ошибкой I рода.

Вероятность такой ошибки обозначается α (или p).

Ошибка второго рода

Ошибка, состоящая в том, что мы приняли нуль-гипотезу, в то время как она неверна, называется ошибкой II рода.

Вероятность такой ошибки обозначается β .

Мощность критерия - это его способность не допустить ошибку II рода. Поэтому

$$\text{мощность} = 1 - \beta.$$

Мощность критерия определяется эмпирическим путем.

Можем ли мы узнать каким-то образом, что совершаем ошибку I или II рода?



Увы! НИКАК!

КЛАССИФИКАЦИЯ МЕТОДОВ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

☞ Цель (вид решаемой задачи)

☞ Тип измерительной шкалы

☞ Количество выборок

☞ Виды выборок

☞ Закон распределения

КЛАССИФИКАЦИЯ МЕТОДОВ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ ПО ТИПУ РАСПРЕДЕЛЕНИЯ

- ▣ Параметрический – основан на конкретном типе распределения генеральной совокупности или использует параметры этой совокупности (средние, дисперсии)
- ▣ Непараметрический – не использует тип распределения генеральной совокупности и ее параметры

Выбор статистического метода

ЦЕЛЬ	ТИП ДАННЫХ		
	Шкала интервалов или отношений (распределение Гаусса)	Шкала порядка или негауссово распределение	Шкала наименований
Описание одной группы	Среднее, стандартное отклонение	Медиана	Мода
Определение зависимости между двумя переменными	Коэффициент корреляции Пирсона	Коэффициент корреляции Спирмена	Коэффициент сопряженности Крамера
Сравнение двух независимых выборок	Непарный t-критерий Стьюдента	U-критерий Манна-Уитни	Критерий χ^2 Пирсона
Сравнение двух зависимых выборок	Парный t-критерий Стьюдента	G-критерий знаков	M-критерий Макнамары