

Элементы статистики для построения гипотез A/B-тестирования

Построение гипотезы для A/B-тестирования

Цель урока

Понять, как сформулировать нулевую и альтернативную гипотезы в A/B-тестировании.

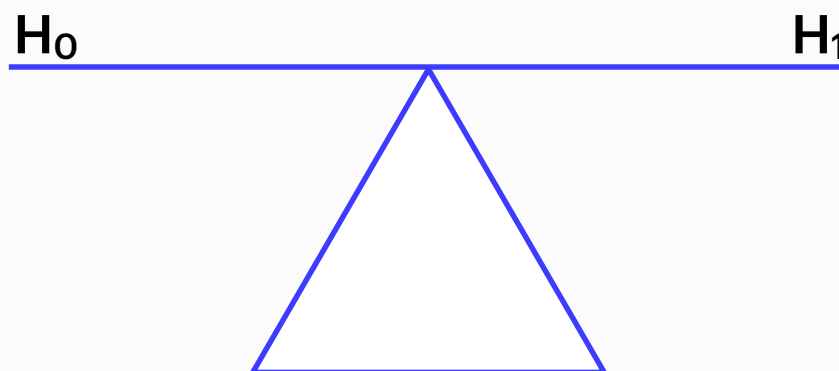
Задачи урока

- ✓ Познакомиться с понятиями нулевой и альтернативной гипотезы в A/B-тестировании
- ✓ Увидеть принцип формулировки гипотезы
- ✓ Разобраться, какие основные ошибки возникают при принятии гипотезы

Статистические гипотезы: нулевая и альтернативные

Нулевая гипотеза H_0 — это предположение о том, что изменения в тестовой группе В никак не повлияют на метрику продукта, а различия, которые мы можем наблюдать между контрольной группой А и тестовой группой В после эксперимента, случайны.

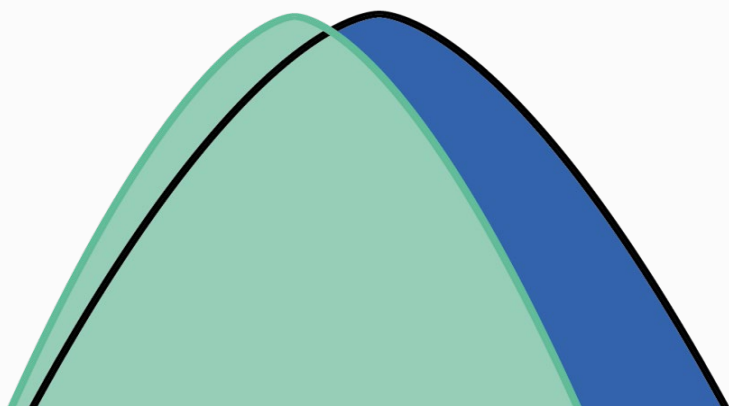
Альтернативная гипотеза H_1 отрицает H_0 , то есть изменения в тестовой группе повлияли на метрику продукта и различия между группами неслучайны.



Нулевая и альтернативные гипотезы

Нулевая гипотеза — H_0

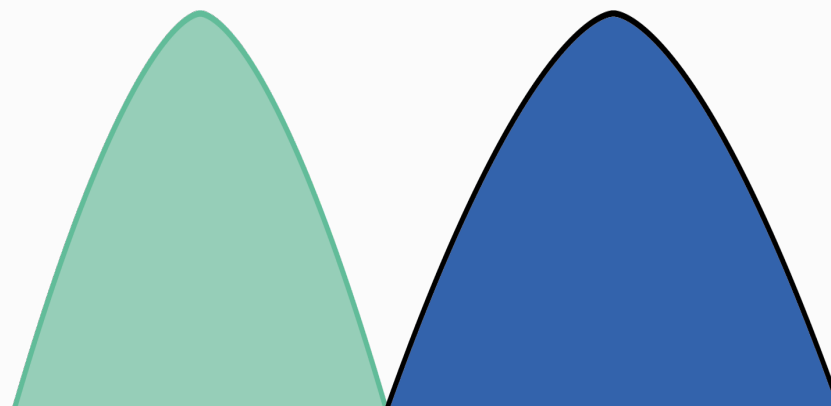
Пример: количество кликов **не** увеличится, если изменить цвет кнопки.



Группы **не** отличаются
Эффект случайный

Альтернативная гипотеза — H_1

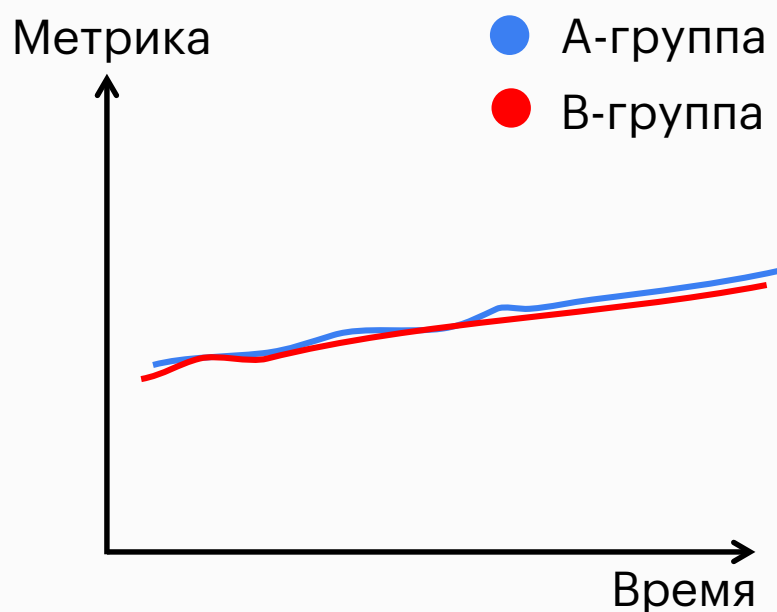
Пример: количество кликов увеличится, если изменить цвет кнопки.



Группы отличаются
Эффект неслучайный

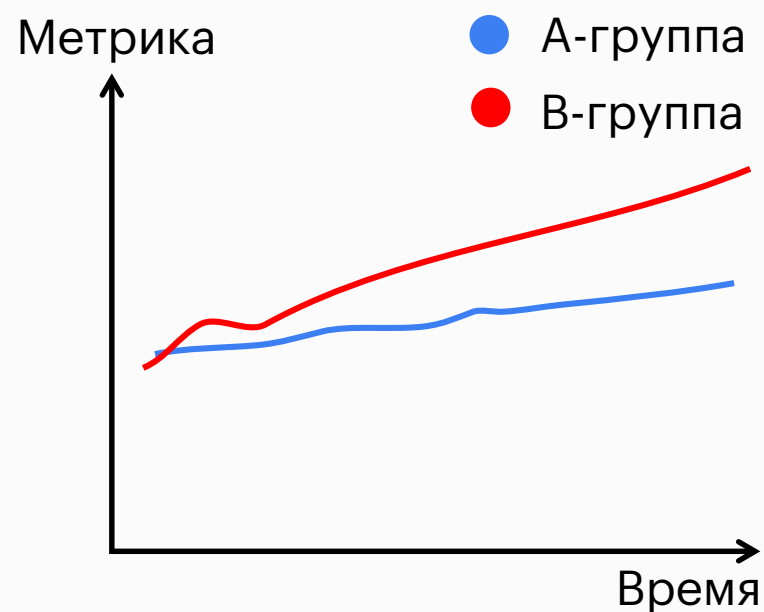
Нулевая и альтернативные гипотезы

Нулевая гипотеза — H_0



Видим, что **эффекта нет**
Графики не отличаются

Альтернативная гипотеза — H_1



Видим, что **эффект есть**
Графики отличаются

Статистическая значимость

Выдвигаем нулевую гипотезу H_0 : различие между выборками недостаточно значимо, и наше изменение не оказывает никакого влияния.

Альтернативная гипотеза H_1 : различие значимо.

Принцип: пытаемся опровергнуть нулевую гипотезу и, тем самым, подтвердить альтернативную.

Нулевая гипотеза считается верной, пока не доказано обратное.

Нельзя на 100 % принять нулевую гипотезу, можно сказать, что данные согласуются с нулевой гипотезой.

Ошибки I и II рода

		Верна ли H_0 на самом деле?	
		Верна	Не верна
Принимаем ли мы H_0 (в результате теста)?	Принимается	H_0 верно принята	Ошибка II рода
	Отвергается	Ошибка I рода	H_0 верно отвергнута

Ошибка I рода (обнаружить различия там, где их на самом деле нет) жёстко ограничивается и критичнее ошибки II рода.

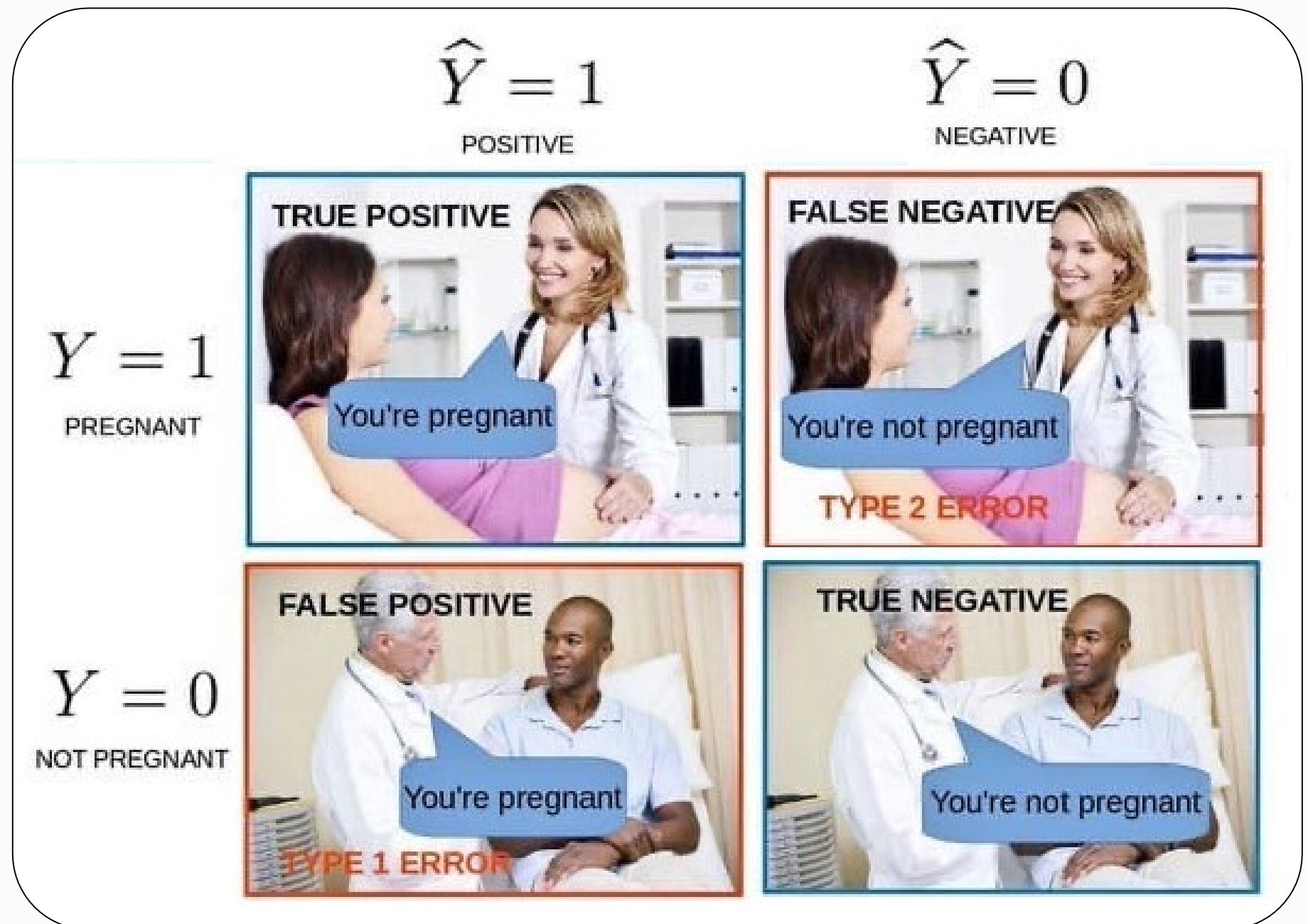
Вероятность ошибки I рода обозначают как **p-value**.

Пороговую вероятность ошибки I рода обозначают как α .

Ошибка II рода (не обнаружить различия или связи, которые на самом деле есть) минимизируется.

Вероятность ошибки II рода обозначают как β .

Построение гипотезы для A/B-тестирования



Ошибки I и II рода

Вероятность ошибки I рода обозначают как **p-value**.

Пороговую вероятность ошибки I рода обозначают как α .

Вероятность ошибки II рода обозначают как β .

Мощность статистического критерия ($1 - \beta$).

Если увеличим выборку, то мощность возрастает.

		Истина	
		H_0	H_1
Принята	H_0	$1 - \alpha$	β
	H_1	α	$1 - \beta$

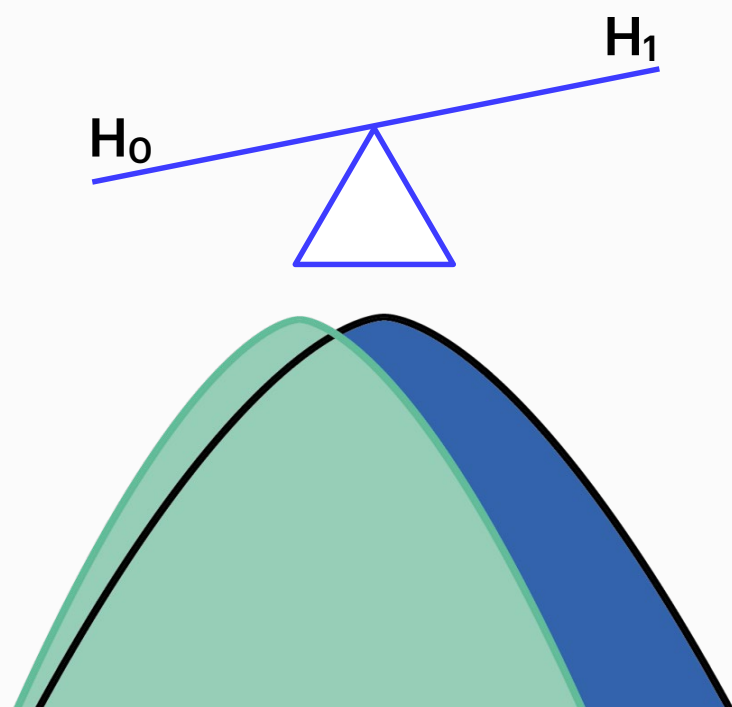
Важность ошибки I рода и p-value

Вероятность ошибки I рода обозначают как **p-value**.

$P(\text{ошибка I рода}) = \text{p-value}$.

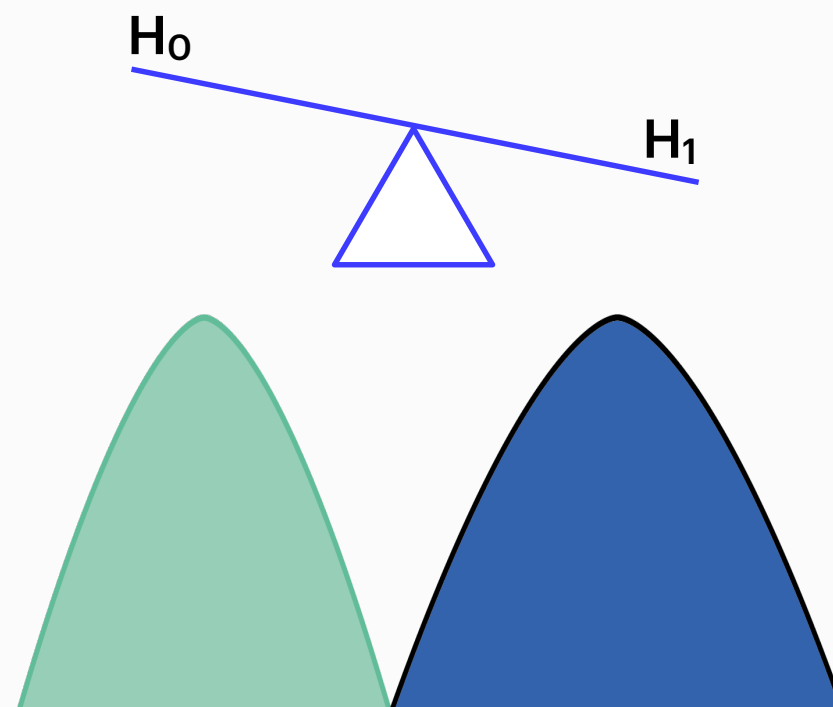
Пороговую вероятность ошибки I рода обозначают как α .

Важность ошибки I рода и p-value



Группы не отличаются
Эффект случайный

$$p\text{-value} > \alpha$$



Группы отличаются
Эффект неслучайный

$$p\text{-value} < \alpha$$

Как правило, за α берётся 5 %

Итоги урока



При построении гипотезы есть нулевая и альтернативная гипотезы



В результате проверки гипотез могут возникать ошибки I и II рода



Ошибкой I рода является p-value