

Skillbox

АВ-тестирование: часть 1

Алла Тамбовцева

Введение в АВ- тестирование

Что такое АВ-тестирование?

Какой дизайн сайта лучше?

- Старый
- Новый

Как понять? Эксперимент!

Что такое АВ-тестирование?

АВ-тестирование

Метод маркетингового исследования, который используется для оценки эффективности каких-либо изменений, а именно изменений *целевых показателей* – тех характеристик продукта, которые нас интересуют.

Примеры АВ-тестирования

- Интернет-маркетинг: ведет ли изменение дизайна сайта к большей конверсии?
- Обычный маркетинг: ведет ли изменение дизайна упаковки к увеличению процента желающих его купить?

Дизайн АВ-тестирования

Как отобрать пользователей для тестирования?

Простой случайный выбор:

случайным образом распределяем пользователей по страницам.

Дизайн АВ-тестирования

Как поделить пользователей на группы?

- **Контрольная группа:** группа, которая видит сайт с текущим дизайном (до изменений)
- **Тестовая группа:** группа, которая видит сайт с новым дизайном (после изменений)

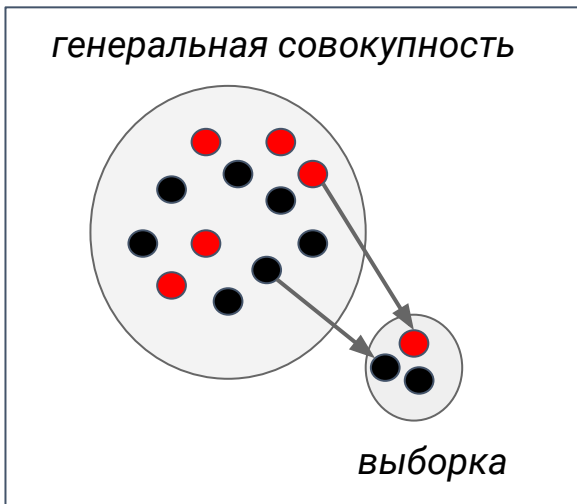
Что дальше?

Научиться правильно понимать разницу между этими двумя группами

Что такое выборочное обследование?

Генеральная совокупность vs выборка

- **Генеральная совокупность** – все объекты интереса.
- **Выборка** – объекты интереса, которые непосредственно обследуем.



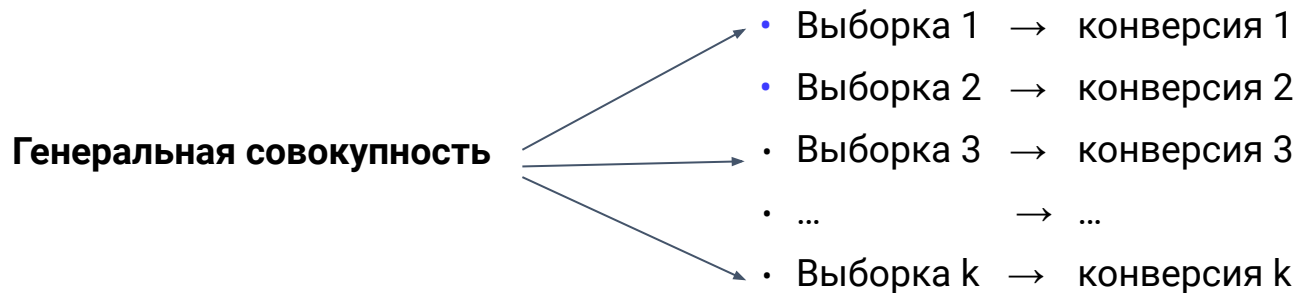
Как понять, хорошая ли выборка?

- Оценить **репрезентативность выборки**: хорошо ли выборка отражает свойства генеральной совокупности?
- Оценить **размер выборки**: достаточно ли выборка большая?

Статистические законы, которые обеспечивают точность результатов при изучении случайной выборки вместо всей генеральной совокупности, начинают работать с объема выборки $n > 30$.

Выбор объема выборки для АВ-тестирования

Вариативность доли



Поэтому рассмотрим не единственное значение доли, а интервал, в котором она может лежать.

Как построить такой интервал?

- В силу ограниченности ресурсов не сможем повторить эксперимент (тестирование) много-много раз и сообщить, в каком интервале лежит доля. Что делать?
- Самостоятельно до эксперимента зафиксировать степень уверенности в результатах.

Уровень доверия: мера надежности статистического вывода.

Доверительный интервал: интервал, в котором с определенной долей уверенности может лежать интересующий нас показатель.

Как построить такой интервал?

Степень уверенности учли. Что дальше? Подумать о погрешности!

$$\text{conversion} \pm \text{error} = \text{conversion} \pm 3\%$$

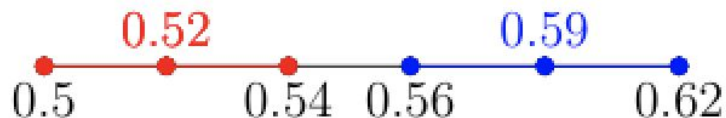
Какой объем выборки взять?

$$n = \frac{z^2 p(1 - p)}{error^2}$$

p – ожидаемое значение доли (которое ожидаем еще до проведения тестирования),
error – максимальная погрешность (ошибка), которую мы хотим допускать при получении результатов,
z – значение стандартной нормальной величины, которое соответствует выбранному уровню доверия.

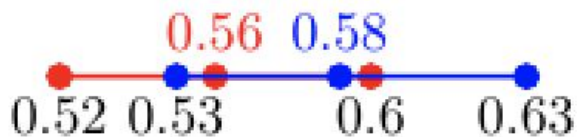
Доверительные интервалы

Что дают доверительные интервалы?

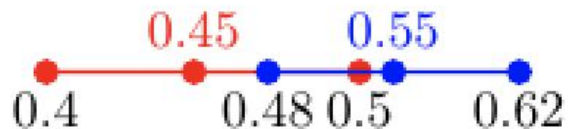


интервалы не пересекаются → разница есть

Что дают доверительные интервалы?



интервалы пересекаются → разницы нет



интервалы пересекаются → разница есть