

**Урок 6**

**Расчет длительности теста и проверка гипотез**

# На этом уроке

-Разберем как можно рассчитывать сколько нужно проводить эксперимент

-Как проверять ваши данные на нормальность

-Нюансы множественных сравнений

-Что такое бутстрап и когда его можно и нужно применять

# Оглавление

[На этом уроке](#_bwlphkh9hbyk)

[Оглавление](#_omlmxk18y8e1)

[Базовые понятия статистики](#_bgagsep32m02)

[Определение необходимой численности выборки](#_otqex5bzsvc1)

[Определение численности выборки для доли](#_p698ljwv31yu)

[Определение численности выборки для среднего](#_bg5vb2w64c6f)

[Дополнительные материалы](#_6pp2pbixtce)

[Используемые источники](#_th507mlvtdq1)

# Базовые понятия статистики

На предыдущем занятии мы разбирали следующие понятия:

**Нулевая гипотеза** – принимаемое предположение о том, что не существует связи между наблюдениями в двух (или более) событиях (выборках, феноменах, совокупностях). Гипотезу отвергают, если данные показывают разницу между выборками.

**True Positive** = говорим истина, когда по факту истина (факт)True Negative = говорим не истина, когда по факту тоже не истина (факт)

**False Positive (ошибка I рода) =** говорим истина, когда по факту не истина. Отклонение верной нулевой гипотезы. Риск совершить такую ошибку равен выбранному уровню статистической значимости (например, α=0.05) (ошиблись)

**False Negative (ошибка II рода) =** говорим не истина, когда по факту истина (ошиблись). Принятие неверной нулевой гипотезы. Вероятность ошибочно сохранить неверную нулевую гипотезу обозначают буквой β (например, 0.2).

**Мощность критерия/теста** - это вероятность, что тест правильно засечёт эффект там, где он и правда есть. (т.е. 1-β) .

Введем несколько новых:

**Зависимые выборки** – это те выборки, в которых каждому респонденту одной выборки поставлен в соответствие по определенному признаку респондент другой выборки.   
Пример: пользователь уже был в эксперименте X, и на нем же проводим эксперимент Y.

**Независимые выборки** – это те выборки, в которых вероятность отбора любого респондента одной выборки не зависит от отбора любого из респондентов другой выборки.  
Пример: рандомно взятые новые пользователи

Эти понятия нам пригодятся в дальнейшем, важно знать их и понимать их суть.

Мы в продуктах не хотим принимать ложноотрицательные(отклонять крутые гипотезы) и ложноположительные(раскатывать плохие гипотезы) решения и хотим минимизировать ошибку первого и второго рода .Выбираем alpha и ищем самый мощный критерий среди корректных.

Мощность и точность зависят от размера выборки - чем большую мощность и точность мы хотим - тем больше выборка нам нужна - поговорим об этом в следующей главе. Вам сейчас важно понять суть различных ошибок - поскольку в дальнейшем они будут использоваться регулярно.

# Определение необходимой численности выборки

Расчет численности выборки очень важный этап планирования A/B-тестов, который может с:

1. Понимать, сколько нужно юзеров для проведения эксперимента

2. Понимать сколько продлится эксперимент

3. Приоритезировать эксперименты

## Определение численности выборки для доли

Предположим, что вы предложили запустить A/B-тест, в котором в версии (B) будет другая форма подтверждения заказа. Нынешняя конверсия сайта = 1% Мы предполагаем, что в новой версии конверсия вырастет до 1,5% Сколько юзеров нужно отправить на каждую из версий? Воспользуемся [калькулятором](https://mindbox.ru/academy/ab-test-calculator/)

Теперь мы знаем, что для того, чтобы получить статистически значимый результат при заданном уровнем конверсии на обеих выборках, нам нужно направить на каждый вариант теста (A и B) по 6 216 юзеров.

Пример :Предположим, что вы предложили запустить A/B-тест, в котором в версии (B) будет другая форма подтверждения заказа. Нынешняя конверсия сайта = 0,7% Мы предполагаем, что в новой версии конверсия вырастет до 1,1% Сколько юзеров нужно отправить на каждую из версий, если мы хотим провести тест с уровнем доверия 99%? Воспользуемся калькулятором

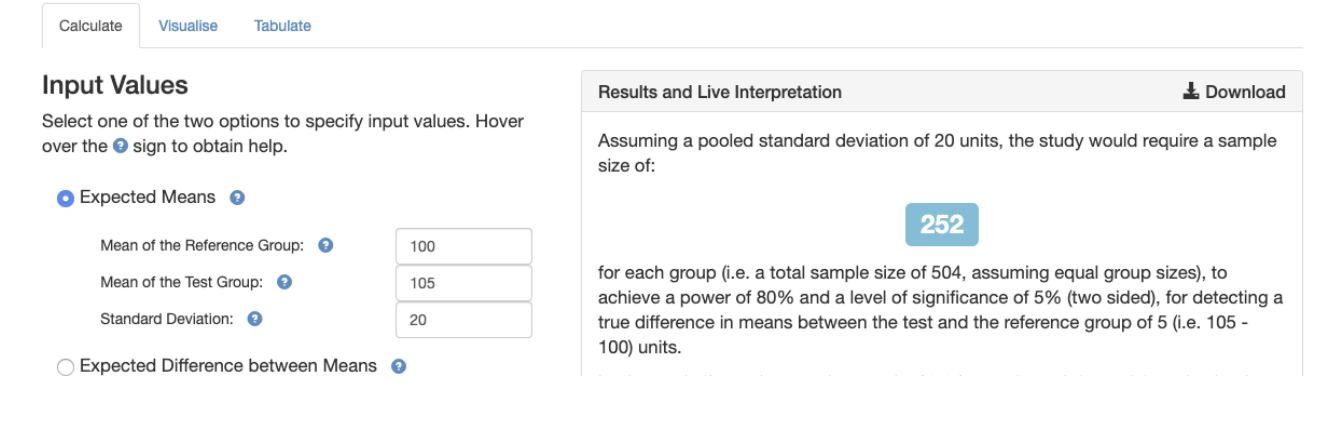
Это задание для студентов. Ответ: 10 147

## Определение численности выборки для среднего

Предположим, что ARPU нынешней версии продукта равен 100. SD (стандартное отклонение) от ARPU равен 20

Мы предполагаем, что в новой версии продукта ARPU будет равна 105. Необходимо рассчитать численность выборки для этого теста

Воспользуемся калькулятором [Sample Size Calculator for Comparing Two Independent Means](https://statulator.com/SampleSize/ss2M.html)



Теперь мы знаем, что для того, чтобы получить статистически значимый результат при заданном уровнем конверсии на обеих выборках, нам нужно направить на каждый вариант теста (A и B) по 252 юзера.

Пример Предположим, что LTV нынешней версии продукта равен 1000. SD (стандартное отклонение) от LTV равен 200 Мы предполагаем, что в новой версии продукта LTV будет равен 1050. Необходимо рассчитать численность выборки для этого теста с помощью [калькулятора](https://statulator.com/SampleSize/ss2M.html). Это задание для студентов.

Ответ: 252 Теперь мы умеем рассчитывать необходимую численность выборки для получения статистически значимых результатов. Не забывайте всегда анализировать, какой тип метрики мы сравниваем: метрику о доли или метрику о среднем.

# Дополнительные материалы

1. [Видео про описательную статистику](https://www.youtube.com/watch?v=4SEX9Bozktw)
2. [Книга Голая статистика](https://www.livelib.ru/book/1001518877-golaya-statistika-samaya-interesnaya-kniga-o-samoj-skuchnoj-nauke-charlz-uilan)
3. [Книга статистика и котики](https://www.livelib.ru/book/1002721779-statistika-i-kotiki-vladimir-savelev)
4. [Книга Statistics in Nutshell](https://www.oreilly.com/library/view/statistics-in-a/9781449361129/)
5. [Базовые концепты статистики](https://medium.com/nuances-of-programming/8-%D0%B1%D0%B0%D0%B7%D0%BE%D0%B2%D1%8B%D1%85-%D0%BF%D0%BE%D0%BD%D1%8F%D1%82%D0%B8%D0%B9-%D1%81%D1%82%D0%B0%D1%82%D0%B8%D1%81%D1%82%D0%B8%D0%BA%D0%B8-%D0%B4%D0%BB%D1%8F-%D0%BD%D0%B0%D1%83%D0%BA%D0%B8-%D0%BE-%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85-14a1ed2e31ae)
6. [Хорошее описание ЦПТ](https://www.marketing.spb.ru/lib-around/stat/Naked_Statistics.htm)
7. [Введение в основы статистики от Анатолия Карпова](https://www.youtube.com/watch?v=9jL5JifQ1WI)

8 ) [Z score и p-value](https://medium.com/@ashishtuteja87/what-is-z-score-and-p-value-b7daf7db008f)

9) [Лекции ВШЭ по проверке гипотез](http://math-info.hse.ru/f/2017-18/ps-ms/hypo-test.pdf)

10) [Про интерпретацию p-value на русском](https://habr.com/ru/company/stepic/blog/250527/)

11) [Про интерпретацию p-value на английском](http://www.perfendo.org/docs/BayesProbability/twelvePvaluemisconceptions.pdf)

12) [Про выбор критериев](https://lit-review.ru/biostatistika/vybor-statisticheskogo-kriteriya/)

13)[Лекции МШЭ МГУ по проверке гипотез](https://mse.msu.ru/wp-content/uploads/2020/03/%D0%9B%D0%B5%D0%BA%D1%86%D0%B8%D1%8F-6-%D0%9F%D1%80%D0%BE%D0%B2%D0%B5%D1%80%D0%BA%D0%B0-%D0%B3%D0%B8%D0%BF%D0%BE%D1%82%D0%B5%D0%B7.pdf)

14) [Для желающих почитать про непараметрику](https://mse.msu.ru/wp-content/uploads/2020/04/%D0%9B%D0%B5%D0%BA%D1%86%D0%B8%D1%8F-9-%D0%BD%D0%B5%D0%BF%D0%B0%D1%80%D0%B0%D0%BC%D0%B5%D1%82%D1%80%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B5-%D0%BC%D0%B5%D1%82%D0%BE%D0%B4%D1%8B.pdf)

15) [Про доверительные интервалы от Демешева](https://www.youtube.com/watch?v=GoK7sod8dhU)

16) [Про доверительные интервалы от Кирилла Мильчакова](https://www.youtube.com/watch?v=MstzroncW28)

17) [Про MDE и ошибки 1 и 2 типа - вводная статья](https://medium.com/@PAPP/significant-power-d9c3bc3dd3c)

18) [Про MDE посложнее](https://rpubs.com/hedbergec/sirc)

# Используемые источники

1. [Видео про описательную статистику](https://www.youtube.com/watch?v=4SEX9Bozktw)
2. [Книга Голая статистика](https://www.livelib.ru/book/1001518877-golaya-statistika-samaya-interesnaya-kniga-o-samoj-skuchnoj-nauke-charlz-uilan)
3. [Книга статистика и котики](https://www.livelib.ru/book/1002721779-statistika-i-kotiki-vladimir-savelev)
4. [Книга Statistics in Nutshell](https://www.oreilly.com/library/view/statistics-in-a/9781449361129/)
5. [Базовые концепты статистики](https://medium.com/nuances-of-programming/8-%D0%B1%D0%B0%D0%B7%D0%BE%D0%B2%D1%8B%D1%85-%D0%BF%D0%BE%D0%BD%D1%8F%D1%82%D0%B8%D0%B9-%D1%81%D1%82%D0%B0%D1%82%D0%B8%D1%81%D1%82%D0%B8%D0%BA%D0%B8-%D0%B4%D0%BB%D1%8F-%D0%BD%D0%B0%D1%83%D0%BA%D0%B8-%D0%BE-%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85-14a1ed2e31ae)
6. [Хорошее описание ЦПТ](https://www.marketing.spb.ru/lib-around/stat/Naked_Statistics.htm)
7. [Введение в основы статистики от Анатолия Карпова](https://www.youtube.com/watch?v=9jL5JifQ1WI)

8 ) [Z score и p-value](https://medium.com/@ashishtuteja87/what-is-z-score-and-p-value-b7daf7db008f)

9) [Лекции ВШЭ по проверке гипотез](http://math-info.hse.ru/f/2017-18/ps-ms/hypo-test.pdf)

10) [Про интерпретацию p-value на русском](https://habr.com/ru/company/stepic/blog/250527/)

11) [Про интерпретацию p-value на английском](http://www.perfendo.org/docs/BayesProbability/twelvePvaluemisconceptions.pdf)

12) [Про выбор критериев](https://lit-review.ru/biostatistika/vybor-statisticheskogo-kriteriya/)

13)[Лекции МШЭ МГУ по проверке гипотез](https://mse.msu.ru/wp-content/uploads/2020/03/%D0%9B%D0%B5%D0%BA%D1%86%D0%B8%D1%8F-6-%D0%9F%D1%80%D0%BE%D0%B2%D0%B5%D1%80%D0%BA%D0%B0-%D0%B3%D0%B8%D0%BF%D0%BE%D1%82%D0%B5%D0%B7.pdf)

14) [Для желающих почитать про непараметрику](https://mse.msu.ru/wp-content/uploads/2020/04/%D0%9B%D0%B5%D0%BA%D1%86%D0%B8%D1%8F-9-%D0%BD%D0%B5%D0%BF%D0%B0%D1%80%D0%B0%D0%BC%D0%B5%D1%82%D1%80%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B5-%D0%BC%D0%B5%D1%82%D0%BE%D0%B4%D1%8B.pdf)

15) [Про доверительные интервалы от Демешева](https://www.youtube.com/watch?v=GoK7sod8dhU)

16) [Про доверительные интервалы от Кирилла Мильчакова](https://www.youtube.com/watch?v=MstzroncW28)

17) [Про MDE и ошибки 1 и 2 типа - вводная статья](https://medium.com/@PAPP/significant-power-d9c3bc3dd3c)

18) [Про MDE посложнее](https://rpubs.com/hedbergec/sirc)

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 