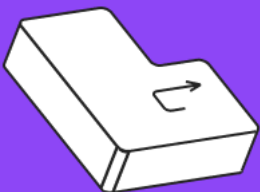


# Введение в статистику

Курс “Введение в  
продуктовую аналитику”





# Оглавление

Введение	2
Термины, используемые в лекции	2
Что такое статистика?	3
Важность статистики	3
A/B тестирование	4
Что нужно для проведения A/B тестирования	4
Типы переменных	5
Формирование гипотез для A/B теста и типы ошибок	7
Framework формирования гипотез	8
Типы ошибок	8
Дизайн эксперимента	10
Порядок проведения A/B теста	11
Пример: A/B-тест для использования поиска	12
Ключевые статистические термины	12
Статистическая значимость	13
Доверительный интервал —	14
Выборка и генеральная совокупность	14
Дополнительные материалы	14

# Введение

На этой лекции вы найдете ответы на такие вопросы как / узнаете:

- Что нужно для проведения A/B тестирования
- Формирование гипотез
- Определение необходимой численности выборки
- Алгоритм проверки гипотез



## Термины, используемые в лекции

- Ошибка первого рода (False Positive) - отклонение нулевой гипотезы, хотя гипотеза верная
- Ошибка второго рода (False Negative) - принятие нулевой гипотезы, хотя верна альтернативная
- ARPU (Average Revenue Per User) - средний доход с одного юзера
- LTV (Lifetime Value) - доход, который получает продукт за всё время работы с клиентом
- t-критерий Стьюдента – общее название для класса методов статистической проверки гипотез (статистических критериев), основанных на распределении Стьюдента. Наиболее частые случаи применения t-критерия связаны с проверкой равенства средних значений в двух выборках
- Bounce rate (показатель отказов) - показывает процент пользователей, которые сразу же ушли с сайта при посещении первой страницы
- Стандартное отклонение (среднеквадратическое отклонение) - в теории вероятностей и статистике наиболее распространенный показатель рассеивания значений случайной величины относительно её математического ожидания.

## Что такое статистика?

### Важность статистики

Математическая статистика - это раздел математики, в котором разрабатываются различные методы для описания и анализа наблюдений с целью использования для научных и практических выводов.

Мы будем изучать статистику через призму A/B тестов.

Математическая статистика - фундамент A/B тестов, без правильного понимания которого резко возрастает риск принятия неверных решений в продукте. И в этом мы ни раз убедимся в рамках курса.

## A/B тестирование

A/B тестирование - это метод исследования, где сравнивается 2 разных варианта, не обязательно для онлайн.



A/B-тестирование или иначе сплит-тестирование — метод маркетингового исследования, в котором:

1. Пользователей случайным образом делят на 2 группы
2. Группа 1 направляется на страницу "А" (существующую)
3. Группа 2 направляется на страницу "Б" (измененную)
4. Сравниваются метрики по обеим группам



## Что нужно для проведения А/В тестирования




1. Уметь рассчитывать размер выборки для теста
2. Понимать, что означает мощность теста
3. Понимать, насколько страшны ошибки I и II рода
4. Понимать, что означает p-value и доверительный интервал
5. Знать основные статистические критерии
6. Уметь корректно посчитать результаты теста

## Типы переменных

### 1. Количественные

- Непрерывные — это переменная, которая может принимать любое значение, например, вес испытуемого. Допустим, вес может принимать любое значение от 50 до 120 кг.
- Дискретные — это переменная, которая может принимать не любое значение, самый распространенный пример, когда переменная может принимать только целые числа, например, количество покупок на одного юзера. Количество покупок не может равняться 2.5.

Примеры дискретной случайной величины:

		
<b>Пример 1</b>	<b>Пример 2</b>	<b>Пример 3</b>
Число учащихся по школам	В результате 100-кратного подбрасывания орел выпал 50 или 51 раз	Число метеоритов, упавших на землю за год
X: 375, 869 ... 687	X: 50, 51, ..., 78	X: 0, 1 ... 1

- число учащихся по школам;



- какое количество раз выпал орел в результате 100-кратного подбрасывания монеты;
- число метеоритов, упавших на землю за год

**Задание на закрепление материала:**

Давайте закрепим тему с дискретными показателями, которую изучали на курсе “Работа с гипотезами и продуктовые исследования” и повторили на этом курсе

## Какие переменные являются дискретными?

1. Число слушателей, присутствующих на семинаре
2. Пол человека
3. Среднее количество просмотров этого видео
4. Имена ваших родственников
5. Число детей в семье
6. Средний доход продуктовых менеджеров



Напишите свои ответ в комментариях под уроком.  
Время на размышление -2 минуты



Давайте закрепим тему с дискретными показателями, которую изучали на курсе “Работа с гипотезами и продуктовые исследования” и повторили на этом курсе

## Какие переменные являются дискретными?

1. Число слушателей, присутствующих на семинаре
2. Пол человека
3. Среднее количество просмотров этого видео
4. Имена ваших родственников
5. Число детей в семье
6. Средний доход продуктовых менеджеров



2. Качественные (номинативные)

3. Ранговые



# Формирование гипотез для А/В теста и типы ошибок

## Framework формирования гипотез

<b>IF</b> если	<b>IF:</b> Если упростить процесс подачи заявки
<b>BY</b> посредством (при помощи чего)	<b>BY:</b> при помощи ликвидации обязательных полей,
<b>WILL</b> то это приведет к	<b>WILL:</b> то это увеличит проходимость людей по форме,
<b>BECAUSE</b> потому что	<b>BECAUSE:</b> так как людям будет проще

Действие [X] позволит увеличить метрику [Y] на величину [N], потому что [Z].

Нулевая гипотеза (H0) — это проверяемое предположение, которое обычно формулируется как отсутствие различий, отсутствие влияние фактора, отсутствие эффекта, равенство нулю значений выборочных характеристик и т.п.

В противовес к ней рассматривают альтернативную гипотезу (H1).

## Типы ошибок

- 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,  $\alpha=0.05$ ) (ошиблись)





- False Negative (ошибка II рода) = говорим не истина, когда по факту истина (ошиблись). Принятие неверной нулевой гипотезы. Вероятность отклонить реально работающее изменение

### **Задание на закрепление материала:**

Сформулируйте гипотезы под ваш продукт либо для продукта онлайн кинотеатра, которую можно подвергнуть А/Б тестированию?



Пример гипотезы для онлайн-кинотеатра:

Гипотеза:

- если мы будем давать уникальный оффер, когда пользователь отказывается от подписки
- при помощи отдельной страницы с описанием преимуществ предложения,
- то это снизит количество отписок на \_\_%,



- так как пользователь поймет, что мы все еще можем удовлетворить его потребности.

Пример таких оферов:

- тем, кто не заходил на платформу за последний месяц делать скидку, которая актуальна только при условии автопродления подписки
- тем, кто заходил, но проводил мало времени и не находил, что посмотреть, даем месяц бесплатной расширенной подписки на сериалы или эксклюзивный контент, чтобы он понял, что ему есть что посмотреть там

## Дизайн эксперимента

Дизайн (планирование) эксперимента (experimental design techniques) — комплекс методов математической статистики, направленных на постановку опытов и проведение рациональных измерений, подверженных случайным ошибкам.

Дизайн эксперимента можно разбить на пять этапов.

1. Этап 1 — гипотеза

**IF**  
если

**Действие [X]**

**BY**  
посредством (при помощи чего)

**позволит увеличить метрику [Y]**

**WILL**  
то это приведет к

**на величину [N],**

**BECAUSE**  
потому что

**потому что [Z].**



2. Этап 2 - Что изменяем в продукте

- Какие делаем изменения
- На каком этапе юзеры попадают в эксперимент

3. Этап 3 — Определение аудитории эксперимента

- Определение генеральной совокупности
- Определение выборочной совокупности
- Какой процент войдёт в эксперимент

4. Этап 4 - Определение метрик

- Какие метрики будем использовать
- Какой эффект на метрики мы ожидаем
- Сколько нужно проводить тест

5. Этап 5 - План действий.

Что делаем, когда получим определённый результат по определённым в тестировании метрикам

## Порядок проведения А/В теста

- 1) Сформировать гипотезу
- 2) Основание (не обязательно) — подробно расписываем, почему мы считаем, что ключевая метрика вырастет
- 3) Определить, какие изменения делаем в продукте
- 4) Определить, на каких аудиториях тестируем
- 5) Определить ключевые метрики



6) Ожидаемый эффект

7) План действий

## Пример: А/В-тест для использования поиска

Гипотеза	Если перенести поиск в меню, то общий ретенш повысится за счёт снижения количества человек, которые его использовали.
Основание	Люди использующие поиск, показывают более низкий процент удержания по сравнению с людьми, которые его не использовали. Снизив их количество, поднимем общий ретеншн.
Что делаем в продукте	Контрольная версия: оставляем всё как есть. Тестовая версия: переносим поиск в меню.
На каких пользователях тестируем	Все новые пользователи. Выборка 50/50.
Метрики	Retention, число просмотренных страниц, время сессии.
План действий	<ul style="list-style-type: none"> <li>• Если Retention и другие метрики вырастут, то раскатим В версию на всех;</li> <li>• Если Retention вырастет, а другие метрики упадут, то раскатим В версию на всех;</li> <li>• Если Retention значимо не изменится, а другие метрики упадут, то оставим А версию;</li> <li>• Если Retention и другие метрики значимо не изменятся, то раскатим В версию на всех;</li> <li>• Если Retention снизится, а другие метрики вырастут, то оставим А версию.</li> </ul>

# Ключевые статистические термины

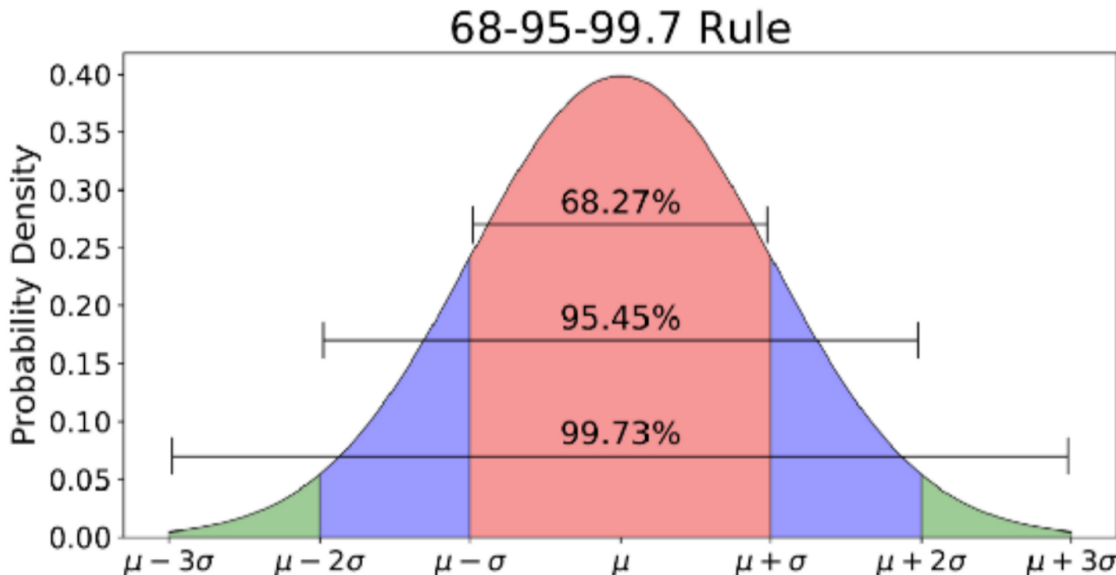
## Статистическая значимость

Статистическая значимость — это оцененная мера уверенности в том, что полученный результат не является случайным

Суть статистической значимости состоит в определении того, существует ли реальное основание в разнице между выбранными для исследования показателями, или это случайность? С данным понятием тесно связаны «нулевая» и «альтернативная» гипотезы.

Для лучшего понимания термина «статистическая значимость» необходимо понять, что такое «проверка гипотез». Эти два термина тесно взаимосвязаны.

Нормальное распределение — это функция плотности вероятности, используемая для просмотра распределения данных. Нормальное распределение имеет два параметра — среднее ( $\mu$ ) и стандартное отклонение, также называемое сигма ( $\sigma$ ).



Уровень значимости (p-value) статистического теста — это вероятность отклонить нулевую гипотезу, когда на самом деле она верна.

Уровень доверия равен  $1-\alpha$ , где  $\alpha$  — уровень значимости.

Уровень доверия — означает вероятность того, что доверительный интервал содержит истинное значение оцениваемого параметра распределения.



**Задача р-значения** ответить на этот вопрос: какова вероятность того, что CR меньше или равен 5%.

## Доверительный интервал —

это способ оценки конверсии, в результате которого мы получаем интервал значений с реальной конверсией

## Выборка и генеральная совокупность

Генеральная совокупность - совокупность всех объектов или наблюдений, относительно которых исследователь намерен делать выводы при решении конкретной задачи. В ее состав включаются все объекты, которые подлежат изучению.

Выборка - часть генеральной совокупности с помощью определённой процедуры выбранных из генеральной совокупности для участия в исследовании

Чтобы переносить выводы с выборки на генеральную совокупность, выборка должна быть репрезентативной, отражать пропорции и особенности генеральной совокупности.

## Дополнительные материалы

1. Видео "Доверительный интервал за 15 мин"
2. Калькулятор численности выборки для среднего
3. Калькулятор численности выборки для доли
4. Калькулятор t-теста. Определение стат значимости
5. Калькулятор p-value из вычисленного t-значения и степени свободы (df)
6. Критерии нормальности
7. Видео "Т-критерий Стьюдента за 12 минут"
8. Таблица стандартного отклонения:  
[https://docs.google.com/spreadsheets/d/1jIv-G6tgDV8UuvrzLCnNXUFZ0AxInnf0T\\_yAGAniJIw/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1jIv-G6tgDV8UuvrzLCnNXUFZ0AxInnf0T_yAGAniJIw/edit?usp=sharing)