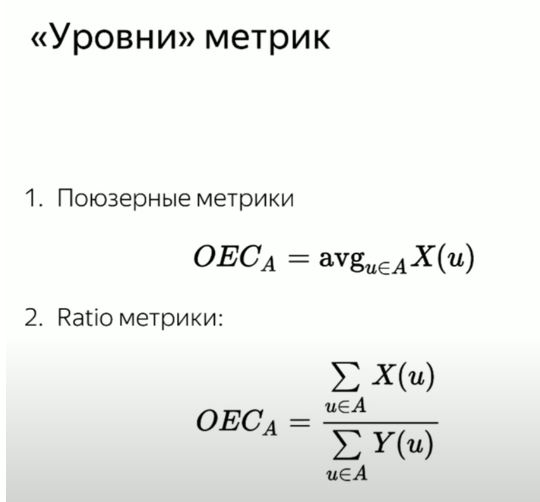
Линеаризация ratio метрик

Интро

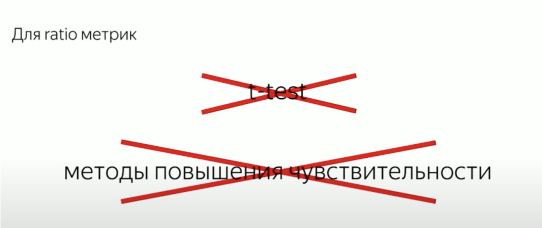
Мы можем считать что-то для каждого отдельного пользователя, например LTV, тогда наша метрика **поюзерная**.

Мы можем считать сумму каких-то значений для всех пользователей и делить на сумму других значений тоже по всем пользователям. Например, долю всех поисковых запросов, завершенных хотя бы одним активным действием. Такую метрику будем называть **Ratio**.



Прикольно! Но есть подвох?

Ага!



Теперь в числителе (да и в знаменателе) наблюдения у нас зависимы, так как могут содержать несколько значений для одного и того же пользователя. Это больше не позволит использовать родной t-test, да и с методами повышения чувствительности есть *проблемки**.

**не то чтобы, нам сейчас актуально, но для истории зафиксируем: обычно такие методы (вычитание предсказания, cuped) используют информацию о пользователя, а в таких метриках у нас может быть несколько значений для пользователя. Можем ошибиться как в самих значениях, так и в их количестве (например, пытаясь предсказать, сколько сессий будет и какой длительности), а что с ними делать дальше – совсем не ясно.*

Грустно… Но ведь умные ребята за нас уже решили эти проблемки, да?

Да! Вот вам целых три метода

- **Bootstrap**, но слишком медленный для нас
- **Delta method**, но не выдает поюзерную метрику (нельзя применить методы повышения чувствительности)
- “Наивное” поюзерное среднее и дальше брать среднее средних, но это вообще сюр, не всегда даже сонаправленность с реальной метрикой сохраняет (из **не**уважения даже жирным выделять не будем)

Много но

а мы просто готовили тебя к десерту, на сладкое у нас **Линеаризация.** Это такое преобразование, которое сделает из ratio метрики поюзерную, да еще и сохранит сонаправленность и значимость (стат. значимую разницу в ней будем видеть в тех же случаях, что и у оригинальной ratio)

Мы нашли такое преобразование $F(x, y)$, что

$$\frac{\sum\limits_{u \in A} X(u)}{\sum\limits_{u \in A} Y(u)} \approx \mathrm{avg}_{u \in A} F(X(u), Y(u))$$

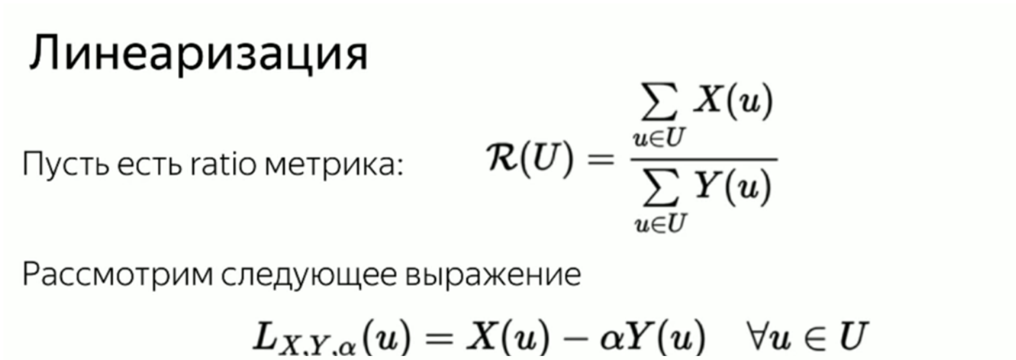
Ratio метрика $\longrightarrow$ поюзерная метрика

Вот вам еще картинка, чтобы повысить сахар в крови:

Сравнение подходов

	Easy to compute	Correct pvalue	Make user-level metric	Directionality
Bootstrap test	✗	✓	✗	✓
Delta method	✓	✓	✗	✓
User average	✓	✗	✓	✗
Linearization method	✓	✓	✓	✓

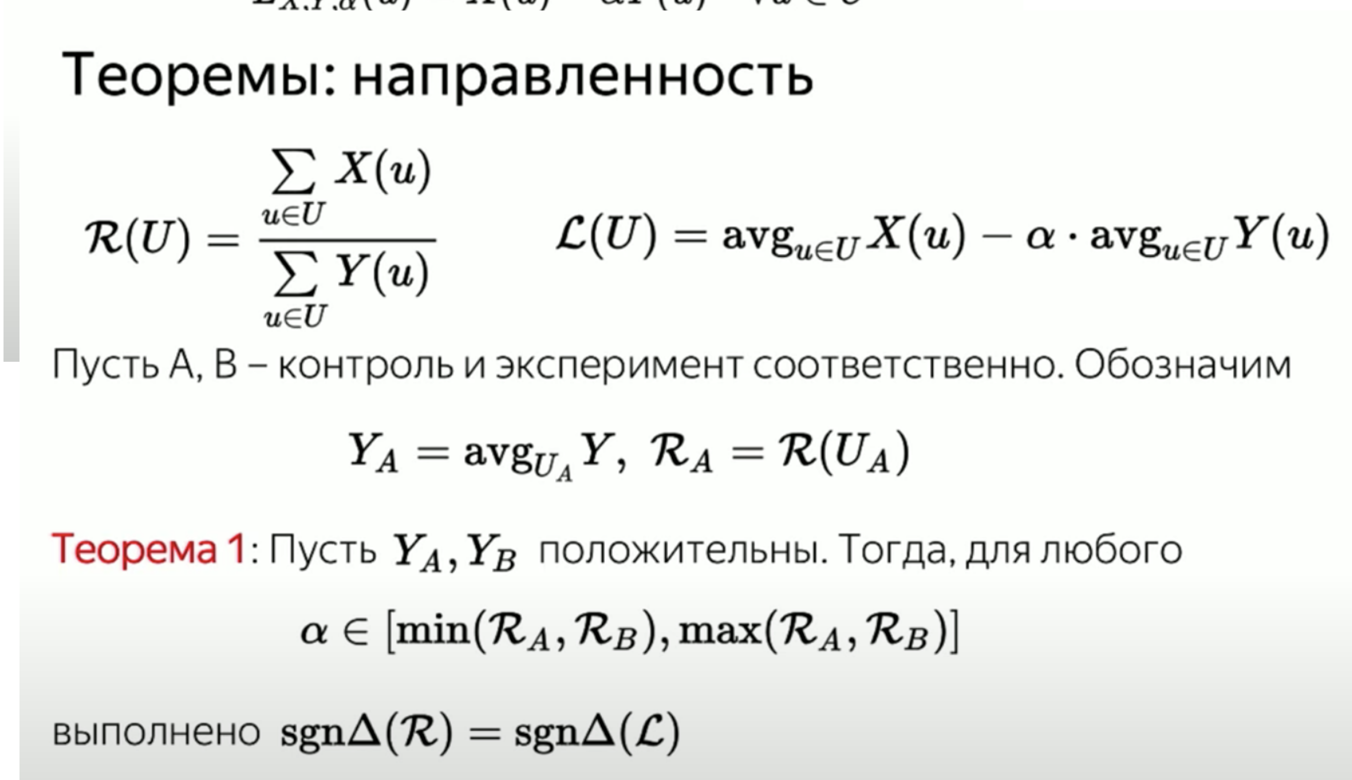
Пристегнитесь, будут формулы



Да тут все на слайде классно расписано.

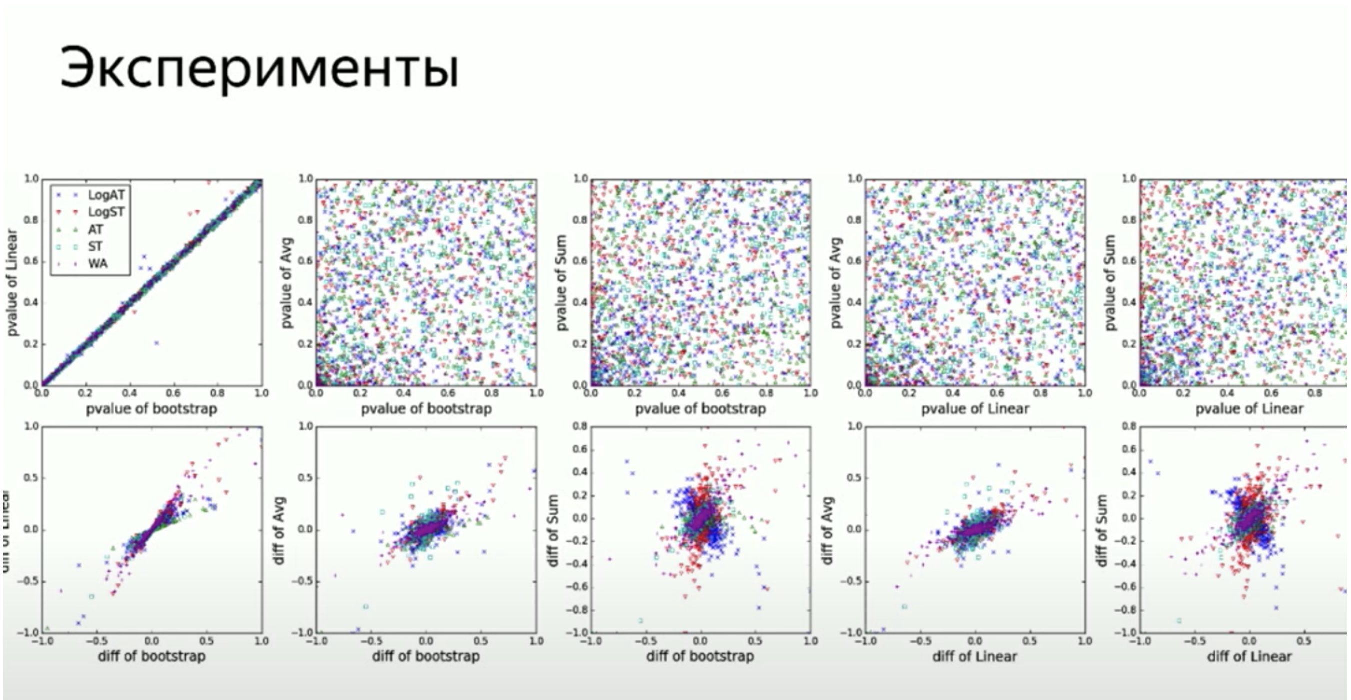
Только добавлю, что α – это такой коэффициент, который должен находится между минимальным и максимальным значением нашей ratio метрики. Обычно берут значение в контрольной группе

Сонаправленность доказывается лайтово:



За “мясом” вам сюда https://www.researchgate.net/publication/322969314_Consistent_Transformation_of_Ratio_Metrics_for_Efficient_Online_Controlled_Experiments

Целительные свойства линеаризации доказываются на практике:



Как то читать: на первых 3 картинках по оси x отмечаются значения, полученные с помощью bootstrap, а по оси y - другим методом. В первой строке картинки с p-value, во второй – с разницей между группами. Самые “классные” графики тут – это первые картинки: p-value совпадает почти полностью, график разниц меньше всего похож на облаке. Я думаю, вы уже догадались, что это графики с линеаризацией по вертикальной оси 🙂

Убедили, ну сяду как-нибудь закодить это…

покажем, как это просто, чтобы ты не откладывал в долгий ящик:

```
def linearization(control, experiment):
 numerator = 0
 denominator = 0
 for row in control:
  numerator += sum(row)
  denominator += len(row)

 control_mean = numerator / denominator
 new_control = [sum(row) - len(row) * control_mean for row in control]
 new_experiment = [sum(row) - len(row) * control_mean for row in experiment]
 return new_control, new_experiment
```

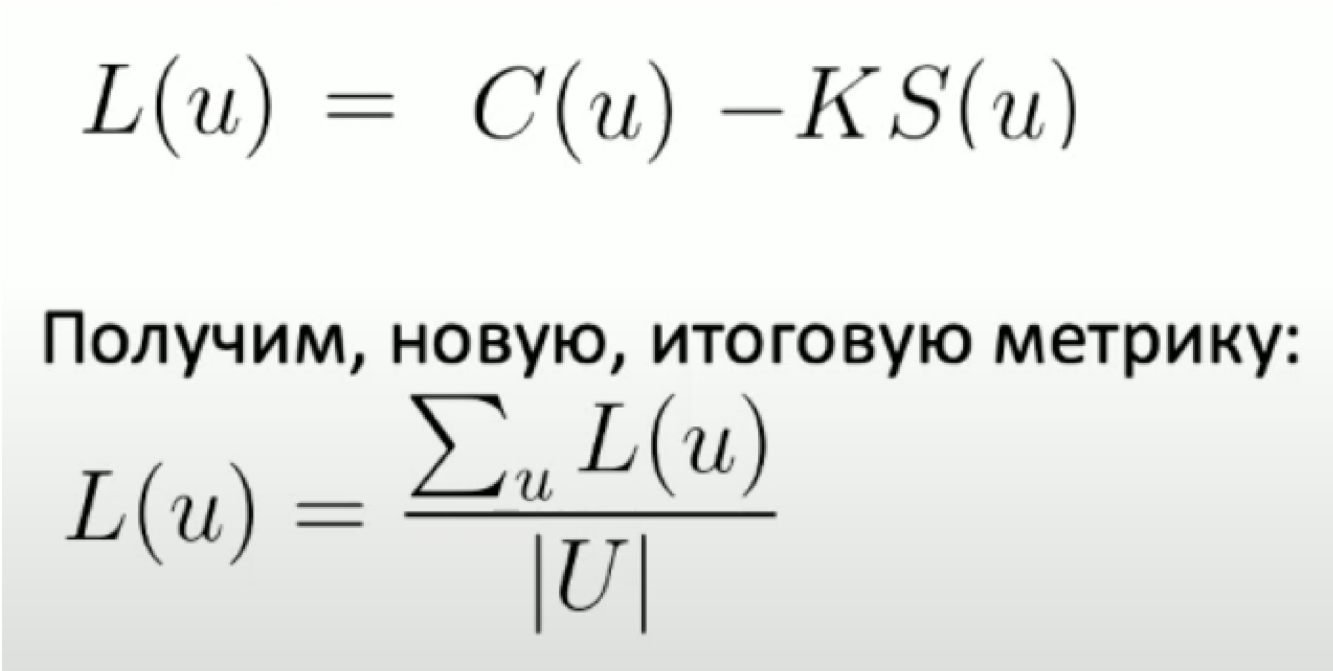
Так, а что дальше?

А дальше можешь спокойно использовать t-test поверх линеаризованных метрик!

Все было так хорошо, пока не появился ~~Навальный~~ Нерсес

В оригинальном докладе об этом не сказали, но после нашего радужного сведения ratio метрик к поюзерным мы столкнемся с проблемой *(про эту проблему рассказывает Head of Advanced Analytics в Raiffaisen CIB – Нерсес Багиян, поэтому заголовок так и называется)*

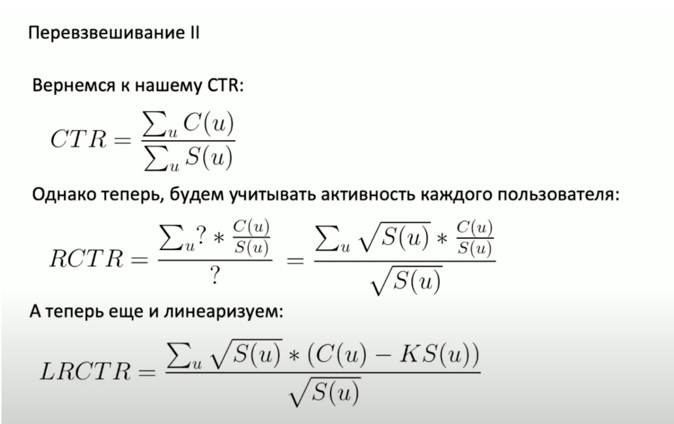
Что значит, когда мы берем среднее в нашей формуле для линеаризованной метрики? Можем ли мы как-то это переписать? Можем!



Теперь видите? Пользователь, который сделал тысячу кликов, и тот, который сделал всего один, будут учитываться в знаменател

Так, ну тут же тоже есть решение, да?

Обижаешь!



Снова слайд расскажет лучше меня, но вот вам пара обозначений.

C(u) – количество кликов, S(u) – количество показов. А вместо здесь используется обозначение **K**

Так, а откуда берется корень из количества показов? Почему именно он? Давай порассуждаем!

– Как учесть активность пользователя?
– Домножим его метрику на какой-нибудь коэффициент, а потом нормализуемся на этот коэффициент!

– Супер. А какой он может быть? Единичка подойдет?
– Нет, бро, получим сумму CTR-ов

– Может, показы?
– Тогда получим обычной CTR.. Ты точно аналитик??

– А давай тогда что-то между? Что посоветуешь?
– Да корень обычно берут..

Бинго! Вот мы и разобрались, почему!

Вы же помните ~~Руса~~ альфу??

Напоним: ***α*** *– это такой коэффициент, который должен находится между минимальным и максимальным значением нашей ratio метрики. Обычно берут значение в контрольной группе*

А мы вот только что взяли и поменяли способ подсчета значений для каждой из групп! Как в таком случае считать ***α***?

У тебя есть два стула..

- Сначала перевзвесить, потом посчитать ***α***
- Сначала посчитать ***α***, потом перевзвесить
- Вспомнить, что ожидания укрепляют отношения и запустить скрипт на boostrap

Если тебе кажется, что разницы не будет, то нет, она будет.
(не то чтобы я успела проверить, но Нерсес так сказал, Нерсесу доверяю)

Ответ: перевзвешиваем контрольную группу получаем ***α*** считаем и перевзвешиваем тестовую группу

Фух, дочитал, вроде даже разобрался

Молодец! Похвастайся в комментах, что прочитал, и поблагодари автора (опционально)!

Я понял, но не понял

Понимаем тебя 🙂 Вот тебе источники:

Основное видео от автора метода https://www.youtube.com/watch?v=vIdwgJFz5Mk&feature=youtu.be&t=1129

Все то же самое, но еще и про перевзвешивания https://www.youtube.com/watch?v=Wxw1lseUXVU

Если ты математик https://www.researchgate.net/publication/322969314_Consistent_Transformation_of_Ratio_Metrics_for_Efficient_Online_Controlled_Experiments