

Ресурсы по материалам 1 недели:

Если на прошедшей неделе Вам нелегко далось понимание определения достигаемого уровня значимости, возможно, Вам будет приятно узнать, что дело не в Вас, а в нём! [Посмотрите, как участники конференции METRICS в Стэнфорде пытаются его сформулировать.](#)

Если первая неделя далась Вам очень легко и хочется чего-то посложнее, посмотрите вот [эту интереснейшую статью о гладких байесовских обобщениях наивного бутстрэпа](#). Оказывается, наивный бутстреп — это частный случай байесовского с априорным распределением, задаваемым через дельта-функции Дирака. Потрясающе, не правда ли?

Ресурсы по материалам 2 недели:

На второй неделе мы говорили об АБ-тестах, позволяющих с помощью небольших и недорогих экспериментов оптимизировать продукт, чтобы, например, заработать на нём больше денег. Вот [здесь](#) можно почитать, как с помощью АБ-тестов, сделанных с помощью продукта компании Optimizely, EA Games на 43% повысила долю покупок игры SimCity 5. Можно ли доверять этим оценкам?

[История о том, как с помощью АА-тестов в продукте Optimizely были выявлены нарушения базовых статистических предположений](#), из-за которых оценки эффекта были завышенными. Был скандал!

Ресурсы по материалам 3 недели:

- [Сезонные эффекты в продажах;](#)
- [Пример нетранзитивности корреляции от Терренса Тао;](#)
- [xkcd про корреляцию;](#)
- [xkcd про множественную проверку гипотез;](#)
- [Использование методов множественной проверки гипотез при построении инвестиционных портфолио;](#)
- [Самый безумный вопрос про регрессию в истории Cross Validated;](#)
- [Ещё о недостатках регрессионных моделей при построении выводов о связи между одним из признаков и откликом.](#)

Список литературы

ОСНОВЫ

Если вы забыли, что такое случайная величина или статистика, какие бывают распределения, как вероятности оцениваются по выборкам, что утверждают ЗБЧ или ЦПТ, — попробуйте вернуться к конспектам четвёртой недели первого курса специализации.

Справочники по статистическим критериям

- Кобзарь. *Прикладная математическая статистика* (2006)
- Kanji. *100 statistical tests* (2006)

Хорошие вводные учебники по статистике

- Гланц. *Медико-биологическая статистика* (1999) — базовые методы объясняются на простых примерах
- Лагутин. *Наглядная математическая статистика* (2007)
- Good, Hardin. *Common Errors in Statistics (and How to Avoid Them)* (2003)

Множественная проверка гипотез

- Bretz, Hothorn, Westfall. *Multiple Comparisons Using R* (2010) — попроще
- Dickhaus. *Simultaneous Statistical Inference With Applications in the Life Sciences* (2014) — посложнее

Линейная регрессия

- Wooldridge. *Introductory Econometrics - A Modern Approach* (2012)

Дополнительно

- Hesterberg, Monaghan, Moore, Clipson, Epstein. *Bootstrap methods and permutation tests*. In *Introduction to the Practice of Statistics* (2005). <http://statweb.stanford.edu/~tibs/stat315a/Supplements/bootstrap.pdf> — доступно про бутстреп
- Davison, Hinkley. *Bootstrap Methods and their Application* (1997) — исчерпывающе про бутстреп
- Good. *Permutation, Parametric and Bootstrap Tests of Hypotheses: A Practical Guide to Resampling Methods for Testing Hypotheses* (2005) — доступно про перестановочные критерии
- Tabachnick, Fidell. *Using Multivariate Statistics* (2012) — многомерные статистические методы, в частности, дисперсионный анализ (ANOVA),

использующийся, когда нужно сравнить не две выборки, а сразу несколько. Одна из авторов в молодости выступала с танцем живота под псевдонимом ANOVA.

- Agresti. *Categorical Data Analysis* (2013) — всё о работе с категориальными данными
- Hosmer, Lemeshow, Sturdivant. *Applied Logistic Regression* (2013) — статистические основы работы с логистической регрессией (по аналогии с тем, что мы делали с линейной на третьей неделе)
- Cameron, Trivedi. *Regression Analysis of Count Data* (2013) — то же для регрессии со счётным откликом
- Pearl, Glymour, Jewell. *Causal Inference in Statistics: A Primer* (2016) — выявление причинно-следственных связей
- Schutt, O'Neil. *Doing Data Science: Straight Talk from the Frontline* (2013) — хорошая глава про причинно-следственные связи и отличия между экспериментальными данными и обсервационными