

Байес в действии

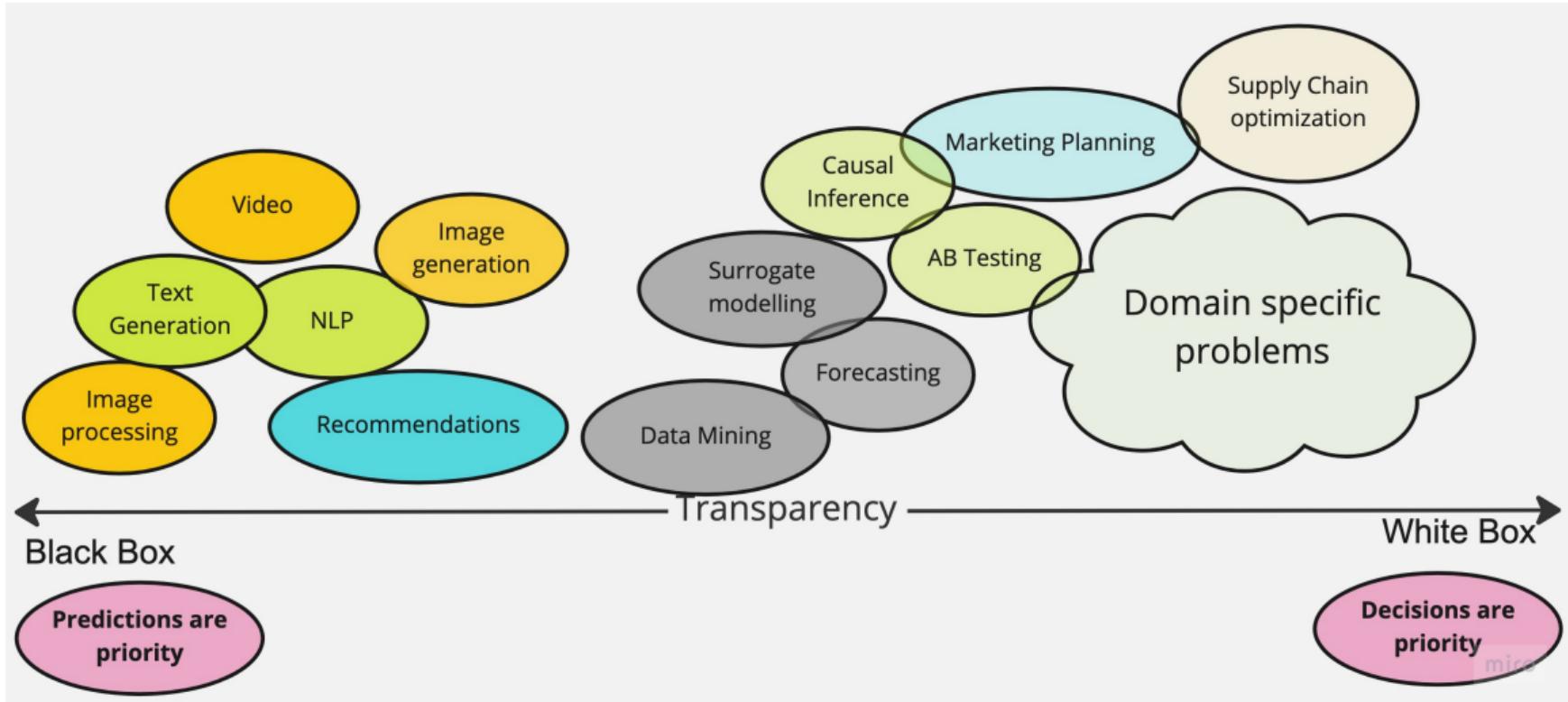
Максим Кочуров

Центральный Университет

- Образование:
 - бакалавриат – ЭФ МГУ (2018)
 - магистратура - Skoltech DS (2020)
- Один из разработчиков PyMC
- Ведущий data сайентист в PyMC Labs
- Выпускник BayesGroup
- Имею опыт работы с:
 - машинным обучением и компьютерным зрением;
 - дифференциальной геометрией для графовых нейросетей;
 - байесовскими методами для глубокого обучения;
 - прикладной байесовской статистикой в индустрии (A/B тестирование, биоинформатика).



Обзор задач



Классическое машинное обучение имеет ограничения.

- Мало контроля параметризации моделей
- Сложно добавить иерархию
- Гибкость – не приоритет

Фокус на алгоритмах, не на решаемой задаче

Оговорка

Обычно это не проблема, но не все задачи решаются стандартными способами.



Думайте о задаче

"White box" модели более гибкие

- Контроль параметризации
- Структура модели – Ваше решение
- Возможности покажутся магией

Фокус на задаче, алгоритмы – лишь инструмент.

Оговорка

Удостоверьтесь, что Ваша задача не типа "fit-predict". В противном случае от байесовских методов не будет дополнительных плюсов.



Идентификация задачи

Данные обычно структурированы

- Вложенные/иерархические
- Временные ряды
- Прочая нетривиальная структура

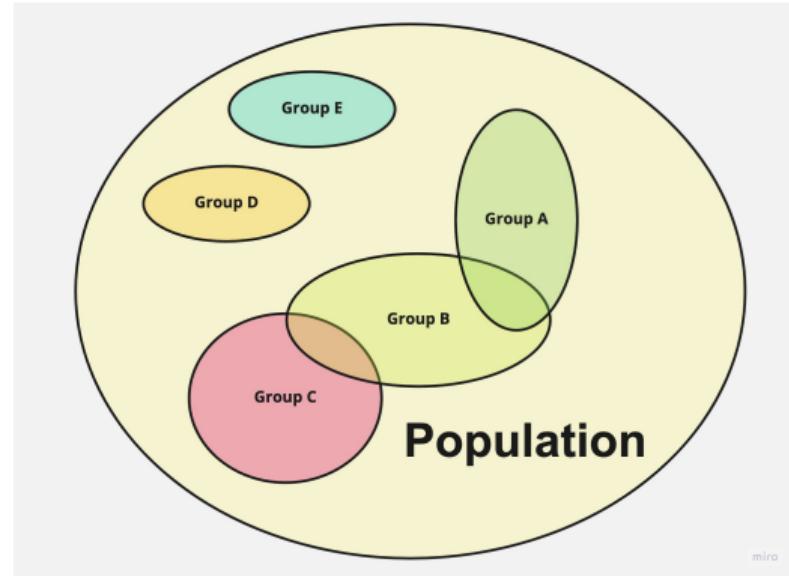
Ваша задача – не просто предсказание:

- Если Вам нужны только предсказания, ML для Вас
- Если Вам не нужна интерпретация, DL для Вас



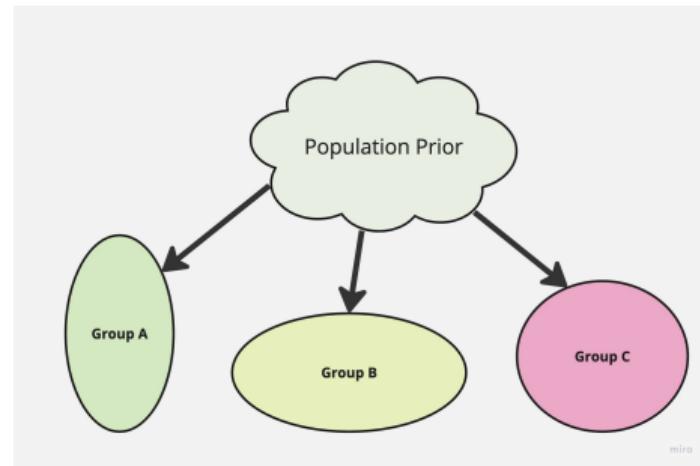
Иерархия

- Прогноз объёма продаж для нескольких магазинов
- Анализ распространения инфекционных заболеваний
- Прогноз потребительских предпочтений



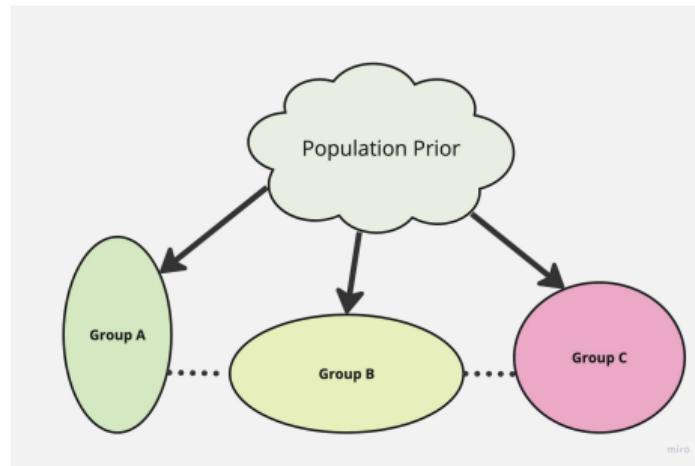
Используем весь потенциал иерархий

- Неравные группы вносят вклад в общее априорное распределение
- Байесовский вывод: распределение знания об общей структуре между группами
- Можно симулировать ненаблюдаемые группы



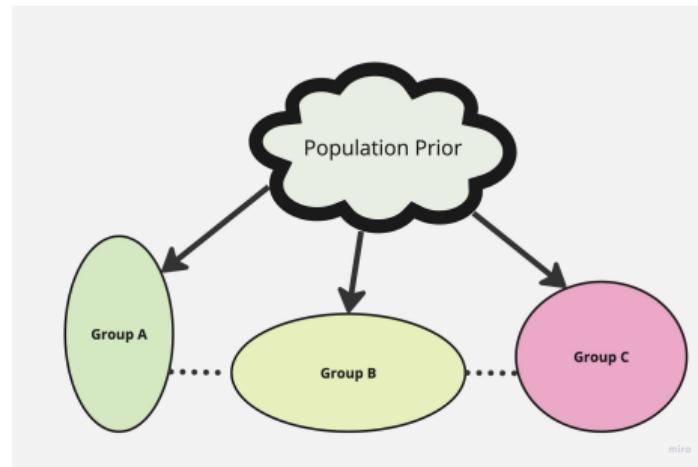
Используем весь потенциал иерархий

- Неравные группы вносят вклад в общее априорное распределение
- Байесовский вывод: распределение знания об общей структуре между группами
- Можно симулировать ненаблюдаемые группы



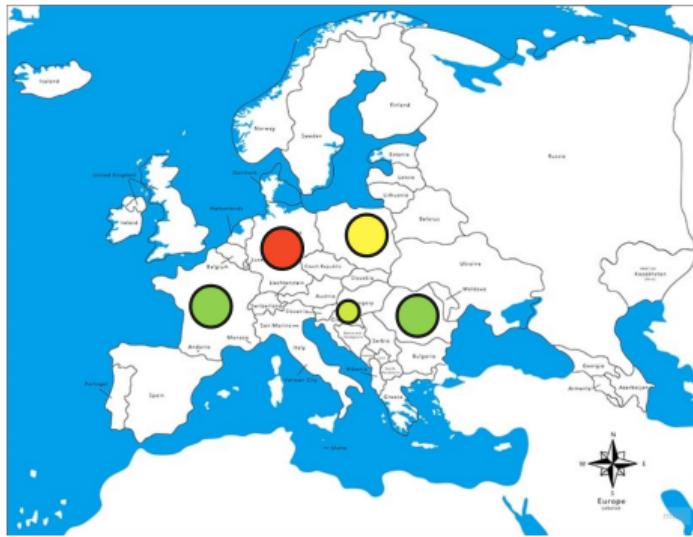
Используем весь потенциал иерархий

- Неравные группы вносят вклад в общее априорное распределение
- Байесовский вывод: распределение знания об общей структуре между группами
- **Можно симулировать ненаблюдаемые группы**



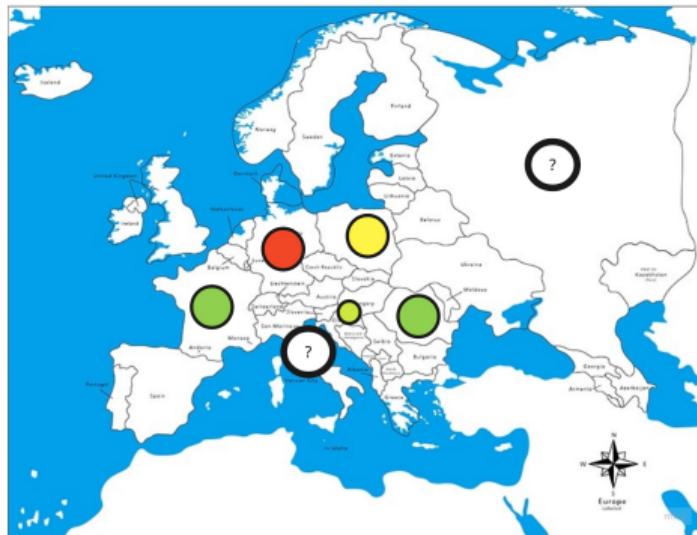
Прогноз для ненаблюдаемых групп

- Вы работаете на нескольких региональных рынках и хотите расширяться
- Нужно угадать предпочтения людей в новых регионах
- Используйте байесовские методы, чтобы симулировать ненаблюдаемое



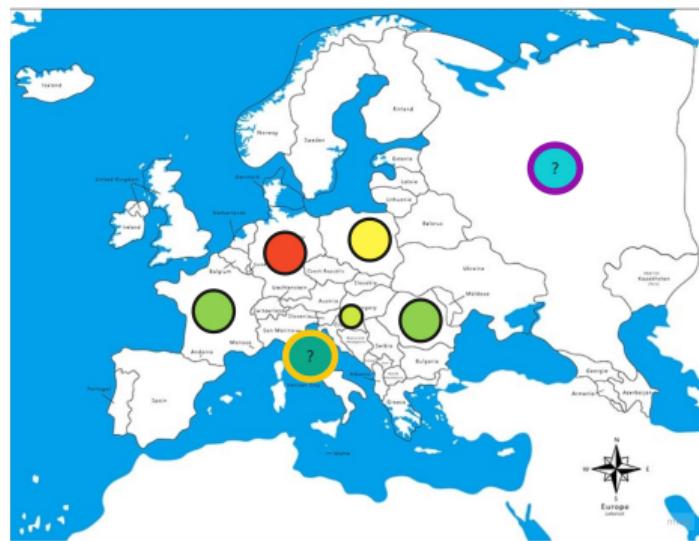
Прогноз для ненаблюдаемых групп

- Вы работаете на нескольких региональных рынках и хотите расширяться
- Нужно угадать предпочтения людей в новых регионах
- Используйте байесовские методы, чтобы симулировать ненаблюдаемое



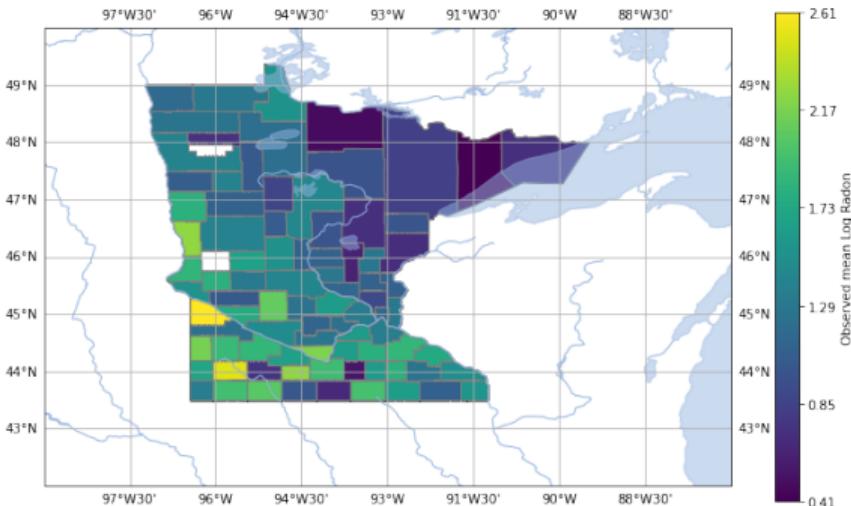
Прогноз для ненаблюдаемых групп

- Вы работаете на нескольких региональных рынках и хотите расширяться
- Нужно угадать предпочтения людей в новых регионах
- Используйте байесовские методы, чтобы симулировать ненаблюдаемое



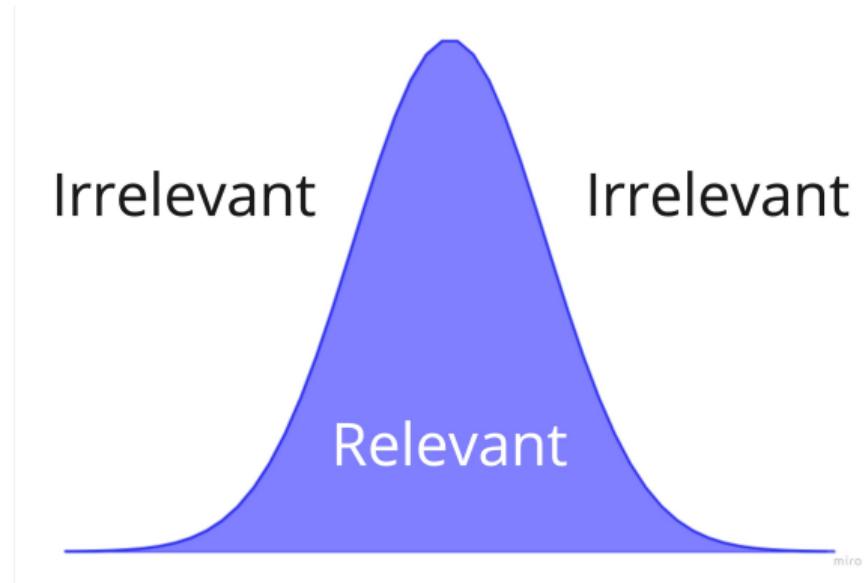
Бонус: иерархии в географии

- Группы взаимозависимы
- Можно использовать эту взаимозависимость, чтобы легко повысить качество предсказаний



Априорные распределения – это круто

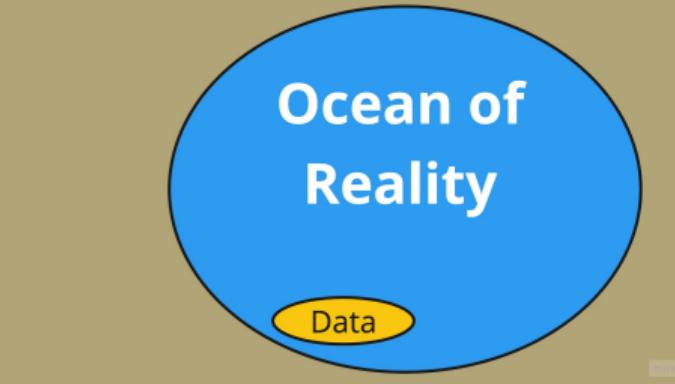
- Априорные распределения – как регуляризация
- Мягкое ограничение пространства решений
- Помогают внести в модель экспертные знания



Когда мало данных

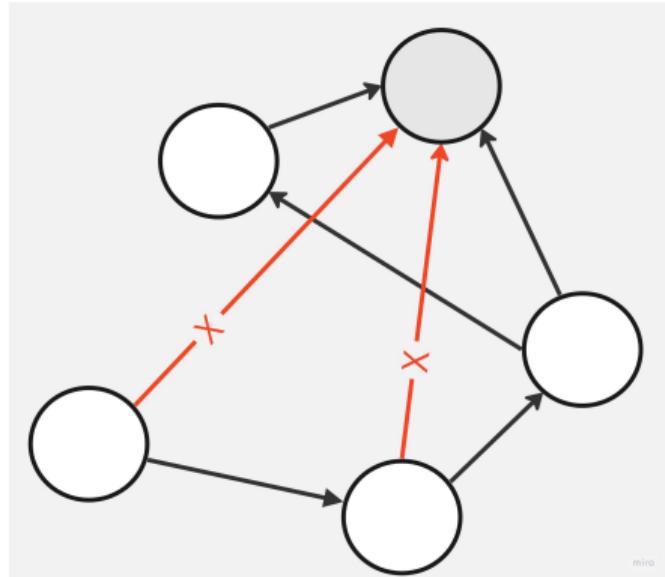
- Остаётся опираться на экспертные знания.
- Можно вложить эти знания в априорные распределения
- Каждый пример важен

Ground of Uncertainty



Структура модели: причинность

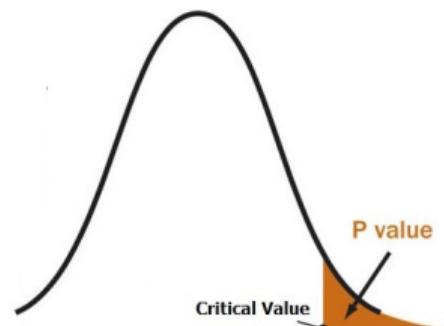
- Не давайте "чёрному ящику" принимать решения
- Задайте собственную структуру
- Структуры разные: "одна на всех" не подойдёт



р-значение в А/Б тестах

Что из нижеперечисленного верно?

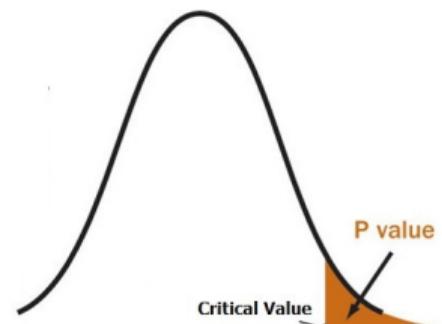
- ① p – вероятность того, что верна нулевая гипотеза.
- ② $(1 - p)$ – вероятность того, что верна альтернативная гипотеза.
- ③ $p \leq 0.05$ значит, что нулевая гипотеза неверна и должна быть отвергнута.
- ④ Значение $p > 0.05$ значит, что эффект отсутствует.



p-значение в А/Б тестах

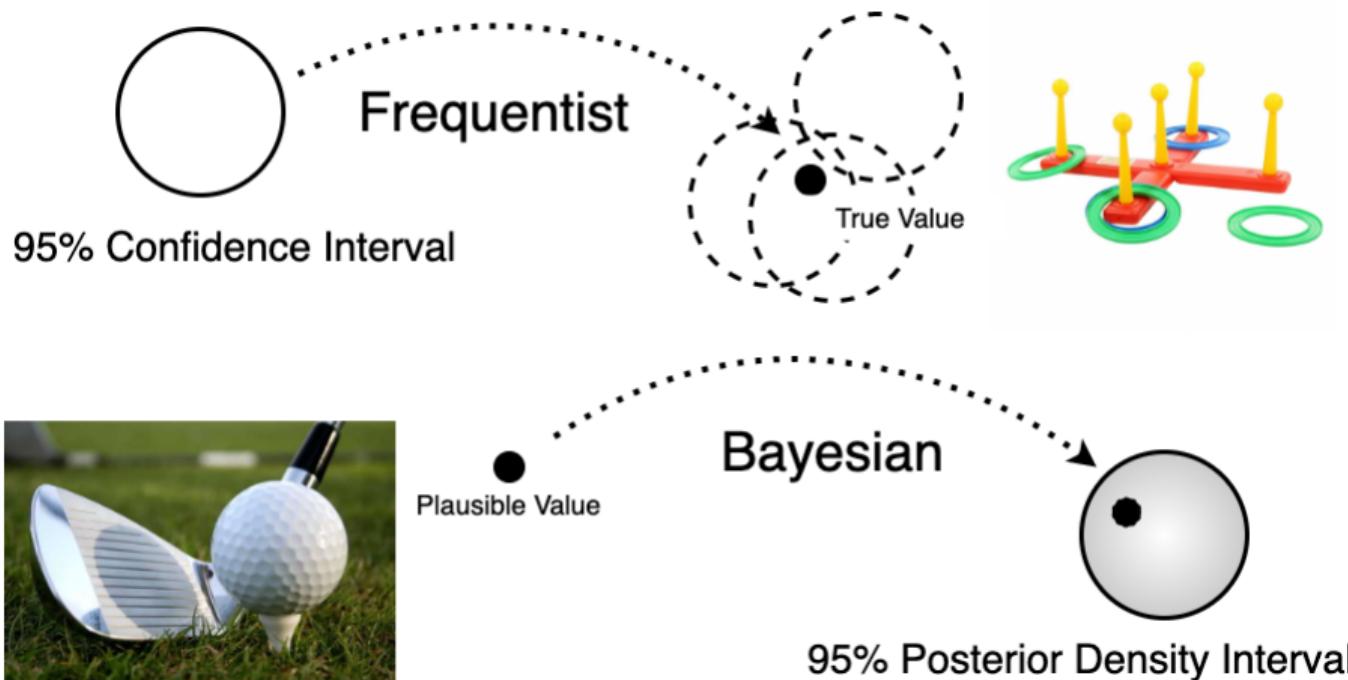
Что из нижеперечисленного верно?

- ① p – вероятность того, что верна нулевая гипотеза.
- ② $(1 - p)$ – вероятность того, что верна альтернативная гипотеза.
- ③ $p \leq 0.05$ значит, что нулевая гипотеза неверна и должна быть отвергнута.
- ④ Значение $p > 0.05$ значит, что эффект отсутствует.



Использовать p-значение плохо?

Я не предлагаю отказаться от p-значений, я предлагаю лучше понять их



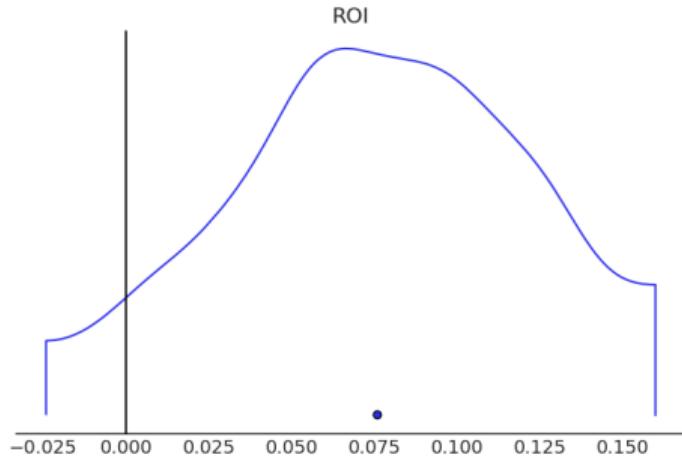
Неопределённость

$$ROI = \frac{NPV - CapEx}{CapEx}$$

$$NPV = (CLV(r) - OpEx(r)) \cdot \Delta N$$

- $CLV(r)$ - Discounted User Value
- $OpEx(r)$ - Operational Cost

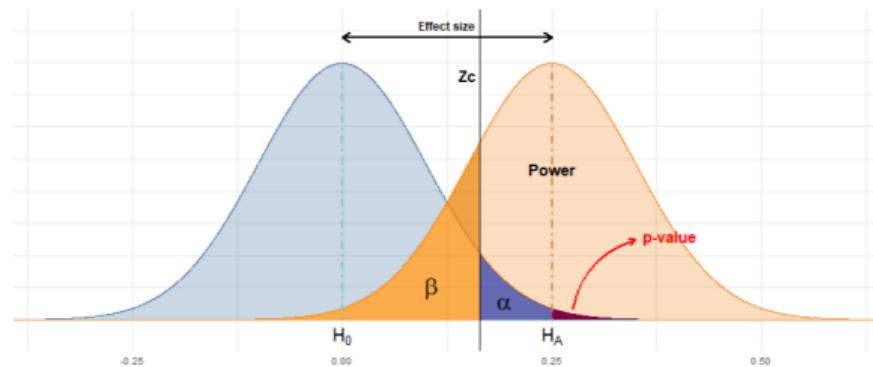
- В финансах нужно оценивать риски
- Делайте симуляции, не доверяйте точечным оценкам
- Байес помогает принимать информированные решения



Кейс: АВ(С) тест

Проблемы

- Работает при равных (50/50) объёмах выборок
 - Сложнее при объёмах 10/90
- Точечные оценки размера эффекта
 - Сложно действовать на их основе
- Бутстрэп делать долго
 - Плюс он не очень помогает, когда мало данных
- Предполагаем простые распределения



Кейс: AB(C) тест

Сложные наблюдения

Тестирование сразу 2-х гипотез, не-нормальное распределение

- Объём покупок



- Вероятность покупки



Множественные тесты

Нет статистических поправок

- 3 или 10 тестов!
- Пробные тесты на малых выборках!
- Объёмы выборок могут быть не сбалансированы!
- Тестируем в онлайн режиме!

Экономика

Можно вычислить экономический эффект реализации

- Какая самая плохая окупаемость инвестиций (ROI)?

Новые результаты: линейная регрессия

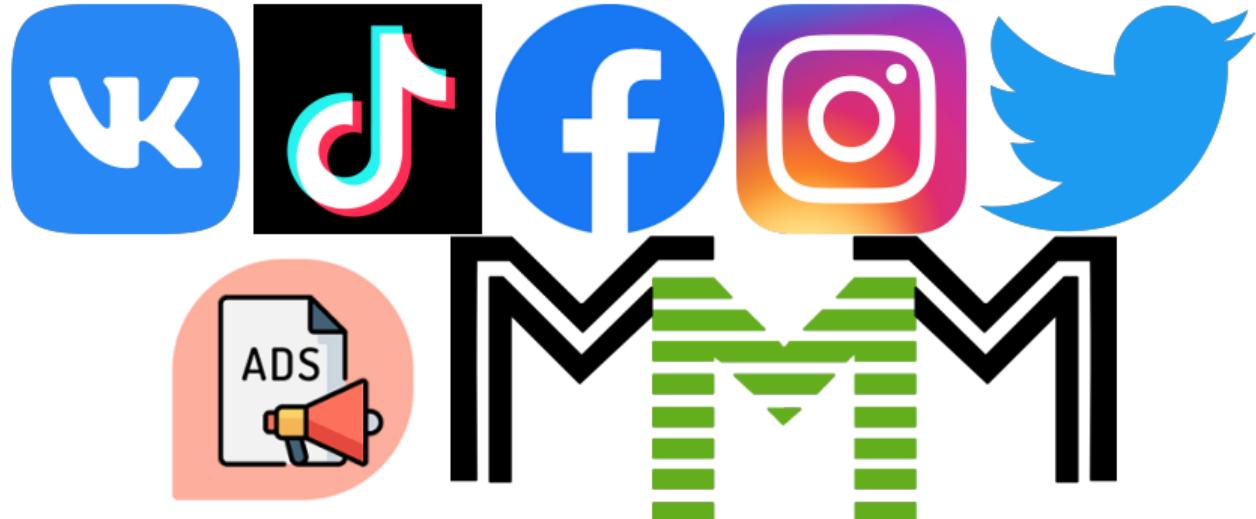
- ① Стандартизируйте данные: $a \mapsto \frac{a - \text{mean}(a)}{\text{std}(a)}$
- ② Выберите R^2
- ③ Выберите важность признаков
- ④ Выберите знак размера эффекта
- ⑤ Готово



Вебинар про линейные регрессии

Ещё больше формул

Это недавно разработанное априорное распределение R2D2M2, см. подробности в [1].
Реализация R2D2M2 в PyMC: [2].



Media Mix Model

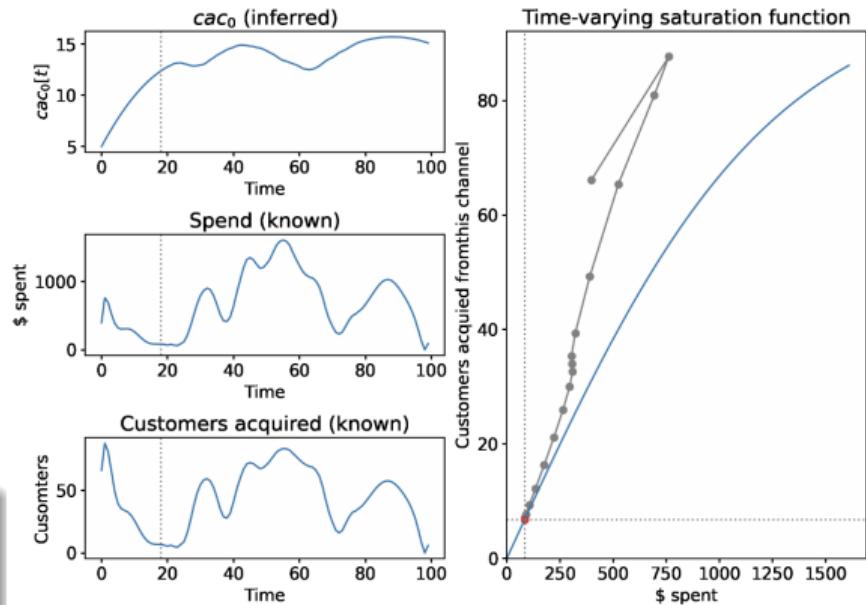
МММ – помогает оценивать маркетинговые каналы по историческим данным

Кейс: маркетинг

- Априорные распределения:
устанавливаем стоимость привлечения клиентов
- Структура: эта стоимость меняется
- Неопределенность: выделяем бюджет обоснованно

Результаты

- Лучшее понимание каналов
- Эффективное привлечение клиентов



PyMC
Labs



**HELLO
FRESH**

Открытые вопросы в выращивании сельхозкультур

- Как выбрать удобрения?
- Как выбрать регион?
- Как комбинировать регионы и удобрения?

Решение

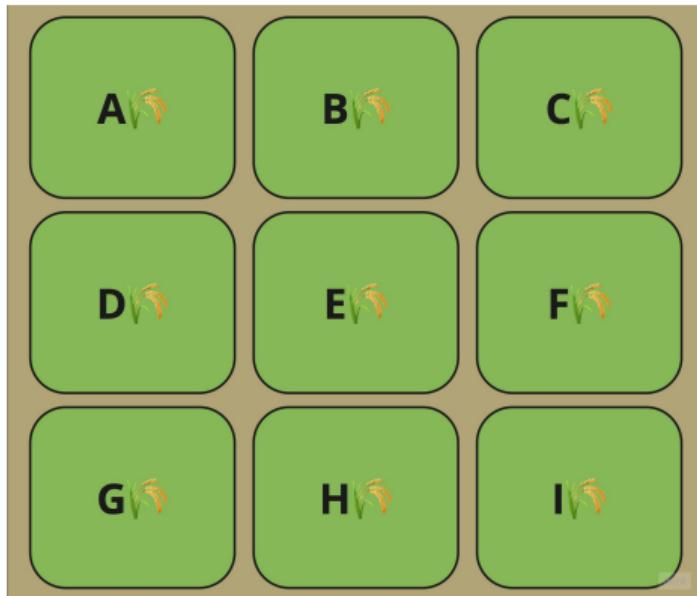
Всестороннее А/Б тестирование



- Иерархия: общий эффект для регионов
- Структура: соседние поля взаимосвязаны
- Прогноз: урожайность нового поля

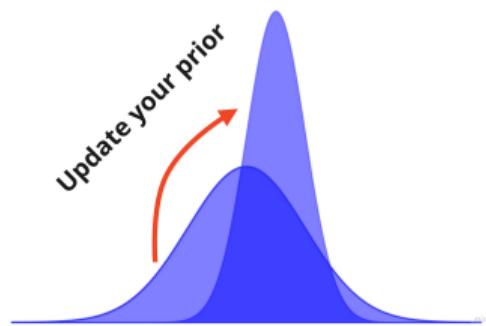
Результат

Пайплайн для экспериментов с удобрениями на уровне континента



Хотите изучить байесовские методы?

Мы начинаем в
сентябре 2025



Библиография I

-  J. E. Aguilar and P.-C. Bürkner.
Intuitive joint priors for bayesian linear multilevel models: The r2d2m2 prior, 2023.
-  M. Kochurov.
[pymc-devs/pymc-experimental: Pull Request 137 R2D2M2CP.](https://github.com/pymc-devs/pymc-experimental/pull/137)
GitHub, 2023.