

# Байесовское моделирование

Максим Кочуров

МГУ им. М.В. Ломоносова

Лекция 2



# Содержание

- ① НМС
- ② Модель Кобба-Дугласа
- ③ Базовая модель
  - Настройка
  - Априорные распределения
- ④ Проверка априорного предиктивного распределения
- ⑤ Иерархии
  - Введение
  - Байесовский подход
  - Параметризация
  - Априорные распределения
- ⑥ Обсуждение



# Сэмплирование из распределения

## Сопряженные модели

- Ограниченнная применимость
- Недостаточная гибкость
- Хорошая масштабируемость

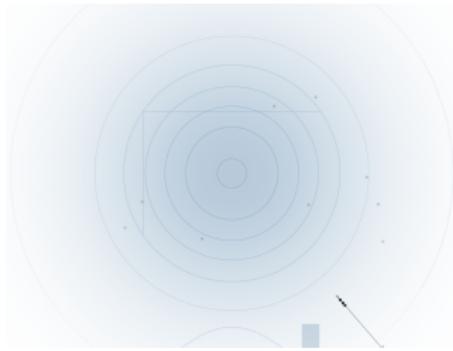


Рис.: Простое распределение

## Большинство моделей

- Нет аналитического решения
- Апостериорные распределения сложные
- Плохая масштабируемость
- Гибкость

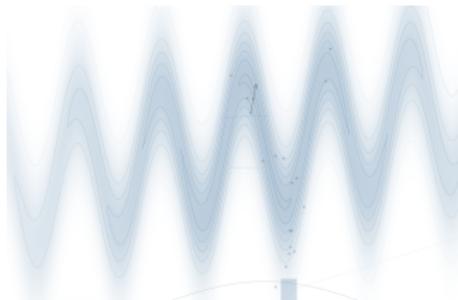


Рис.: Сложное распределение

# Метод Монте-Карло с механикой Гамильтона (НМС)



В НМС сэмплирование происходит из сложных распределений

- ① Берем идеи из физики
- ② Считаем градиенты
- ③ Численно интегрируем

Откалиброванная модель НМС сходится к целевому распределению

## Предупреждение

Мы обещали курс без сложной математики. Но НМС нужен для отладки ваших моделей.

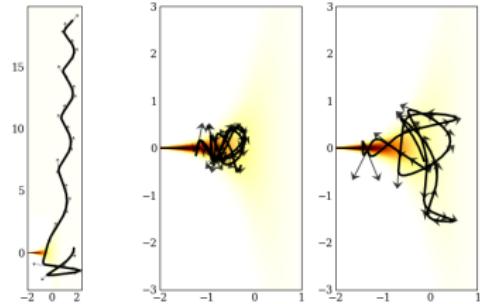


Рис.: интегрирование Leapfrog



# Построение распределений через НМС

- $p(\Theta)$  - Целевое распределение,  $\Theta \in \mathbb{R}^d$  ( $\Theta$  или Координаты)
- $p(\Delta | \Theta)$  - Распределение импульса,  $\Delta \in \mathbb{R}^d$  ( $\Delta$  или Ускорение)

Гамильтониан вычисляется как

$$H(\Delta, \Theta) = -\log p(\Delta, \Theta)$$

## Заметка

- $p(\Delta | \Theta) = \text{Normal}(0, M)$
- $\Delta$  и  $\Theta$  имеют одинаковые размерности

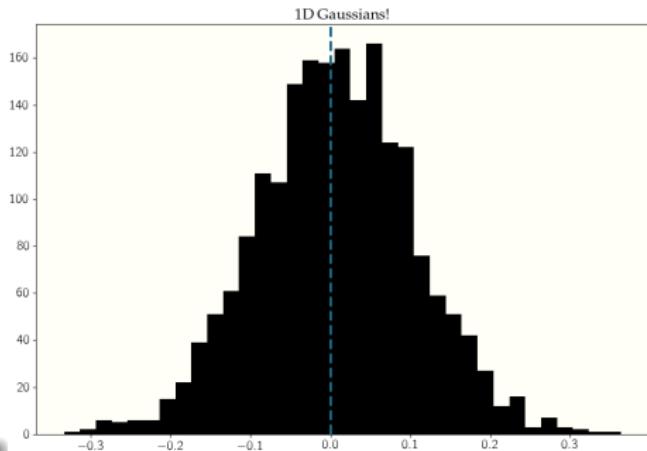


Рис.:  $p(\Theta) = \text{Normal}(0, 1)$



# Система дифференциальных уровней НМС

$$\begin{aligned}
 H(\Delta, \Theta) &= -\log p(\Delta, \Theta) \\
 &= -\log p(\Delta | \Theta) - \log p(\Theta) \\
 &= \underbrace{K(\Delta, \Theta)}_{\text{Kinetic E}} + \underbrace{V(\Theta)}_{\text{Potential E}}
 \end{aligned}$$

Физическая система, задающая движение

$$\begin{aligned}
 \frac{\partial \Theta}{\partial t} &= \frac{\partial H}{\partial \Delta} \\
 \frac{\partial \Delta}{\partial t} &= -\frac{\partial H}{\partial \Theta}
 \end{aligned}$$

Движение содержит полную энергию  $H(\Delta, \Theta)$

Максим Кочуров

Прикладные байесовские методы



Рис.: скейт-парк и НМС



# Дивергенции в НМС

Дивергенции возникают при крупной ошибке численного интегрирования в ходе решения дифференциального уравнения.

## Когда НМС не работает

Неудачный выбор геометрического распределения вероятностей для гамильтониана

Почему геометрическое распределение неудачно

- ① Мультиколлинеарность
- ② Распределения вида воронки
- ③ Сильное влияние функции правдоподобия
- ④ Неоднородное апостериорные распределение

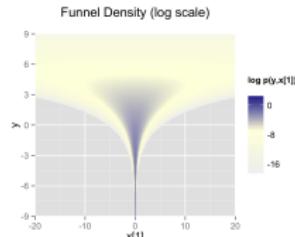


Рис.: Воронка Нила



# Что почитать по НМС

Для дальнейшего погружения

- ① Интерактивное [демо](#)
- ② [Пособие](#) от Colin Carroll
- ③ [Статья](#) Michael Betancourt
- ④ [Статья](#) про NUTS от Matthew D. Hoffman, Andrew Gelman

# Пример



# Простой пример - функция Кобба-Дугласа

Скорее всего, вы знакомы с производственной функцией вида

$$Y \approx A \cdot L^\beta$$

Имеем:

- ① 6 различных групп (иерархия)
- ② Группы различимы
- ③ Мы знаем, что факторная производительность  $A$  отличается между ними
- ④ Производительность труда  $\beta$  отличается не сильно

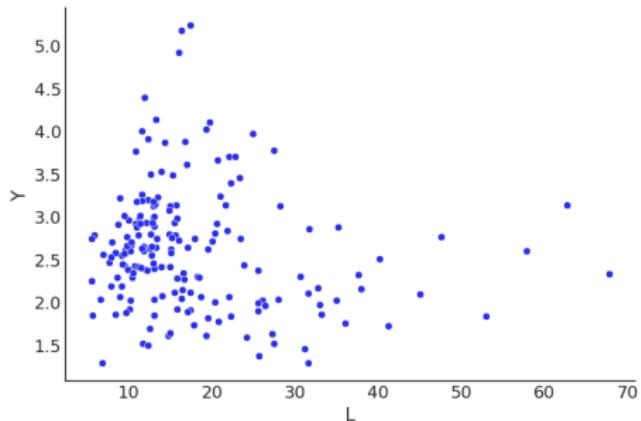


Рис.: Пример агрегированных данных



# Простой пример - производство ковров

Уточним пример

$$Y_g \approx A_g \cdot L^\beta$$

Пусть есть ателье, шьющее ковры, с 6 рабочими:

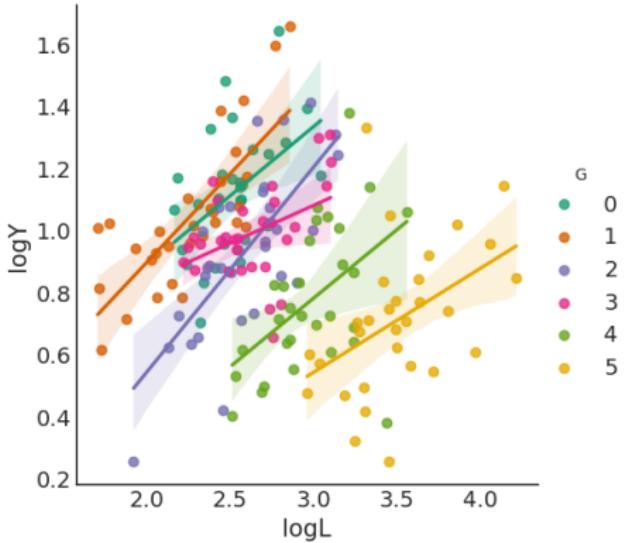
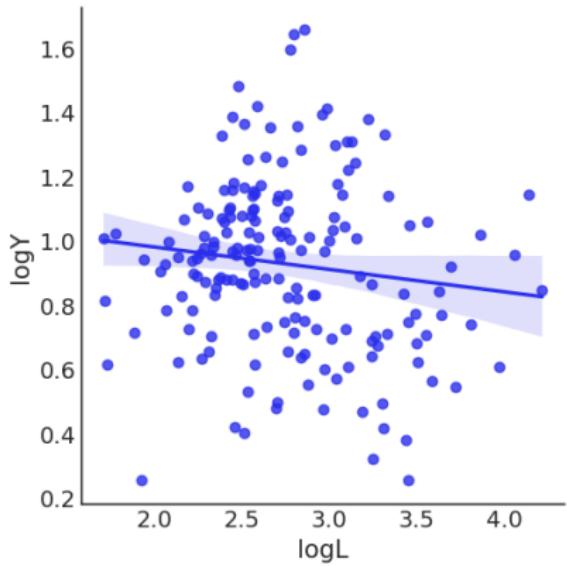
- ① Работники делают разные ковры с различной факторной производительностью  $A$
- ② Производительность труда  $\beta$  можно воспринимать как сосредоточенность, со временем работники устают
- ③ Работник шьет ковер один, без чужой помощи



Рис.: Пример продукта ( $Y$ )



# Парadox Симпсона





# Модель с одной группой

Несколько советов для моделирования.

- Начинать стоит с простой модели
- Убедитесь, что модель сходится
- Усложните модель



# Модель с одной группой

Несколько советов для моделирования.

- Начинать стоит с простой модели
  - Рассмотрите модель с 1 группой
  - Проверьте, что априорные распределения выбраны корректно
- Убедитесь, что модель сходится
- Усложните модель



# Модель с одной группой

Несколько советов для моделирования.

- Начинать стоит с простой модели
  - Рассмотрите модель с 1 группой
  - Проверьте, что априорные распределения выбраны корректно
- Убедитесь, что модель сходится
  - Если не сойдется модель с одной группой – модели с множеством групп тоже не сойдутся
  - Проведите проверку, что модель семплируется корректно
- Усложните модель



# Модель с одной группой

Несколько советов для моделирования.

- Начинать стоит с простой модели
  - Рассмотрите модель с 1 группой
  - Проверьте, что априорные распределения выбраны корректно
- Убедитесь, что модель сходится
  - Если не сойдется модель с одной группой – модели с множеством групп тоже не сойдутся
  - Проведите проверку, что модель семплируется корректно
- Усложните модель
  - Попробуйте сделать параметризацию
  - Проверьте семплирование модели
  - Сравните модели (про это – поговорим далее по курсу)



# Простая модель: с чего начать?

Для начала зададим себе несколько вопросов

$$Y_{g=0} \approx A_{g=0} \cdot L^{\beta}$$

- ① Какое апостериорное распределение взять для  $A$ ?
- ② Какое апостериорное распределение взять для  $\beta$ ?
- ③ Какое апостериорное предиктивное распределение для  $Y_{g=0}$ ?



## Формулируем модель

$$Y_{g=0} \approx A_{g=0} \cdot L^\beta$$

$$\log Y_{g=0} \approx \log A_{g=0} + \log L \cdot \beta$$

Предполагаем следующие распределения

$$\log Y_{g=0} \sim \text{Normal}(\log A_{g=0} + \log L \cdot \beta, \epsilon)$$

$$\epsilon \sim ???$$

$$\beta \sim ???$$

$$A_{g=0} \sim ???$$



## Априорное распределение для $\beta$

Какие свойства мы предполагаем для производительности (или эластичности) труда  $\beta$ ? Подумайте

- ① Может ли она быть  $< 0$ ?
- ② Может ли она принимать большие значения?
- ③ Может ли она быть  $> 1$ ?



## Априорное распределение для $\beta$

Какие свойства мы предполагаем для производительности (или эластичности) труда  $\beta$ ? Подумайте

- ① Может ли она быть  $< 0$ ? Нет
- ② Может ли она принимать большие значения?
- ③ Может ли она быть  $> 1$ ?



## Априорное распределение для $\beta$

Какие свойства мы предполагаем для производительности (или эластичности) труда  $\beta$ ? Подумайте

- ① Может ли она быть  $< 0$ ? Нет
- ② Может ли она принимать большие значения? Нет
- ③ Может ли она быть  $> 1$ ?



## Априорное распределение для $\beta$

Какие свойства мы предполагаем для производительности (или эластичности) труда  $\beta$ ? Подумайте

- ① Может ли она быть  $< 0$ ? Нет
- ② Может ли она принимать большие значения? Нет
- ③ Может ли она быть  $> 1$ ? Нет



## Априорное распределение для $\beta$

Какие свойства мы предполагаем для производительности (или эластичности) труда  $\beta$ ? Подумайте

- ① Может ли она быть  $< 0$ ? Нет
- ② Может ли она принимать большие значения? Нет
- ③ Может ли она быть  $> 1$ ? Нет

Заключение: она лежит в промежутке  $(0, 1)$

Данные априорные предположения – субъективны!

Есть у кого-то замечания по нашим ограничениям?



## Априорное распределение для $\beta$

Какие свойства мы предполагаем для производительности (или эластичности) труда  $\beta$ ? Подумайте

- ① Может ли она быть  $< 0$ ? Нет
- ② Может ли она принимать большие значения? Нет
- ③ Может ли она быть  $> 1$ ? Нет

Заключение: она лежит в промежутке  $(0, 1)$

Данные априорные предположения – субъективны!

Есть у кого-то замечания по нашим ограничениям?

Мы пока еще не вывели априорное распределение

Нужно подобрать распределение, которое удовлетворяет нашим условиям



# Априорное распределение для $\beta$

Что имеем:

- $\beta \in (0, 1)$
- Маловероятно, что истинное значение близко к границам
- Ничего не знаем о том, где конкретно в интервале оно лежит.

## Заметка

Посмотрим распределения, которые удовлетворяют требованиям



# Априорное распределение для $\beta$

Что имеем:

- $\beta \in (0, 1)$
- Маловероятно, что истинное значение близко к границам
- Ничего не знаем о том, где конкретно в интервале оно лежит.

## Заметка

Посмотрим распределения, которые удовлетворяют требованиям

- ➊ Beta( $a, b$ ),  $a > 0, b > 0$  где  $a, b$  удовлетворяют ограничениям



# Априорное распределение для $\beta$

Что имеем:

- $\beta \in (0, 1)$
- Маловероятно, что истинное значение близко к границам
- Ничего не знаем о том, где конкретно в интервале оно лежит.

## Заметка

Посмотрим распределения, которые удовлетворяют требованиям

- ① Beta( $a, b$ ),  $a > 0, b > 0$  где  $a, b$  удовлетворяют ограничениям
- ② LogitNormal( $\mu, \sigma$ ) - всегда удовлетворяет ограничениям



# Априорное распределение для $\beta$

Что имеем:

- $\beta \in (0, 1)$
- Маловероятно, что истинное значение близко к границам
- Ничего не знаем о том, где конкретно в интервале оно лежит.

## Заметка

Посмотрим распределения, которые удовлетворяют требованиям

- ① Beta( $a, b$ ),  $a > 0, b > 0$  где  $a, b$  удовлетворяют ограничениям
- ② LogitNormal( $\mu, \sigma$ ) - всегда удовлетворяет ограничениям
- ③ Uniform(0, 1) - частный случай Beta(1, 1)



# Априорное распределение для $\beta$

Что имеем:

- $\beta \in (0, 1)$
- Маловероятно, что истинное значение близко к границам
- Ничего не знаем о том, где конкретно в интервале оно лежит.

## Заметка

Посмотрим распределения, которые удовлетворяют требованиям

- ① Beta( $a, b$ ),  $a > 0, b > 0$  где  $a, b$  удовлетворяют ограничениям
- ② LogitNormal( $\mu, \sigma$ ) - всегда удовлетворяет ограничениям
- ③ Uniform(0, 1) - частный случай Beta(1, 1)
- ④ Kumaraswamy( $a, b$ ),  $a > 0, b > 0$  но лучше обойтись без него



## Визуализируем априорное распределение

Перед написанием изобразите распределения. Какое лучше?

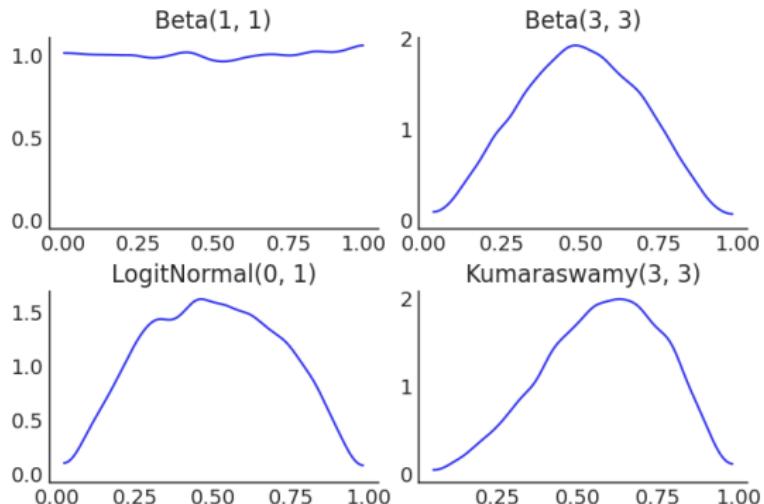


Рис.: Распределения с прошлого слайда

Выбирайте то, которое нравится и согласуется с теорией



## Настраиваем априорное распределение

Для этой задачи я выберу LogitNormal( $0, 1$ ). У него удобная функциональная форма

### Помните

- Априорное распределение при моделировании выбираете **вы**
- Этот выбор должен быть обоснован
- Выбор должен быть разумным, учитывать накопленное знание
- Вы должны иметь возможность защитить свой выбор
- **Априорное распределение – источник неопределенности**



# Простая модель: что имеем на данный момент

$$\log Y_{g=0} \sim \text{Normal}(\log A_{g=0} + \log L \cdot \beta, \epsilon)$$

$$\epsilon \sim ???$$

$$\beta \sim \text{LogitNormal}(0, 1)$$

$$A_{g=0} \sim ???$$



# Априорное распределение $\epsilon$

## Эмпирическое наблюдение

Ошибка обычно не велика. Но при этом не равна нулю.

В нашем случае;

- "маленькая ошибка" означает отклонение в 10-50%

Пусть

$$\epsilon \sim \text{LogNormal}(-2, 1)$$

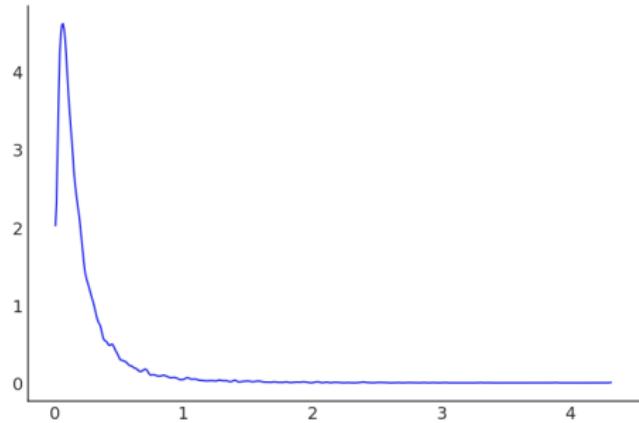


Рис.: Априорное распределение для  $\epsilon$



# Априорное распределение $\epsilon$

## Эмпирическое наблюдение

Ошибка обычно не велика. Но при этом не равна нулю.

В нашем случае;

- "маленькая ошибка" означает отклонение в 10-50%

Пусть

$$\epsilon \sim \text{LogNormal}(-2, 1)$$

## Примечание

В логарифмических моделях  
ошибка – величина относительная

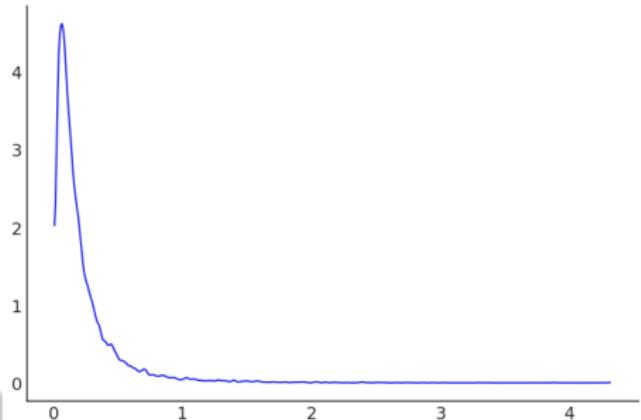


Рис.: Априорное распределение для  $\epsilon$



# Простая модель: что имеем на данный момент

$$\log Y_{g=0} \sim \text{Normal}(\log A_{g=0} + \log L \cdot \beta, \epsilon)$$

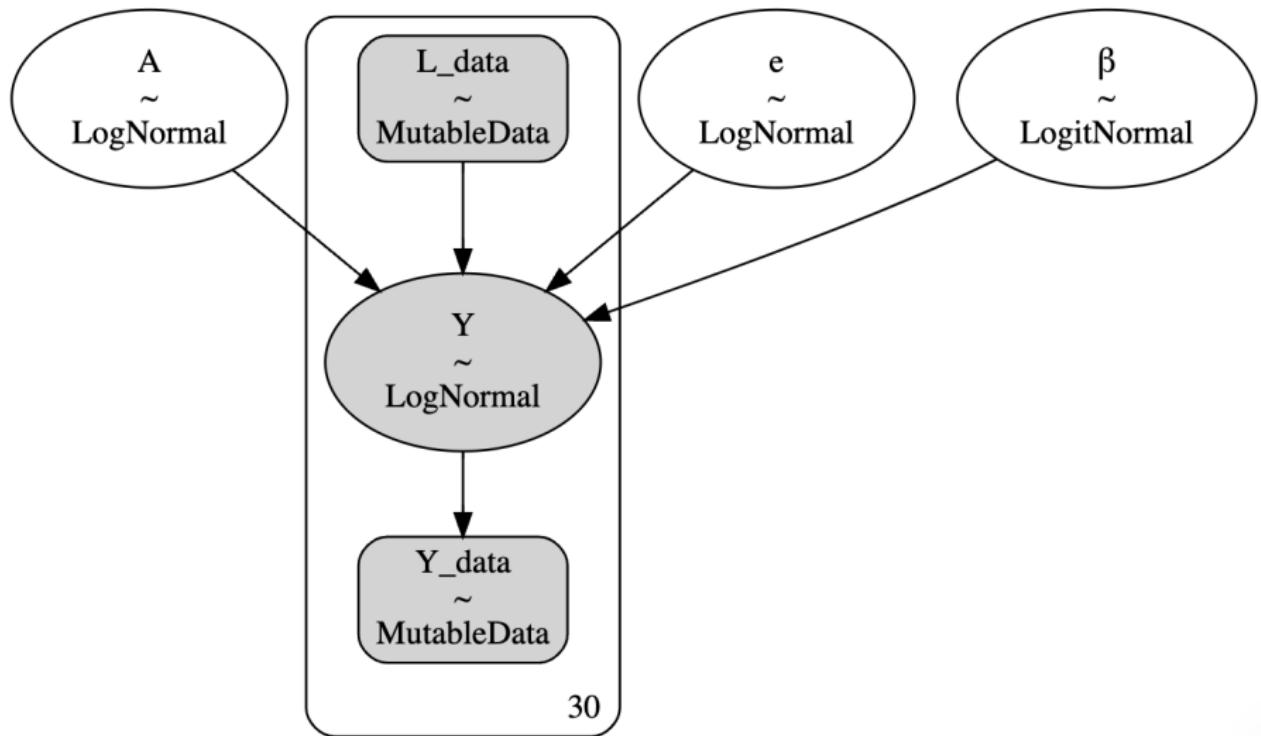
$$\epsilon \sim \text{LogNormal}(-2, 1)$$

$$\beta \sim \text{LogitNormal}(0, 1)$$

$$A_{g=0} \sim ???$$



## Визуализиурем модель





## Априорное предиктивное распределение

Подобрать априорное распределение  $\beta$  было несложно. А для  $A_{g=0}$ ?



## Априорное предиктивное распределение

Подобрать априорное распределение  $\beta$  было несложно. А для  $A_{g=0}$ ?

- Нет никакой информацией об априорном распределении



## Априорное предиктивное распределение

Подобрать априорное распределение  $\beta$  было несложно. А для  $A_{g=0}$ ?

- Нет никакой информацией об априорном распределении
- Но кое-что знаем про  $Y$



## Априорное предиктивное распределение

Подобрать априорное распределение  $\beta$  было несложно. А для  $A_{g=0}$ ?

- Нет никакой информацией об априорном распределении
- Но кое-что знаем про  $Y$
- $Y$  положительно, соответственно  $A$  тоже положительно



## Априорное предиктивное распределение

Подобрать априорное распределение  $\beta$  было несложно. А для  $A_{g=0}$ ?

- Нет никакой информацией об априорном распределении
- Но кое-что знаем про  $Y$
- $Y$  положительно, соответственно  $A$  тоже положительно
- Есть наблюдаемые значения для  $Y$ , значит можно вывести значения для  $A$ ?



## Априорное предиктивное распределение

Подобрать априорное распределение  $\beta$  было несложно. А для  $A_{g=0}$ ?

- Нет никакой информацией об априорном распределении
- Но кое-что знаем про  $Y$
- $Y$  положительно, соответственно  $A$  тоже положительно
- Есть наблюдаемые значения для  $Y$ , значит можно вывести значения для  $A$ ?
- Степени 10 подойдут как значения для  $Y$



# Априорное предиктивное распределение

Подобрать априорное распределение  $\beta$  было несложно. А для  $A_{g=0}$ ?

- Нет никакой информацией об априорном распределении
- Но кое-что знаем про  $Y$
- $Y$  положительно, соответственно  $A$  тоже положительно
- Есть наблюдаемые значения для  $Y$ , значит можно вывести значения для  $A$ ?
- Степени 10 подойдут как значения для  $Y$
- Для проверки – построим априорное предиктивное распределение



# Априорное предиктивное распределение

Подобрать априорное распределение  $\beta$  было несложно. А для  $A_{g=0}$ ?

- Нет никакой информацией об априорном распределении
- Но кое-что знаем про  $Y$
- $Y$  положительно, соответственно  $A$  тоже положительно
- Есть наблюдаемые значения для  $Y$ , значит можно вывести значения для  $A$ ?
- Степени 10 подойдут как значения для  $Y$
- Для проверки – построим априорное предиктивное распределение

## Определение

Априорное предиктивное распределение строится как симуляция распределения без учета данных.

## Горькая правда

Подбор априорных распределений – основа нашей непростой работы



# Подбор распределения "наобум"

Почему бы не предположить,  
например

$$A \sim \text{LogNormal}(0, 1)$$



# Подбор распределения "наобум"

Почему бы не предположить,  
например

$$A \sim \text{LogNormal}(0, 1)$$

## Контринтуитивный результат

Никакой работник не пошьет  
вручную 800 ковров за неделю

Потому что (см. график ниже)

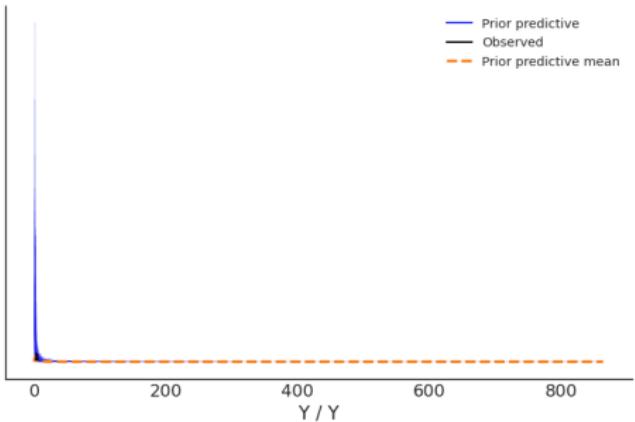


Рис.: Априорное предиктивное  
распределение  $Y$  и распределение  
реальных значений

# Анализируем априорные предиктивные распределения



Еще раз взглянем на нашу модель

$$\log Y_{g=0} \sim \text{Normal}(\log A_{g=0} + \log L \cdot \beta, \epsilon)$$

$$\epsilon \sim \text{LogNormal}(-2, 1)$$

$$\beta \sim \text{LogitNormal}(0, 1)$$

$$A_{g=0} \sim \text{LogNormal}(0, 1)$$

- Наблюдаем избыточную дисперсию
- Она может возникать из-за  $A$  или  $\epsilon$

Что делать?

- ① Попробуем уменьшить дисперсию  $A$
- ② Попробуем уменьшить дисперсию  $\epsilon$

Что получаем

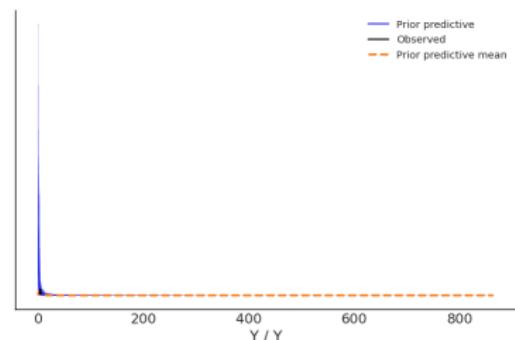


Рис.: Априорное предиктивное распределение  $Y$  и распределение реальных значений



# Пример хорошего априорного предиктивного распределения

На семинаре

Разберем схожий пример на семинаре.

Хорошее априорное предиктивное можно задать следующим образом:

$$\log Y_{g=0} \sim \text{Normal}(\log A_{g=0} + \log L \cdot \beta, \epsilon)$$

$$\epsilon \sim \text{LogNormal}(-2, 0.1)$$

$$\beta \sim \text{LogitNormal}(0, 1)$$

$$A_{g=0} \sim \text{LogNormal}(-0.5, 0.1)$$

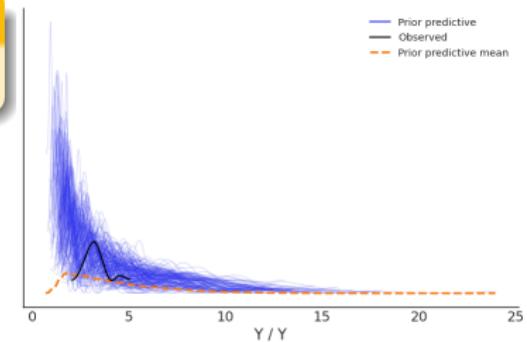


Рис.: Априорное предиктивное распределение  $Y$  и распределение реальных значений



# Как понять, что априорное предиктивное распределение удачное?

- Распределение **достаточно** соответствует данным.
- Данные – ориентир, не самоцель

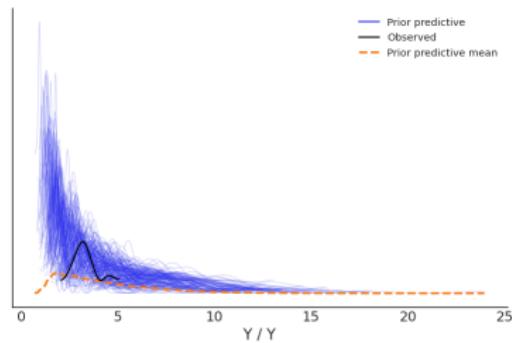


Рис.: Априорное предиктивное распределение  $Y$  и распределение реальных значений



# Как понять, что априорное предиктивное распределение удачное?

- Распределение **достаточно** соответствует данным.
  - нет астрономических скоростей
  - нет маленьких расстояний
  - нет очень больших расстояний
  - нет работников-стахановцев
- **Данные – ориентир**, не самоцель

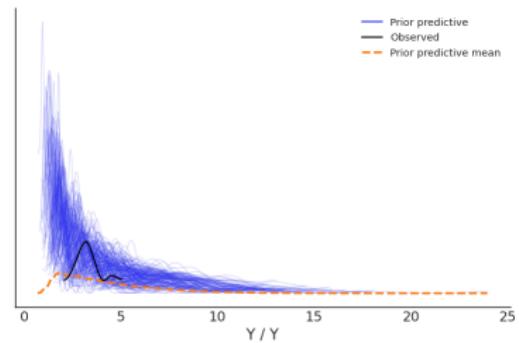


Рис.: Априорное предиктивное распределение  $Y$  и распределение реальных значений



# Как понять, что априорное предиктивное распределение удачное?

- Распределение **достаточно** соответствует данным.
  - нет астрономических скоростей
  - нет маленьких расстояний
  - нет очень больших расстояний
  - нет работников-стахановцев
- **Данные – ориентир**, не самоцель
  - не стоит подгонять распределение точь в точь под данные
  - в 90% данные не нужны
  - в 90% достаточно здравого смысла
  - в 10% можно спросить у экспертов
  - смотреть на данные – крайняя мера

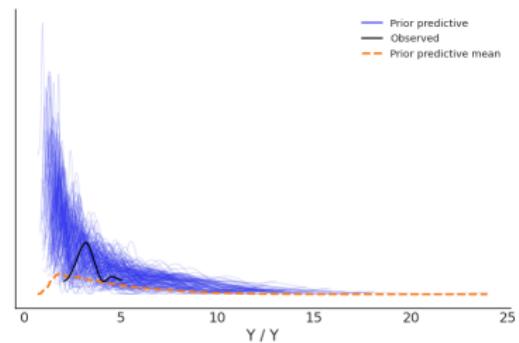


Рис.: Априорное предиктивное распределение  $Y$  и распределение реальных значений

# Пора применять НМС





# Сэмплирование

После всех  
проведенных  
проверок – можно  
начинать  
сэмплирование.

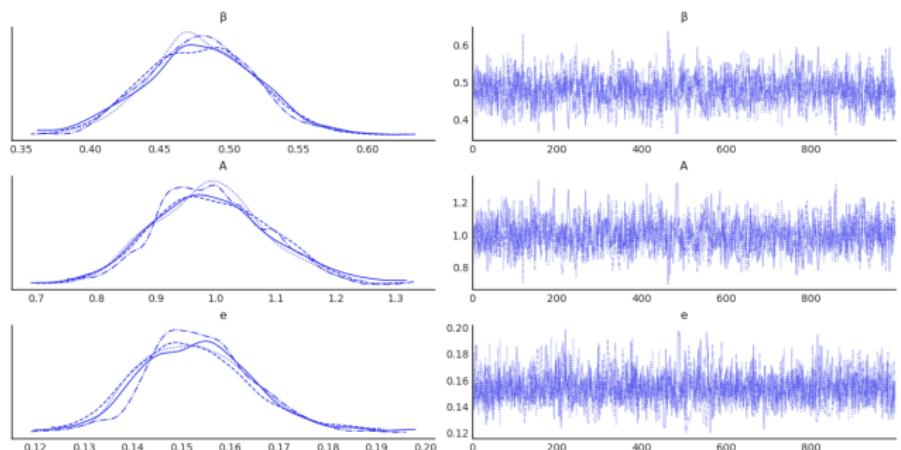


Рис.: Апостериорное распределение, полученное  
МСМС

# Иерархии



# Иерархии

В исходных данных есть группы. Как это учесть?

$$\log Y_{\mathbf{g}} \sim \text{Normal}(\log A_{\mathbf{g}} + \log L \cdot \beta, \epsilon)$$

$$\epsilon \sim \text{LogNormal}(-2, 0.1)$$

$$\beta \sim \text{LogitNormal}(0, 1)$$

$$A_{\mathbf{g}} \sim ???$$



# Что такое иерархии?

## Иерархии

Если в данных есть похожие группы, значит есть иерархии.

Примеры:

- ① Страны, Регионы
- ② Разделение пользователей по полу, возрасту и тд.
- ③ Группы воздействия
- ④ Временные эффекты
- ⑤ Панельные данные

## Наш пример

Рабочие делают разные ковры и обладают совокупной производительностью  $A$



# Как учитывать иерархии

Классический эконометрический подход:

- ① Если все группы независимы. **Регрессия пула**

$$y_{k,i} = \alpha + \beta x_{k,i} + \varepsilon_{i,k}$$

- ② Если в группах есть различия.

**Регрессия с фиксированными эффектами**

$$y_{k,i} = \alpha_k + \beta x_{k,i} + \varepsilon_{i,k}$$

- ③ Разница между группами случайная, незначительная.

**Регрессия со случайными эффектами**

$$y_{k,i} = \alpha + \beta x_{k,i} + u_k + \varepsilon_{i,k}$$

Где

$$\mathbb{E} u_{k,i} = 0, \quad \mathbb{E} \varepsilon_{i,k} = 0$$



## Байесовский подход к иерархиям

Перепишем уравнение

$$y_{k,i} = \alpha + \beta x_{k,i} + u_k + \varepsilon_{i,k}$$

Следующим образом

$$y_{k,i} = (\alpha + u_k) + \beta x_{k,i} + \varepsilon_{i,k}$$

- $\alpha$  - среднее по выборке
- $\alpha_k = \alpha + u_k$  - среднее по группе

Байесовский подход требует априорных распределений. Например

$$\alpha \sim \text{Normal}(\bar{\mu}, \bar{\sigma})$$

$$u_k \sim \text{Normal}(0, 1)$$

$$\alpha_k = \alpha + u_k \cdot \sigma$$

$$\alpha \sim \text{Normal}(\bar{\mu}, \bar{\sigma})$$

$$\alpha_k \sim \text{Normal}(\alpha, \sigma)$$



# Больше про априорные распределения

Нецентрированная параметризация   Центрированная параметризация

$$\alpha \sim \text{Normal}(\bar{\mu}, \bar{\sigma})$$

$$u_k \sim \text{Normal}(0, 1)$$

$$\alpha_k = \alpha + u_k \cdot \sigma$$

$$\alpha \sim \text{Normal}(\bar{\mu}, \bar{\sigma})$$

$$\alpha_k \sim \text{Normal}(\alpha, \sigma)$$

Group specific parameter  $u_k$  is  
disentangled

$\sigma$  отражает разницу между группами

- ①  $\sigma \rightarrow 0$ : Модель пула
- ② Небольшая  $\sigma$ : Случайные эффекты / Частичный пул
- ③ Большая  $\sigma$ : Фиксированные эффекты / Необъединенная модель

$\sigma$  присутствует в каждом типе моделей



# Деградация моделей

Для центрированной параметризации имеем

$$\alpha \sim \text{Normal}(\bar{\mu}, \bar{\sigma})$$

$$\alpha_k \sim \text{Normal}(\alpha, \sigma)$$

## Предупреждение

Для малого объема данных центрирование приводит к появлению воронок

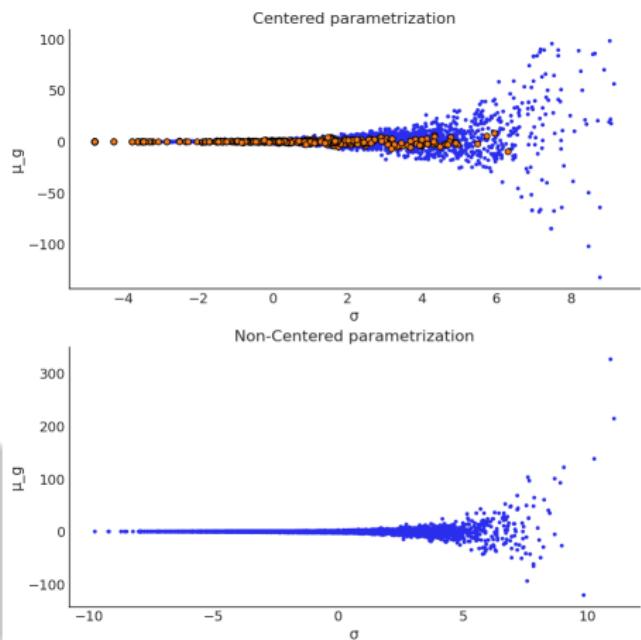


Рис.: Центрирование и расхождения



# Как появляются воронки?

Геометрия распределения имеет значение

- ① При сэмплировании шаг учитывает предыдущие итерации
- ② При неудачно выбранном распределении сэмплеру сложно предсказать следующий шаг

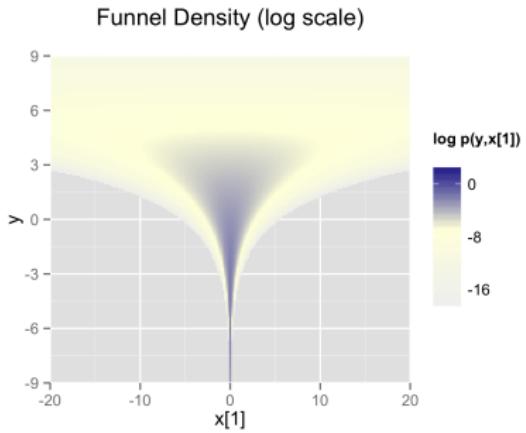


Рис.: Пример воронки

## Рекомендация

Про репараметризацию можно почитать [здесь](#)



# Инвертированная воронкообразная деградация

"Удачная" параметризация тоже имеет свои особенности

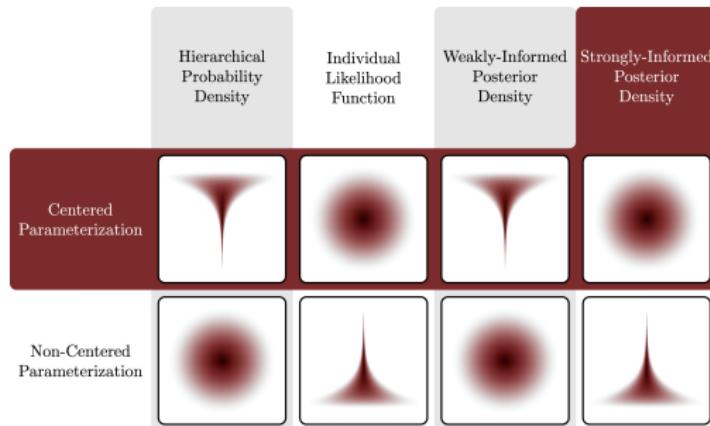


Рис.: Пример инвертированной воронкообразной деградации

Для дальнейшего погружения

Почитать про это можно [здесь](#)



# Настраиваем иерархии в априорных распределениях

- ① Начать лучше с модели с одной группой или модели пула
- ② Добавляем иерархии



# Настраиваем иерархии в априорных распределениях

- ① Начать лучше с модели с одной группой или модели пула
  - Понимаем размерность и особенности параметров
  - Структурируем модель
  - Пока не думаем о предсказательной способности модели
- ② Добавляем иерархии



# Настраиваем иерархии в априорных распределениях

- ① Начать лучше с модели с одной группой или модели пула
  - Понимаем размерность и особенности параметров
  - Структурируем модель
  - Пока не думаем о предсказательной способности модели
- ② Добавляем иерархии
  - Выбираем параметр, для которого реализуем иерархии
  - Выбираем, допустима ли вариативность остальных параметров
  - Поправляем отклонения, если нужно – повторно подбираем параметры



# Настраиваем иерархии в априорных распределениях

- ➊ Начать лучше с модели с одной группой или модели пула
  - Понимаем размерность и особенности параметров
  - Структурируем модель
  - Пока не думаем о предсказательной способности модели
- ➋ Добавляем иерархии
  - Выбираем параметр, для которого реализуем иерархии
  - Выбираем, допустима ли вариативность остальных параметров
  - Поправляем отклонения, если нужно – повторно подбираем параметры

## Практический совет

Не надо хардокдить параметризацию, делайте свой код гибким



# Пример функции Кобба-Дугласа

Модель с единственной группой

$$\log Y_0 \sim \text{Normal}(\log A_0 + \log L \cdot \beta, \epsilon)$$

$$\epsilon \sim \text{LogNormal}(-2, 0.1)$$

$$\beta \sim \text{LogitNormal}(0, 1)$$

$$A_0 \sim \text{LogNormal}(-0.5, 0.1)$$

Иерархическая модель

$$\log Y_k \sim \text{Normal}(\log A_k + \log L \cdot \beta, \epsilon)$$

$$\epsilon \sim \text{LogNormal}(-2, 0.1)$$

$$\beta \sim \text{LogitNormal}(0, 1)$$

$$A_k \sim \text{LogNormal}(\log A_{\text{pop}}, \sigma_A)$$

$$A_{\text{pop}} \sim \text{LogNormal}(-0.5, 0.1)$$

$$\sigma_A \sim \text{LogNormal}(-2, 0.1)$$



# Пример функции Кобба-Дугласа

Модель с единственной группой

$$\begin{aligned}\log Y_0 &\sim \text{Normal}(\log A_0 + \log L \cdot \beta, \epsilon) \\ \epsilon &\sim \text{LogNormal}(-2, 0.1) \\ \beta &\sim \text{LogitNormal}(0, 1) \\ A_0 &\sim \text{LogNormal}(-0.5, 0.1)\end{aligned}$$

Иерархическая модель

$$\begin{aligned}\log Y_k &\sim \text{Normal}(\log A_k + \log L \cdot \beta, \epsilon) \\ \epsilon &\sim \text{LogNormal}(-2, 0.1) \\ \beta &\sim \text{LogitNormal}(0, 1) \\ A_k &\sim \text{LogNormal}(\log A_{\text{pop}}, \sigma_A) \\ A_{\text{pop}} &\sim \text{LogNormal}(-0.5, 0.1) \\ \sigma_A &\sim \text{LogNormal}(-2, 0.1)\end{aligned}$$

## Подсказка

Некоторые параметры можно переиспользовать, нужно лишь добавить немного вариативности с помощью  $\sigma_A$



## Обсуждаем итоги

При подборе априорных распределений стоит помнить следующее

- Не всегда доступны экспертные знания
- Параметризация не всегда помогает подобрать хорошие распределения
- Иногда априорные предиктивные распределения зависят от многих параметров
- Вы часто скованы временем
- Можно применять гипераприорные распределения