Выбор априорного распределения. Распределение Джеффриса.

Линдеманн Никита

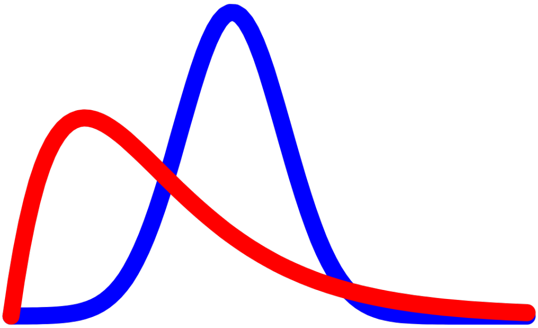
ФУПМ МФТИ
Кафедра интеллектуальных систем

30 декабря 2021 г.

Определение

Априорное распределение

Априорное распределение вероятностей случайной величины p – это распределение вероятностей, которое выражает предположения о p до учёта экспериментальных данных.



Априорное распределение

Виды априорных распределений

1. **Информативное априорное распределение**
   Информативное априорное распределение выражает конкретную информацию о переменной.
   Например, подходящим априорным распределением для температуры воздуха завтра в полдень будет нормальное распределение со средним значением, равным температуре сегодня в полдень, и дисперсией, равной ежедневной дисперсии температуры.
2. **Неинформативное априорное распределение**
   Неинформативное априорное распределение выражает размытую или общую информацию о переменной.
   Простейшим правилом назначения неинформативного априори является принцип безразличия, который назначает равные вероятности для всех возможностей.

Несобственное распределение

Несобственное априорное распределение

Несобственное априорное распределение характеризуется неотрицательной плотностью распределения, интеграл от которой по всему параетрическому прастранству расходится.

Пример

Пусть задана нормальная модель с неизвестным средним и единичной десперсией. Тогда несобственное априорное распределение для среднего:

$$q(\mu) = 1, \mu \in \mathbb{R}.$$

После одного измерения получим апостериорное распределение:

$$p(\mu \mid x) \sim p(x, \mu) \Rightarrow p(\mu \mid x) = \mathcal{N}(\mu \mid x, 1).$$

Апостериорное распределение оказалось собственным.

Распределение Джеффриса

Определение

Априорное распределение Джеффриса – это неинформативное априорное распределение, пропорциональное квадратному корню из определителя информации Фишера:

$$p(\theta) \sim \sqrt{\det\left(\mathcal{I}(\theta)\right)},$$

где $\mathcal{I}(\theta)$ – информация Фишера:

$$\mathcal{I}(\theta) = -\mathbb{E}_\theta \left(\frac{\partial^2 \log L(\theta, X)}{\partial \theta^2} \right).$$

Распределение Джеффриса

Пример с бросанием монетки

Пусть $X \sim \mathrm{Bin}(n, \theta)$, $0 \leq \theta \leq 1$ – биномиальная случайная величина с параметром θ. Тогда

$$L(x, \theta) = C_n^x \theta^x (1-\theta)^{(n-x)},$$

$$\frac{d^2 \log L}{d\theta^2} = -\frac{x}{\theta^2} - \frac{n-x}{(1-\theta)^2},$$

$$\mathbb{E}(X) = n\theta \Rightarrow \mathcal{I}(\theta) = \frac{n}{\theta(1-\theta)},$$

$$p(\theta) \sim \sqrt{\mathcal{I}(\theta)} \sim \theta^{-\frac{1}{2}} (1-\theta)^{-\frac{1}{2}} \sim Beta\left(\frac{1}{2}, \frac{1}{2}\right).$$

Свойства распределения Джеффриса

1. Априорное распределение Джеффриса не зависит от параметризации:
$$p(\theta) \sim \sqrt{\mathcal{I}(\theta)},$$
$$p(\varphi) = p(\theta) \left| \frac{\partial \theta}{\partial \varphi} \right| \sim \sqrt{\mathbb{E}\left(\frac{\partial^2 \log L}{\partial \theta^2} \right) \left(\frac{\partial \theta^2}{\partial \varphi^2} \right)} =$$
$$= \sqrt{\mathbb{E}\left(\frac{\partial^2 \log L}{\partial \varphi^2} \right)} = \sqrt{\mathcal{I}(\varphi)}.$$
2. Априорное распределение Джеффриса может быть несобственным.

Свойства распределения Джеффриса

Инвариантность к масштабу

1. Главное свойство – независимоть от параметризации (это свойство даже можно использовать в качестве определения). Оно означает, что априорное распределение Джеффриса инвариантно к масштабу случайной величины (так как во многих задачах масштаб заранее не известен).
2. В предыдущем примере с бросанием монетки именно распределение $Beta\left(\frac{1}{2}, \frac{1}{2}\right)$ обладает этим свойством инвариантности, хотя интуитивно более подходящим априорным распределением кажется равномерное. В сравнении с равномерным, $Beta\left(\frac{1}{2}, \frac{1}{2}\right)$ имеет вытянутые «хвосты».

Несобственное распределение Джеффриса

Нормальное распределение со средним μ как параметром

$$f(x \mid \mu) = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}},$$

$$p(\mu) \sim \sqrt{\mathcal{I}(\mu)} = \sqrt{\mathbb{E}\left(\frac{x-\mu}{\sigma^2}\right)^2} =$$

$$= \sqrt{\int\limits_{-\infty}^{+\infty} f(x \mid \mu)\left(\frac{x-\mu}{\sigma^2}\right)^2 dx} = \sqrt{\frac{\sigma^2}{\sigma^4}} \sim 1.$$

Априорное распределение Джеффриса для μ является несобственным равномерным распределением на $\mathbb{R}$: оно равно некоторой фиксированной константе (например, 1) для всех точек.

Несобственное распределение Джеффриса

Нормальное распределение со стандартным отклонением σ как параметром

$$f(x \mid \mu) = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}},$$

$$p(\sigma) \sim \sqrt{\mathcal{I}(\sigma)} = \sqrt{\mathbb{E}\left(\frac{(x-\mu)^2-\sigma^2}{\sigma^3}\right)^2} =$$

$$= \sqrt{\int\limits_{-\infty}^{+\infty} f(x \mid \mu)\left(\frac{(x-\mu)^2-\sigma^2}{\sigma^3}\right)^2 dx} = \sqrt{\frac{2}{\sigma^2}} \sim \frac{1}{|\sigma|}.$$

Априорное распределение Джеффриса для $\log\sigma^2$ является несобственным равномерным распределением на действительной оси (логарифмическое априорное распределение).

Несобственное распределение Джеффриса

Распределение Пуассона в стандартной параметризации λ

$$f(x \mid \lambda) = e^{-\lambda} \frac{\lambda^n}{n!},$$

$$p(\mu) \sim \sqrt{\mathcal{I}(\mu)} = \sqrt{\mathbb{E}\left(\frac{n-\lambda}{\lambda}\right)^2} =$$

$$= \sqrt{\sum_{n=0}^{+\infty} f(n \mid \lambda)\left(\frac{n-\lambda}{\lambda}\right)^2} = \sqrt{\frac{1}{\lambda}}.$$

Априорное распределение Джеффриса для $\sqrt{\lambda}$ является несобственным равномерным распределением на $\mathbb{R}_+$.

Спасибо за внимание

Использованаая литература

1. Jeffreys, H. An Invariant Form for the Prior Probability in Estimation Problems.
2. Jeffreys, H. Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences.