Introduction to Bayesian inference

2018.06.04

杜岳華

Outline

- Bayes' theorom
- Bayesian inference
- Maximum likelihood estimation
- Maximum a posteriori
- Markov chain Monte Carlo

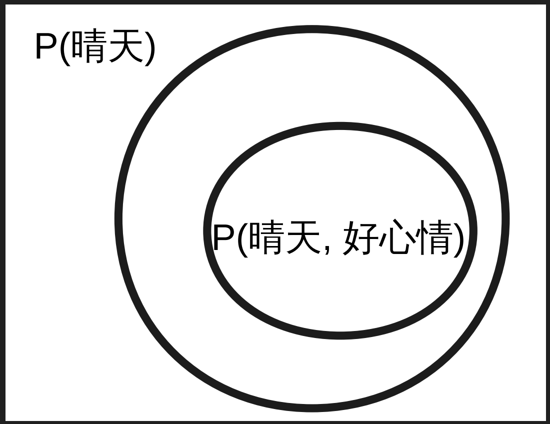
Bayesian statistics

Thomas Bayes (1701–1761)

Conditional Probability

$$P(y \mid x) = \frac{P(x,y)}{P(x)}$$

$$P(\text{好心情} \mid \text{晴天}) = \frac{P(\text{晴天,好心情})}{P(\text{晴天})}$$



Description of cause-effect

$$P(y \mid x)$$

$$x \rightarrow y$$

$$\text{因} \rightarrow \text{果}$$

$$ex.\text{晴天} \rightarrow \text{好心情}$$

Bayes' theorom

$$P(y \mid x) = \frac{P(x,y)}{P(x)}$$

$$\begin{aligned} P(x, y) &= P(x)P(y \mid x) \\ &= P(y)P(x \mid y) \end{aligned}$$

Bayes' theorom

$$P(y \mid x) = \frac{P(y)P(x|y)}{P(x)}$$

Example

檢驗方法

疾病\檢驗	陽性（+）	陰性（-）	Total
有病（T）	89	11	100
沒病（F）	9	152	161
Total	98	163	261

Example

$$疾病 \longrightarrow 檢驗$$

疾病\檢驗	陽性（+）	陰性（-）	Total
有病（T）	0.340	0.043	0.383
沒病（F）	0.035	0.582	0.617
Total	0.375	0.625	1.0

$$P(疾病, 檢驗)$$

Example

對病人來說，檢驗結果為陽性，有多少機率是有病？

$$P(\text{疾病} = T \mid \text{檢驗} = +) = ?$$

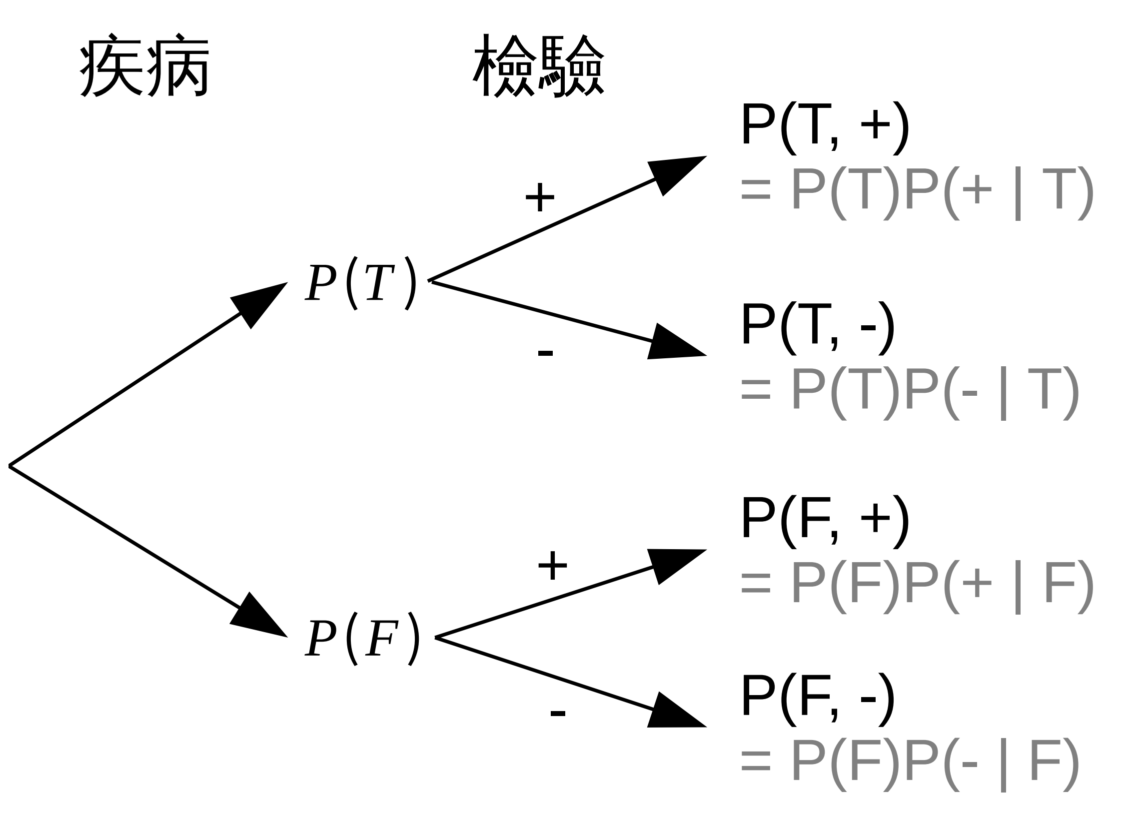
$$P(T \mid +) = \frac{P(T)P(+|T)}{P(+)}$$

$$P(T \mid +) = \frac{P(T)P(+|T)}{P(T)P(+|T)+P(F)P(+|F)}$$

Example

疾病\檢驗	陽性（+）	陰性（-）	Total
有病（T）	$P(T,+)$		$P(T)$
沒病（F）			$P(F)$
Total	$P(+)$	$P(-)$	1.0

Example



Example

$$P(T \mid +) = \frac{P(T)P(+|T)}{P(T)P(+|T)+P(F)P(+|F)}$$

$$P(T \mid +) = \frac{0.383\times P(+|T)}{0.383\times P(+|T)+0.617\times P(+|F)}$$

$$P(T \mid +) = \frac{0.340}{0.340+0.035}$$

$$P(T \mid +) = 0.907$$

Example

對病人來說，檢驗結果為陰性，有多少機率是沒病？

$P(\text{疾病} = F \mid \text{檢驗} = -)$= ?

Example

$$P(F \mid -) = \frac{P(F)P(-|F)}{P(T)P(-|T)+P(F)P(-|F)}$$

$$P(F \mid -) = \frac{0.617 \times P(-|F)}{0.383 \times P(-|T)+0.617 \times P(-|F)}$$

$$P(F \mid -) = \frac{0.582}{0.043+0.582}$$

$$P(F \mid -) = 0.931$$

Bayesian Inference

$$x \to y$$

$$P(y \mid x) = \frac{P(y)P(x|y)}{P(x)}$$

Bayesian Inference

$$P(Model \mid Data) = \frac{P(Model)P(Data|Model)}{P(Data)}$$

$$Posterior = \frac{Prior \times Likelihood}{Evidence}$$

Prior distribution

先驗機率

在看過資料或證據**之前**的假設或信念

獨立於資料或證據的

$$P(Model)$$

$$P(Model = 1), P(Model = 2) \ldots$$

Posterior distribution

後驗機率

在看過資料或證據**之後**的假設或信念
依賴資料或證據的
$$P(Model \mid Data)$$

Evidence

證據

考慮所蒐集到的證據多寡

Normalizing factor

$$P(Data)$$

Likelihood (function)

可能性

根據這些 model，有多少的可能性可以符合這些 data

$$P(Data \mid Model)$$

Maximum likelihood estimation

$$P(Data \mid Model)$$

$$\mathcal{L}(\theta) = P(X \mid \theta)$$

找出一個 model，讓符合這些 data 的可能性最大

$$Model(\text{因}) \rightarrow Data(\text{果})$$

Maximum likelihood estimation

找出一個 model，讓符合這些 data 的可能性最大

$$arg\,max_{\theta}\ \mathcal{L}(\theta)$$

Maximum likelihood estimation

實務上，為了計算方便

$$arg\,max_{\theta}\ \log \mathcal{L}(\theta)$$

Maximum a posteriori

$$P(Model \mid Data) = \frac{P(Model)P(Data|Model)}{P(Data)}$$

$$Posterior = \frac{Prior \times Likelihood}{Evidence}$$

最大化後驗機率的 model

Maximum a posteriori

跟 MLE 的差別？

Prior distribution!
希望加入人類知識進行運算

Markov chain Monte Carlo method

一種 sampling method，是一類方法的總稱

- Metropolis-Hastings method
- Gibbs sampling
- Slicing method
- Hamiltonian Monte Carlo (HMC)
- No-U-Turn Sampler (NUT, 2011)

Markov chain Monte Carlo method

我沒有要講

Q & A