Bayesian inference and data analysis
05/11/2021

Conditional probability:

$I =$ "normal 52-card deck"

$p(\text{black-card} \mid I) = \frac{26}{52} = \frac{1}{2}$

given

Ques. Probability of drawing 2 black-suit card from a pack of 52 cards?

Trial 1: $T_0 = \frac{26}{52} = \frac{1}{2}$

Trial 2: $T_1 = \frac{26-1}{52-1} = 25/51$

Probability of drawing 2 black-suit card is

$T_1 \times T_0 = \frac{25}{102}$

Mathematically,

$$P(T_0, T_1 \mid I) = \underbrace{P(T_1 \mid T_0, I)}_{= 25/51} \underbrace{P(T_0 \mid I)}_{= 1/2} = \frac{25}{102}$$

$$P(T_1 \mid T_0, I) = \frac{\text{number of black cards}}{\text{number of cards}}$$

$$= 25/51$$

Conditional probability: $P(A, B \mid I) = P(A \mid B, I)\, P(B \mid I)$

$= P(B \mid A, I)\, P(A \mid I)$

Bayes' Theorem $\Rightarrow P(A \mid B, I) = \dfrac{P(B \mid A, I)\, P(A \mid I)}{P(B \mid I)}$

Inference: figuring something from the data.

Parameter estimation

Model selection.

I. Parameter estimation: figuring out model parameters θ given some model M_A.

Let's denote data by $\bar{d}$.

$$P(\theta \mid M_A, \bar{d}) = \frac{P(\bar{d} \mid M_A, \theta)\, P(\theta \mid M_A)}{P(d \mid M_A)}$$

$P(d \mid M_A) \rightsquigarrow$ normalizes the distribution.

or, $P(\theta \mid M_A, \bar{d}) \propto P(\bar{d} \mid M_A, \theta)\, P(\theta \mid M_A)$

In general,

$$\underbrace{P(\theta \mid M_A, \bar{d})}_{\text{posterior}} \propto \underbrace{\mathcal{L}(d \mid M_A, \theta)}_{\text{Likelihood}}\, \underbrace{\pi(\theta \mid M_A)}_{\text{prior}}$$

Example: One-dimensional posterior distribution.

Consider specific data $\bar{y} = [y_0, y_1, y_2, \ldots, y_N]$ observed at times

$\bar{t} = [t_0, t_1, \ldots, t_N]$

Model : M_A :

$$y_A(t;w) = \sin(wt)$$

One model parameter, $\theta = [w]$

Jupyter notebook.

Data + Noise : Try to figure out frequency (w).

Gaussian

$$y_{obs} = y_{det} + \text{noise} \Rightarrow \text{noise} = y_{obs} - y_{det}$$

For a single data point y_i, t_i, then given a particular w

$$\mathcal{L}(y_i, t_i | w, M_A, \sigma = 0.5) = \text{Normal}(y - y_A(t;w); \sigma = 0.1)$$
$$= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - y_A(t_i))^2}{2\sigma^2}\right)$$

Then. log-likelihood

$$\ln \mathcal{L}(y_i, t_i | w, M_A, \sigma = 0.5) = -\frac{1}{2}\left(\frac{(y_i - y_A(t_i))^2}{\sigma^2} + \ln(2\pi\sigma^2)\right)$$

For all the data,

$$\mathcal{L}(\bar{y}, \bar{t} | w, M_A, \sigma = 0.5) = \underset{\text{product}}{\prod_i} \mathcal{L}(y_i, t_i | w, M_A, \sigma = 0.5)$$

$$\ln \mathcal{L}(\bar{y}, \bar{t} | w, M_A, \sigma = 0.5) = \underset{\text{sum}}{\sum_i} \ln \mathcal{L}(y_i, t_i | w, M_A, \sigma = 0.5)$$

Posteriors

$$P(w|\bar{d}) \propto \mathcal{L}(\bar{d}|w) \times \pi(w)$$

data.

- If $\pi(w) \propto$ constant, then $P(w|\bar{d}) \propto \mathcal{L}(\bar{d}|w)$
- If not, then we need to multiply by the prior.

Consider the easier case.