Bayesian Statistics

Recall: [conditional probability]

$$P(a|b) = \frac{P(b|a)\,P(a)}{P(b)}$$

we write:

$$p(a|b) \propto p(b|a)\,p(a)$$

This stems from

(1) Product rule: $p(b|a)\,p(a) = p(a,b)$

(2) Marginilization: if we model joint distribution $p(a,b)$ we compute $\sum_a p(a,b)$ or $\int_a p(a,b)$

(3) Normalization: If we can model the numerator of Bayes

we compute the left-hand side, by normalizing the values of the numerator over all possible values of α

By (1) + (2) $\quad z = p(b) = \sum_{\alpha} p(b|\alpha)\, p(\alpha)$

$$\Rightarrow \quad p(\alpha|b) = \frac{1}{z}\, p(b|\alpha)\, p(\alpha)$$

if z is fixed $\Rightarrow$

$$p(\alpha|b) \propto p(b|\alpha)\, p(\alpha)$$

Given dataset $D = \{y^{(i)}\}_{i=1}^{n}$ and ϑ to be out parameter of interest

$$p(\vartheta|D) = \frac{p(D|\vartheta) \cdot p(\vartheta)}{p(D)}$$

$p(D) \rightarrow$ fixed