Bayesian inference

likelihood function for the parameters

prior probability

$$P(\Theta \mid \mathbf{x}) = \frac{P(\mathbf{x} \mid \Theta) P(\Theta)}{P(\mathbf{x})}$$

posterior probability

normalizer, often is not of interest

Simple example*

- What you know
  - John is coughing
- What do you conclude?
  - John has a cold
  - John has lung cancer
  - John has stomach problems

*Adopted from Josh Tenenbaum

Why this approach

Separates representation, modeling, and inference

Model is separated into prior and likelihood

Handles fitting and learning similarly

What is known is always represented as a distribution

Probability review

Formulas that you should be very comfortable with are marked by *.

Interpretations of probability

1)Representation of expected frequency
2)Degree of belief

Basic terminology and rules

Space of outcomes (often denoted by Ω)

Event (subset of Ω)

Set of measurable events, S

$$a \in S, \quad P(a) \in [0,1]$$

$$P(a \cup b) = p(a) + p(b) \quad \text{if } a \cap b = \varnothing$$

$$P(\Omega) = 1$$

From this we get the usual fact that probabilities over disjoint sets that cover P(Ω) sum to 1.

Basic terminology and rules

Conditional probability (definition)

$$P(A|B) \equiv \frac{P(A \cap B)}{P(B)} \qquad *$$

Chain (Product) Rule

$$P(A_1 \cap A_2) = P(A_1)P(A_2|A_1) \qquad *$$

$$P(A_1 \cap A_2 \cap \;\ldots\; A_N) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \;\ldots\; P(A_N|A_1 \cap A_2 \cap \;\ldots\; A_{N-1})$$

Bayes Rule

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \qquad *$$

Example (continued)

Probability of disease given symptoms

Suppose a TB test is 95% accurate
Suppose that TB is in 0.1% of population

What is P (TB | positive)?

$$P(TB \mid positive)$$

$$= \frac{P(positive \mid TB)P(TB)}{P(positive)}$$

$$= \frac{P(positive \mid TB)P(TB)}{P(positive \mid TB)P(TB) + P(positive \mid \widetilde{TB})P(\widetilde{TB})}$$

$$= \frac{(0.95)(0.001)}{(0.95)(0.001) + (0.05)(0.999)}$$

$$\cong 0.0187$$