Bayesian Statistics: basics and elementary examples

Václav Šmídl

February 20, 2018

Course layout

Aim: practical use of general methodology

- minimum level of formality

Literature: Bishop, Ch.M. Pattern recognition and machine learning. Springer, 2006.

Details: scholar.google.com

Marks: individual assignment

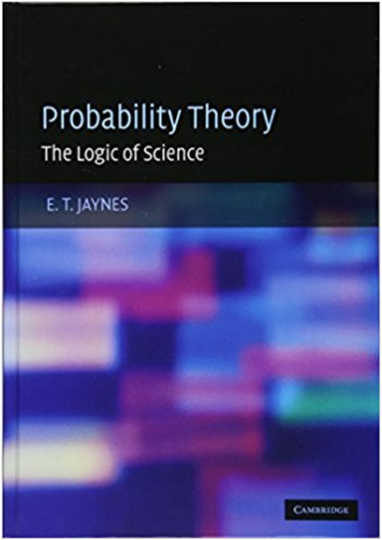
Probability theory

Probability theory as:

- extension of logic
- language
- necessity for making decisions under uncertainty

Alternatives:

- Dempster Shafer
- fuzzy logic



Example

Probability of patients illness:

illness	Dr. #1 likelihood	Dr.#2 likelihood	posterior
brain tumor	99.9%	0%	
concussion	0%	99.9%	
meningitis	0.1%	0.1%	

Example

Probability of patients illness:

illness	Dr. #1 likelihood	Dr.#2 likelihood	posterior
brain tumor	99.9%	0%	0%
concussion	0%	99.9%	0%
meningitis	0.1%	0.1%	100%

Example

Probability of patients illness:

illness	Dr. #1 likelihood	Dr.#2 likelihood	posterior
brain tumor	99.9%	0%	0%
concussion	0%	99.9%	0%
meningitis	0.1%	0.1%	100%

Counterintuitive results

1. Is probability theory wrong?

Example

Probability of patients illness:

illness	Dr. #1 likelihood	Dr.#2 likelihood	posterior
brain tumor	99.9%	0%	0%
concussion	0%	99.9%	0%
meningitis	0.1%	0.1%	100%

Counterintuitive results

1. Is probability theory wrong?
2. Is model wrong?

Probability calculus discrete

Random variables:

$$X \in \{x_1, \dots, x_M\}$$
$$Y \in \{y_1, \dots, y_L\}$$

Joint probability

$$P(X = x_i, Y = y_j) = \frac{n_{i,j}}{N}$$

where N ($N \to \infty$) is the number of realizations and $n_{i,j}$ is the number of trials where $X = x_i$, $Y = y_j$.
Rules:

1. sum rule
$$P(X = x_i) = \sum_{j=1}^{L} P(X = x_i, Y = y_i),$$
2. product rule
$$P(X, Y) = p(Y|X)p(X)$$

Probability calculus continuous

Random variable: X, realization x
Probability density function: $p_X(x)$,

Probability calculus continuous

Random variable: X, realization x
Probability density function: $p_X(x)$, $\int p_X(x)dx = 1$

Probability calculus continuous

Random variable: X, realization x
Probability density function: $p_X(x)$, $\int p_X(x)dx = 1$
Cumulative density function: $P(X < x) = \int_{-\infty}^{x} p_X(t)dt$

Probability calculus continuous

Random variable: X, realization x
Probability density function: $p_X(x)$, $\int p_X(x)dx = 1$
Cumulative density function: $P(X < x) = \int_{-\infty}^{x} p_X(t)dt$
Rules:

1. sum rule
$$p(x) = \int p(x, y)dy$$
2. product rule
$$p(x, y) = p(y|x)p(x) = p(x|y)p(y)$$

Probability calculus continuous

Random variable: X, realization x
Probability density function: $p_X(x)$, $\int p_X(x)dx = 1$
Cumulative density function: $P(X < x) = \int_{-\infty}^{x} p_X(t)dt$
Rules:

1. sum rule
$$p(x) = \int p(x, y)dy$$
2. product rule
$$p(x, y) = p(y|x)p(x) = p(x|y)p(y)$$

Engineering (ML) notation: meaning given by context.

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}$$

Toy problem

Noisy observation:

$$d = m + e,$$

What is the distribution of m given d?

Toy problem

Noisy observation:

$$d = m + e,$$

What is the distribution of m given d?

$$p(m|d) \propto p(m, d)$$

Toy problem

Noisy observation:

$$d = m + e,$$

What is the distribution of m given d?

$$p(m|d) \propto p(m, d) = \int p(d|m, \omega) p(m, \omega) d\omega$$

Toy problem

Noisy observation:

$$d = m + e,$$

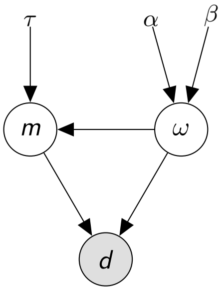
What is the distribution of m given d?

$$p(m|d) \propto p(m, d) = \int p(d|m, \omega) p(m, \omega) d\omega$$

Noise model (likelihood): Gaussian

$$p(e_i) = \mathcal{N}(0, \omega^{-1}) = \sqrt{\frac{\omega}{2\pi}} \exp\left(-\frac{1}{2}\omega e^2\right)$$

$$p(d_i) = \mathcal{N}(m, \omega^{-1}) = \sqrt{\frac{\omega}{2\pi}} \exp\left(-\frac{1}{2}\omega (d - m)^2\right)$$



Prior:

$$p(m|\omega) = \mathcal{N}(0, \tau\omega) = \sqrt{\frac{\omega\tau}{2\pi}} \exp\left(-\frac{1}{2}\tau\omega m^2\right)$$

$$p(\omega) = G(\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \omega^{\alpha-1} \exp(-\beta\omega)$$

Toy problem

Noisy observation:

$$d = m + e$$

What is the distribution of m given d?

$$p(m|d) \propto p(m, d) = \int p(d|m, \omega) p(m, \omega) d\omega$$

Toy problem

Noisy observation:

$$d = m + e$$

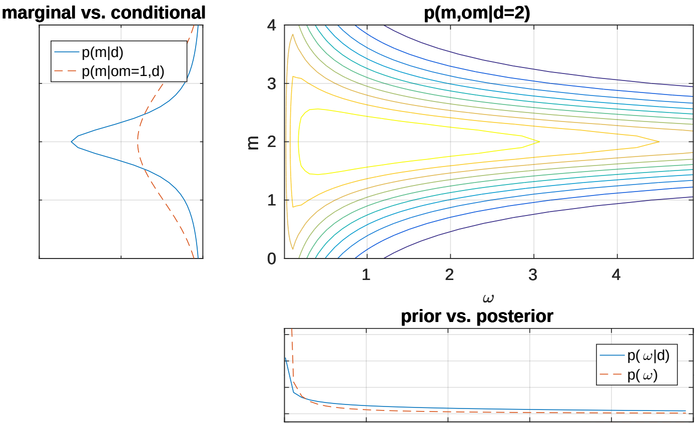
What is the distribution of m given d?

$$p(m|d) \propto p(m,d) = \int p(d|m,\omega)p(m,\omega)d\omega$$

Find:

$$\begin{aligned} p(m,\omega|d) &=? \\ p(m|\omega,d) &=? \\ p(m|d) &=? \\ p(\omega|d) &=? \end{aligned}$$

Results: $d = 2, \alpha = 0.1, \beta = 0.1, \tau = 0.0001$



Choices:

Likelihood:

1. Gaussian
2. Other

Prior:

- informative
- non-informative

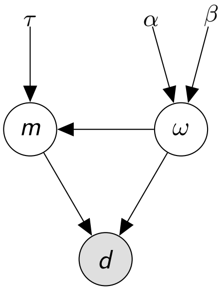
Choices:

Likelihood:

1. Gaussian
2. Other

Prior:

- informative
- non-informative



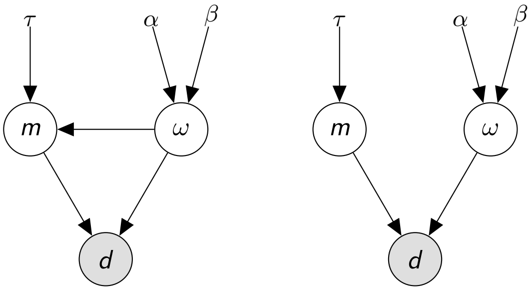
Choices:

Likelihood:

1. Gaussian
2. Other

Prior:

- informative
- non-informative



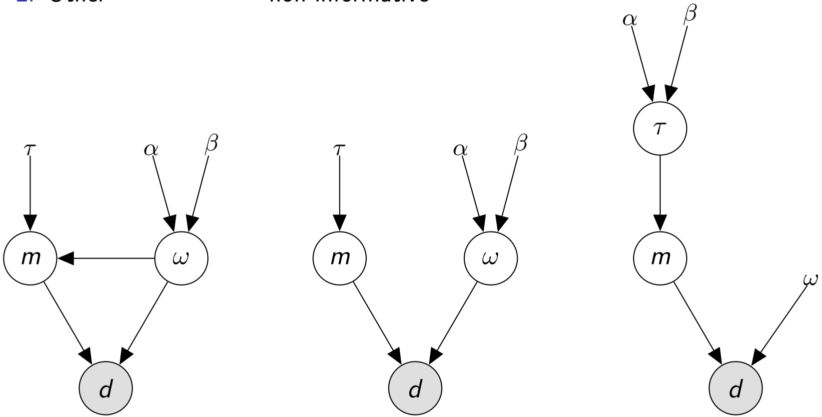
Choices:

Likelihood:

1. Gaussian
2. Other

Prior:

- informative
- non-informative



Alternative model

- $p(m|\omega) = \int p(d|m,\omega)p(m|\alpha)p(\alpha)d\alpha$
- graphical model