SIGNAL & DATA ANALYSIS IN NEUROSCIENCE

ESTIMATION

biu.sigproc@gmail.com

Agenda

- MLE & MAP – Terms and definitions
- Samples questions

Estimation methods

- **Maximum a posteriori estimator (MAP):**

$$\hat{\Theta} = \text{argmax } P(\Theta|x)$$

$$p(\Theta|x) = \frac{p(x|\Theta)\cdot p(\Theta)}{p(x)}$$

$$\hat{\Theta} = \text{argmax } P(x|\Theta)\cdot p(\Theta)$$

- **Bayesian approach:**
  Given cost function

$$C(\epsilon = \Theta - \hat{\Theta}) \text{ and } x:$$

$$\hat{\Theta} = \text{argmin}_{\hat{\Theta}} \int_{\Theta} C(\Theta, \hat{\Theta}) \cdot p(\Theta|x) d\Theta$$

Estimation methods

- **Maximum Likelihood Estimator (MLE):**
  - Estimating model parameters given an observation – similar to MAP without the prior…
  - For large n, prior is less relevant an MLE can be used.

$$
\begin{aligned}
\hat{\Theta} &= \text{argmax } P(\Theta|x) \\
p(\Theta|x) &= \frac{p(x|\Theta) \cdot p(\Theta)}{p(x)} \\
\hat{\Theta} &= \text{argmax } P(x|\Theta)
\end{aligned}
$$

Example MAP: Exam 2007

- The time difference (in weeks) between occurrences of amnesia in Foergetis Homeworkis patients may generally be estimated by an exponential distribution:

  $$p(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

  The parameter λ describing the distribution is different for each patient. The five observations of amnesia in patient X are spaced by 2, 1, 5, 1, 3 weeks.

  Assuming the prior distribution $P_0(\lambda) = ½\,\lambda$, $(0 \leq \lambda \leq 2)$

  known for the general patient population, what is the maximum a-posteriori (MAP) estimator?

Solution

Given: p(x) = λ exp(-λ x), $\underline{X}$ = [1, 1, 2, 3, 5], p(λ)=0.5(λ)

Find: MAP for λ

$$\begin{aligned}
\hat{\lambda} &= argmax_{\lambda} P(\lambda|\underline{X}) \text{ / MAP} \\
P(\lambda|\underline{X}) &= \frac{P(\underline{X}|\lambda)\cdot P(\lambda)}{P(\underline{X})} \text{ /Bayes} \\
F = \prod_i P(x_i|\lambda)\cdot P(\lambda) &= \lambda^5 \exp^{-12\lambda} \cdot \frac{\lambda}{2} = \frac{1}{2}\lambda^6 \exp^{-12\lambda} \text{ / } x_i \text{ iid} \\
\frac{dF}{d\lambda} &= \frac{1}{2}\lambda^5 \exp^{-12\lambda} [6 - 12\lambda] \\
\lambda_1 = 0 & \quad P(\lambda_1|\underline{X}) = 0 \\
\lambda_2 = \frac{1}{2} & \quad P(\lambda_2|\underline{X}) > 0 \text{ and max point,} \\
\text{Hence, } \hat{\lambda} &= \frac{1}{2}
\end{aligned}$$

Example MLE: Exam 2005

- During 5 trials the recorded rates of the neuron were: 70, 62, 95, 59, 65. Assuming that the recorded rate is the result of a Poisson distribution, find the MLE for the parameter λ.

Solution

Given: X~ Poisson distribution. X = [59, 62, 65, 70, 95]

Find: MLE λ

Solutio

$$
\begin{aligned}
\hat{\lambda} &= \mathrm{argmax}_{\lambda} P(\underline{X}; \lambda) \ / \ \mathrm{MLE} \\
f(x_i; \lambda) &= \frac{\lambda^{x_i} exp^{-\lambda}}{x_i!} \ / \ \text{Piosson distribution} \\
f(\underline{X}; \lambda) &= \prod_i \frac{\lambda^{x_i} exp^{-\lambda}}{x_i!} \\
ln(f(\underline{X}; \lambda)) &= \sum_i ln\left(\frac{\lambda^{x_i} exp^{-\lambda}}{x_i!}\right) \\
F = \sum_i ln(\lambda^{x_i} exp^{-\lambda}) &= \sum_i x_i ln(\lambda) - \lambda \cdot N \\
\frac{dF}{d\lambda} &= \frac{\sum_i x_i}{\lambda} - N = 0 \\
\lambda &= \frac{\sum_i x_i}{N} \ / \ \text{max point} \\
\hat{\lambda} &= E[x_i]
\end{aligned}
$$

Example: MAP +MLE exam 2006

- The time difference (in hours) between occurrences of hiccups in Hiccupitis Neuralitis Syndrome patients may be estimated by the following uniform distribution:
  $p(x) = 1/n \qquad 0 \le x \le n$
  $p(x) = 0 \qquad x > n$
  The parameter n describing the distribution is different for each patient. The five observations of hiccups in patient X are spaced by 2, 3, 7, 2, 1 hours.
- Find the maximum likelihood (ML) estimator for n in patient X and explain the results.
- Assuming the priors $P_0(7)=0.1$ & $P_0(8)=0.9$ known for the general patient population. Is the maximum aposteriori (MAP) estimator different from the ML estimator calculated in section (a)?

Solution

Given: $P(x_i) = 1/n$, X = [1, 2, 2, 3, 7]

Find: MLE for n

Solution:

$$\begin{aligned} \hat{n} &= \text{argmax}_n P(\underline{X}|n) \\ &= \prod_i P(x_i|n) = \left(\frac{1}{n}\right)^N \end{aligned}$$

(i) $n \geq 7$

(ii) Minimal n will maximize $P(\underline{X}|n)$

n = 7

Solution cont

Given prior for n, find MAP for n

Solution:

$$
\begin{aligned}
\hat{n} &= \text{argmax}_n P(n|\underline{X}) \\
P(n|\underline{X}) &= \frac{P(\underline{X}|n) \cdot P(n)}{P(\underline{X})} \\
F = P(\underline{X}|n) \cdot P(n) &= \prod_i P(x_i|n) \cdot P(n) \\
F &= \left(\frac{1}{n_0}\right)^N \cdot P(n = n_0) \\
F(n=7) &= \left(\frac{1}{7}\right)^5 \cdot 0.1 = 6e^{-6} \\
F(n=8) &= \left(\frac{1}{8}\right)^5 \cdot 0.9 = 2.7e^{-5} \\
\hat{n} &= 8
\end{aligned}
$$