Adnane
Ezouhri

01351331

Machine Learning
HW4

Problem 1

$$p(x|\lambda) = \begin{cases} \lambda \exp(-\lambda x), & x > 0 \\ 0, & x < 0 \end{cases}$$

$$p(\lambda) = \begin{cases} \alpha \exp(-\alpha \lambda), & \lambda > 0 \\ 0, & \lambda < 0 \end{cases}$$

a)

Likelihood:

$$L(\lambda) = \prod_{i=1}^{N} \lambda e^{-\lambda x_i}$$

take log:

$$\ln(L(\lambda)) = \sum_{i=1}^{N} \ln(\lambda) - \lambda x_i$$

derivative and set to 0:

$$\frac{d}{d\lambda} \ln(L(\lambda)) = 0 \Leftrightarrow \frac{N}{\lambda} - \sum_{i=1}^{N} x_i = 0 \Leftrightarrow \boxed{\lambda_{MLE} = \frac{N}{\sum_{i=1}^{N} x_i}}$$

b) MAP estimate of λ:

- Proportional to prior:

$$\underbrace{p(\lambda | x_1 \dots x_N)}_{A} \propto L(\lambda)\, p(\lambda)$$

- log step:

$$\ln(A) = \ln(L(\lambda)) + \ln(p(\lambda))$$

We then have:

$$\underbrace{\ln(A)}_{B} = \sum_{i=1}^{N} \ln(\lambda) - \lambda x_i + \ln(\alpha) - \alpha\lambda$$

· Take derivative and set it to 0:

$$\frac{d}{d\lambda} B = 0 \Leftrightarrow \frac{N}{\lambda} - \sum_{i=1}^{N} x_i - \alpha = 0 \Leftrightarrow \boxed{\hat{\lambda}_{MAP} = \frac{N}{\sum_{i=1}^{N} x_i + \alpha}}$$

Problem 2

$$p(x|\theta) = \begin{cases} \frac{1}{\theta} & , \ 0 < x < \theta \\ 0 & , \ \text{else} \end{cases}$$

a)

$$L(\theta) = \prod_{i=1}^{N} p(x_i|\theta) = \prod_{i=1}^{N} \frac{1}{\theta} \boxed{= \left(\frac{1}{\theta}\right)^N}$$

However θ must be greater than the max observat°, x_{max}. To ensure that all the values are in range.

Constraint: $\theta \geq x_{max}$

b) The joint proba is maximized when θ is minimized, Therefore,

$$\boxed{\hat{\theta}_{MLE} = x_{max}}$$

Problem 3

d)

$$MSE_{\theta \text{ Sample Mean}} = \frac{\theta}{N} = \frac{3^2}{10} \boxed{= 0,9}$$

Yes, I think that the unbiased estimator yields the lowest MSE.

e)

In situations where there is a significant risk that the model might overfit due to a large variance in the estimator, increasing the value of λ can be beneficial for reducing the anticipated error.