JHU Engineering for Professionals
Applied and Computational Mathematics
Data Mining 625.740

Homework for Module 2

1. Let x have an exponential distribution

$$p(x|\vartheta) = \begin{cases} \vartheta e^{-\vartheta x}, & x \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

   (a) Sketch $p(x|\vartheta)$ versus x for a fixed value of the parameter ϑ.
   (b) Sketch $p(x|\vartheta)$ versus $\vartheta, \vartheta > 0$ for a fixed value of x.
   (c) Suppose that n samples $x_1, \ldots, x_n$ are drawn independently according to $p(x|\vartheta)$. Show that the maximum likelihood estimate for ϑ is given by

$$\hat{\vartheta} = \frac{1}{\frac{1}{n}\sum_{k=1}^{n} x_k}.$$

2. Let x have a uniform distribution

$$p(x|\vartheta) = \begin{cases} \frac{1}{\vartheta}, & 0 \leq x \leq \vartheta \\ 0, & \text{otherwise.} \end{cases}$$

   (a) Sketch $p(x|\vartheta)$ versus ϑ for an arbitrary value of x.
   (b) Suppose that n samples $x_1, \ldots, x_n$ are drawn independently according to $p(x|\vartheta)$. Show that the maximum likelihood estimate for ϑ is $\max_k x_k$.
   (c) Find the method of moments estimator for ϑ.

3. Let $\mathbf{x}$ be a binary $(0, 1)$ vector with multivariate Bernouli distribution

$$p(\mathbf{x}|\boldsymbol{\vartheta}) = \prod_{i=1}^{d} \vartheta_i^{x_i}(1 - \vartheta_i)^{1-x_i},$$

   where $\boldsymbol{\vartheta} = (\vartheta_1, \ldots, \vartheta_d)^T$ is an unknown parameter vector, ϑ_i being the probability that $x_i = 1$. Show that the maximum likelihood estimate for $\boldsymbol{\vartheta}$ is

$$\hat{\boldsymbol{\vartheta}} = \frac{1}{n}\sum_{k=1}^{n} \mathbf{x}_k.$$

4. Let x have a Gamma distribution

$$p(x|\alpha, \beta) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}, & x > 0 \text{ and } \alpha, \beta > 0 \\ 0, & \text{otherwise.} \end{cases}$$

   (a) Suppose that n samples $x_1, \ldots, x_n$ are drawn independently according to $p(x|\alpha, \beta)$. Find the method of moments estimator for α and β.
   (b) Show that the exponential distribution is $\Gamma(1, 1/\vartheta)$.