JHU Engineering for Professionals
Applied and Computational Mathematics
Data Mining 625.740

Module 2 Homework Solutions

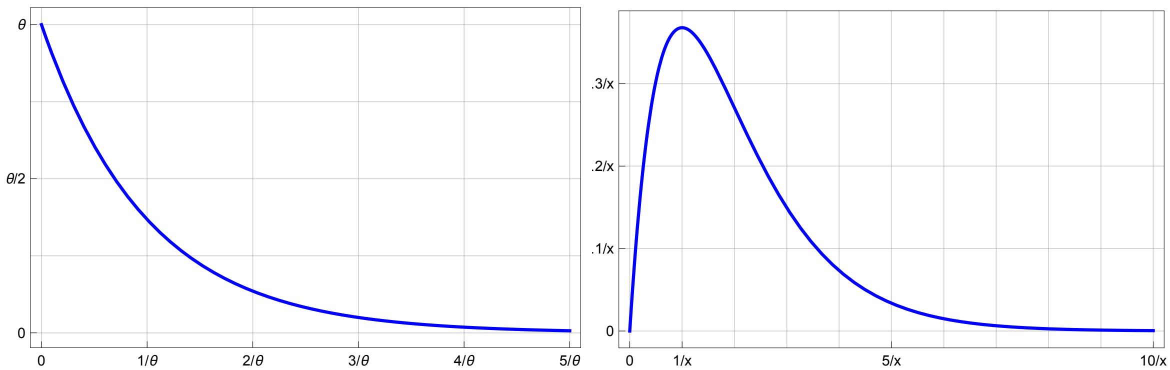

Figure 1. (a) The sketch of $p(x|\vartheta)$ versus x for fixed ϑ and (b) the sketch of $p(x|\vartheta)$ versus ϑ for fixed x.

1. The random variable x has an exponential distribution

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

   (a) The sketch of $p(x|\vartheta)$ versus x is shown in figure 1(a).

   (b) The sketch of $p(x|\vartheta)$ versus ϑ is shown in figure 1(b). The maximum of $p(x|\vartheta)$ is at $\vartheta_1 = \frac{1}{x}$ and there is an inflection point at $\vartheta_2 = \frac{2}{x}$. The second derivative, $\frac{\partial^2 p}{\partial \vartheta^2}$, equals zero and has opposite signs to the left and right of ϑ_2.

   (c) The log-likelihood function is $l_n(\vartheta) = n \log \vartheta - \vartheta \sum_k x_k$. Differentiating with respect to ϑ and setting the derivative equal to zero, we have

$$0 = \frac{dl_n}{d\vartheta} = \frac{n}{\vartheta} - \sum_k x_k,$$
$$\hat{\vartheta} = \frac{n}{\sum_k x_k}.$$

   The second derivative is strictly negative, so $\hat{\vartheta}$ is a maximum.

$$\frac{d^2 l_n}{d\vartheta^2} = -\frac{n}{\vartheta^2} < 0.$$

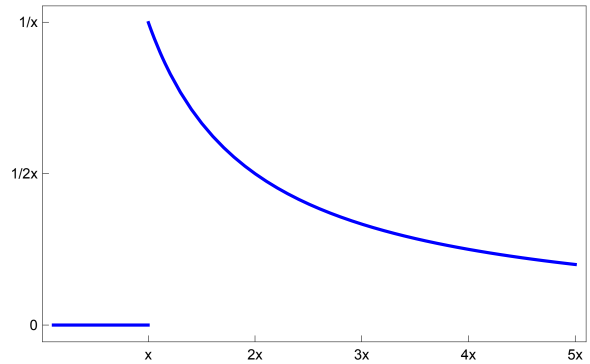

Figure 2. For the uniform distribution, the sketch of $p(x|\vartheta)$ versus ϑ for fixed x.

2. The random variable x has uniform distribution

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

   (a) The sketch of $p(x|\vartheta)$ versus ϑ is shown in figure 2.

   (b) The likelihood function,

$$\mathcal{L}(\vartheta) = \begin{cases} \frac{1}{\vartheta^n}, & 0 \le \max_k x_k \le \vartheta \\ 0, & \text{otherwise,} \end{cases}$$

   achieves its maximum at $\vartheta = \max_k x_k$. Thus, this is the maximum likelihood estimate.

   (c) The expected value of x is $E(x) = \frac{1}{\vartheta}\int_0^\vartheta x\, dx = \frac{\vartheta}{2}$ so the method of moments estimator for ϑ is $\hat{\vartheta}_{\text{MME}} = 2\overline{x} = \frac{2}{n}\sum_{k=1}^n x_k$.

3. The random variable $\mathbf{x}$ is a binary $(0,1)$ vector with multivariate Bernouli distribution

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

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

$$\begin{aligned} \mathcal{L}(\boldsymbol{\vartheta}) &= \prod_{k=1}^n \prod_{i=1}^d \vartheta_i^{x_i}(1-\vartheta_i)^{1-x_i} \\ &= \prod_{i=1}^d \vartheta_i^{\sum_k x_{ik}}(1-\vartheta_i)^{n-\sum_k x_{ik}} \end{aligned}$$

The log-likelihood $\ell_n(\boldsymbol{\vartheta})$ is

$$\ell_n(\boldsymbol{\vartheta}) = \sum_{i=1}^d \left[\sum_k x_{ik} \log\theta_i + (n - \sum_k x_{ik}) \log(1-\vartheta_i) \right].$$

Taking the derivative of the log-likelihood with respect to ϑ_i,

$$\frac{\partial \ell_n(\boldsymbol{\vartheta})}{\partial \vartheta_i} = \frac{\sum_k x_{ik}}{\vartheta_i} - \frac{n - \sum_k x_{ik}}{1 - \vartheta_i}.$$

Setting $\frac{\partial \ell_n(\boldsymbol{\vartheta})}{\partial \vartheta_i} = 0$, results in

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

4. The random variable x has 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) The expectation of each sample is

$$E(x_1) = \int_0^\infty x p(x|\alpha,\beta)\,dx = \frac{\beta}{\Gamma(\alpha)}\int_0^\infty \zeta^\alpha e^{-\zeta}\,d\zeta = \frac{\beta\Gamma(\alpha+1)}{\Gamma(\alpha)} = \alpha\beta.$$

The second moment is given by

$$E(x_1^2) = \int_0^\infty x^2 p(x|\alpha,\beta)\,dx = \frac{\beta^2}{\Gamma(\alpha)}\int_0^\infty \zeta^{\alpha+1} e^{-\zeta}\,d\zeta = \frac{\beta^2\Gamma(\alpha+2)}{\Gamma(\alpha)} = (\alpha+1)\alpha\beta^2.$$

We can find an expression for the variance:

$$\hat{\sigma}^2 = \mathrm{Var}(x_1) = E(x_1^2) - (E(x_1))^2 = \alpha\beta^2.$$

By the method of moments, we have two equations in two unknowns:

$$\alpha\beta = \overline{x},$$
$$\alpha\beta^2 = \hat{\sigma}^2.$$

The solution is

$$\alpha = \frac{\overline{x}^2}{\hat{\sigma}^2}, \quad \beta = \frac{\hat{\sigma}^2}{\overline{x}}.$$

(b) Replacing α by 1 and β by $1/\vartheta$, yields

$$\Gamma\left(1, \frac{1}{\vartheta}\right) \sim \theta e^{-x\vartheta}, \quad (x \geq 0),$$

which is the exponential distribution.