ENGG 5202: Assignment #1

Due on Thursday, March 5, 2015

Kai Chen

Problem 1

The likelihood of θ_1

$$p(D_1|\theta_1) = p(x_1|\omega_1, \theta_1)p(x_2|\omega_1, \theta_1)$$

The log-likelihood

$$l(\theta_1) = \log p(x_1|\omega_1, \theta_1) = \log p(x_1|\omega_1, \theta_1) + \log p(x_2|\omega_1, \theta_1)$$

We determine θ_1 by maximizing $l(\theta_1)$

$$\hat{\theta_1} = \arg\max l(\theta_1)$$

Let

$$\nabla l(\theta_1) = 0$$

By substituting symbols with numeral values

$$\begin{aligned} l(\theta_1) &= 2\log\frac{2}{\theta_1} + \log(1-\frac{2}{\theta_1}) + \log(1-\frac{5}{\theta_1}) \\ \nabla l(\theta_1) &= -\frac{4}{\theta_1} + \frac{4}{\theta_1(\theta_1-2)} + \frac{10}{\theta_1(\theta_1-5)} \end{aligned}$$

We get

$$\theta_1 = 8 \quad or \quad \theta_1 = 2.5$$

However, $p(x|\omega_1) = 0$ when $x > \theta_1$ according to the densities form, if $\theta_1 = 2.5$, then $D_1 = \{2, 5\}$ will not occur. Thus $\theta_1 = 8$.
Similarly, we can calculate $\theta_2 \approx 14.2$

Problem 2

2.1

Figure 1 shows $p(x|\theta)$ versus x for $\theta = 1$.

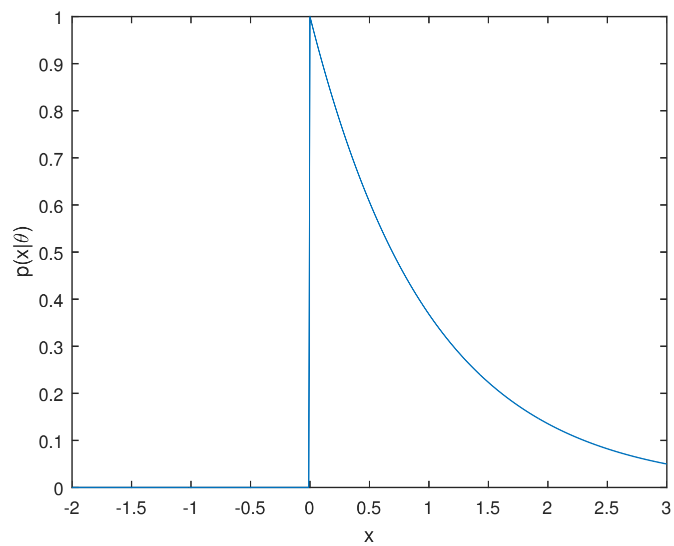

Figure 1: $p(x|\theta)$ versus x for $\theta = 1$

Figure 2 shows $p(x|\theta)$ versus θ for $x = 2$.

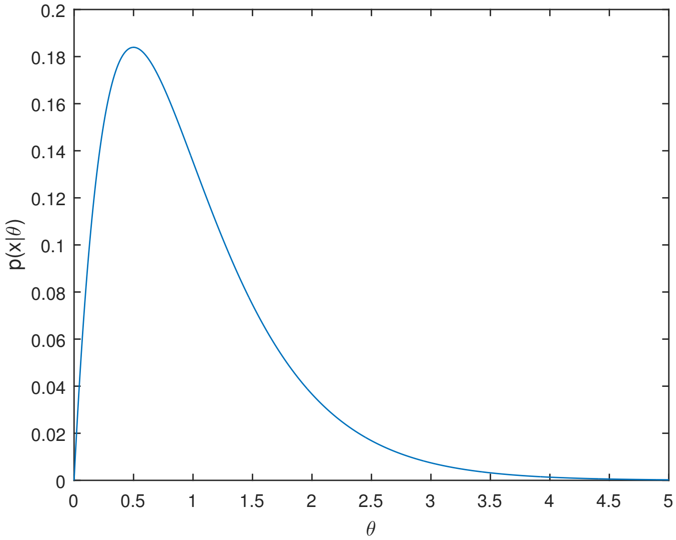

Figure 2: $p(x|\theta)$ versus θ for $x = 2$

2.2

The log-likelihood

$$\begin{aligned} l(\theta) &= \log \prod_{k=1}^{n} p(x_k|\theta) \\ &= \sum_{k=1}^{n} \log \theta \mathrm{e}^{-\theta x_k} \end{aligned} \tag{1}$$

The gradient of $l(\theta)$

$$\begin{aligned} \nabla l(\theta) &= \sum_{k=1}^{n} \frac{(1-\theta x_k)\mathrm{e}^{-\theta x_k}}{\theta \mathrm{e}^{-\theta x_k}} \\ &= \sum_{k=1}^{n} \frac{1-\theta x_k}{\theta} \\ &= \frac{n - \theta \sum_{k=1}^{n} x_k}{\theta} \end{aligned} \tag{2}$$

Let $\nabla l(\theta) = 0$, we can calculate

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

2.3

According to the law of large numbers, when n is very large, the sample average converges to the expected value.

$$\frac{\sum_{k=1}^{n} x_k}{n} \to \mathbb{E}(x) = \int_0^\infty \theta x \mathrm{e}^{-\theta x} \,\mathrm{d}x = \frac{1}{\theta} \tag{3}$$

So $\hat{\theta}$ approach to the true θ when n is very large. If the samples are generated from $p(x|\theta)$ with $\theta = 1$, the maximum-likelihood estimate $\hat{\theta}$ for large n is 1.