Homework 4 - Theory

Lecture: Prof. Qiang Liu

Note: Please typeset your answer (LaTeX recommended) and upload it on edX.

1. Assume X is a discrete random variable that takes values in $\{1, 2, 3\}$, with probability defined by

$$\Pr(X = 1) = \theta_1$$
$$\Pr(X = 2) = 2\theta_1$$
$$\Pr(X = 3) = \theta_2,$$

   where $\theta = [\theta_1, \theta_2]$ is an unknown parameter to be estimated.

   Now assume we observe a sequence $D := \{x^{(1)}, x^{(2)}, \ldots, x^{(n)}\}$ that is *independent and identically distributed* (*i.i.d.*) from the distribution. We assume the number of observations of the values: $1, 2, 3$ in D are s_1, s_2, s_3, respectively.

   (a) [**5 points**] To ensure that $\Pr(X = i)$ is a valid probability mass function, what constraint should we put on $\theta = [\theta_1, \theta_2]$? Write your answers quantitatively as expressions that include θ_1 and θ_2.

   (b) [**5 points**] Write down the joint probability of the data sequence

   $$\Pr(D \mid \theta) = \Pr\left(\left\{x^{(1)}, \ldots, x^{(n)}\right\} \mid \theta\right),$$

   and the log probability $\log \Pr(D \mid \theta)$.

   (c) [**5 points**] Calculate the maximum likelihood estimation $\hat{\theta}$ of θ based on the sequence D.

2. [**10 points**] Let $\{x^{(1)}, \ldots, x^{(n)}\}$ be an *i.i.d.* sample from an exponential distribution, whose the density function is defined as

$$f(x \mid \beta) = \frac{1}{\beta} \exp\left(-\frac{x}{\beta}\right), \quad \text{for} \quad 0 \le x < \infty.$$

   Please find the maximum likelihood estimator (MLE) of the parameter β. Show your work.

3. (a) [**10 points**] Assume that you want to investigate the proportion (θ) of defective items manufactured at a production line. You take a random sample of 30 items and found 5 of them were defective. Assume the prior of θ is a uniform distribution on $[0, 1]$. Please compute the posterior of θ. It is sufficient to write down the posterior density function upto a normalization constant that does not depend on θ.

   (b) [**10 points**] Assume an observation $D := \{x^{(1)}, \ldots, x^{(n)}\}$ is *i.i.d.* drawn from a Gaussian distribution $\mathcal{N}(\mu, 1)$, with an unknown mean μ and a variance of 1. Assume the prior distribution of μ is $\mathcal{N}(0, 1)$. Please derive the posterior distribution $p(\mu \mid D)$ of μ given data D.