StatsI — Exercise 6

1. Let $Y_1, \cdots, Y_n$ be a sample from a Poisson distribution with mean $\theta > 0$ unknown.

   (a) Let $Y = Y_1 + \cdots + Y_n$. Find the mean, the variance and the distribution of Y.
   (**Hint**: Find out the MGF of Y.)

   (b) Obtain the MLE for θ and its standard error.

   (c) Suppose now that only the first m ($m < n$) observations of the sample are known explicitly, while for the other $n - m$ only their sum, Z say, is known, determine the MLE of θ.

2. Find the maximum likelihood estimator of λ given a random sample from the gamma distribution with density function

$$f(x) = \frac{1}{\Gamma(r)} \exp(-\lambda x) x^{r-1} \lambda^r,$$

   where r is a known constant.

3. Find the maximum likelihood estimator for θ from a random sample from the population with density function

$$f(y;\theta) = \frac{2y}{\theta^2} \quad 0 < y \leq \theta, \theta > 0.$$

   Do not use calculus. Draw a picture of the likelihood.

4. Suppose that we have a random sample of size n from the logistic distribution with density function

$$f(y;\mu) = \frac{\exp(y-\mu)}{[1+\exp(y-\mu)]^2}.$$

   Find the likelihood equation for μ, and write the iteration equations for the Fisher scoring method.

   (**Hint**. write quantities as much as possible in terms of the density function and the distribution function for the logistic.)

5. Let $X_1, \cdots, X_n$ be a sample from $U(0,\theta)$, where $\theta > 0$ is an unknown parameter.

   (a) Find the MLE $\widehat{\theta}$ for θ. Derive the distribution for $\widehat{\theta}$, and, therefore, show that $\widehat{\theta}$ is a consistent estimator in the sense that $\widehat{\theta} \xrightarrow{P} \theta$ when $n \to \infty$.
   (**Hint**: $P\{\max_{1\leq i\leq n} X_i \leq y\} = \prod_{1\leq i\leq n} P(X_i \leq y)$.)

   (b) Draw a sample of size $n = 300$ from $U(0,\theta)$ with $\theta = 2$ from R. Set $B = 500$, plot the histograms for bootstrap estimates for θ using both the nonparametric bootstrap and the parametric bootstrap methods. Compare them with the true PDF of $\widehat{\theta}$.

6. Let $X_1, \cdots, X_n$ be a random sample from an exponential distribution with the density function $f(x, \lambda) = \lambda e^{-\lambda x} I(x \geq 0)$, where $\lambda > 0$ is an unknown parameter. Suppose the last $n-m$ $(m < n)$ observations are *censored* in the sense that we only known $X_j \in [u_j, v_j]$ for some given $v_j > u_j > 0$ while X_j itself is unknown, $j = m+1, \cdots, n$. Outline an EM algorithm for estimating the parameter $\lambda > 0$.

7. Let $X_1, \cdots, X_n$ be a random sample from the density function

$$f(x) = \begin{cases} \theta^2 x e^{-\theta x} & x > 0, \\ 0 & x \leq 0, \end{cases}$$

where $\theta > 0$ is an unknown parameter.

   (a) Find the maximum likelihood estimator for θ.

   (b) Suppose now that only the last $n - m$ $(m < n)$ observations of the sample are known explicitly, while for the first m only their sum, $Z = X_1 + \cdots + X_m$, is known. Show that the probability density function of Z is of the form

$$f_Z(z) = \theta^{2m} e^{-\theta z} g(z),$$

   where $g(z)$ is a function independent of θ. Determine the maximum likelihood estimator of θ based on the observations $Z, X_{m+1}, \cdots, X_n$.

   (**Hint**. Find the CDF $F_Z(z)$ first.)

   (c) Suppose now that the first m observations are censored in the sense that we only known $X_j \in [u_j, v_j]$ for some given $v_j > u_j > 0$ while X_j itself is unknown, $j = 1, \cdots, m$. Outline an EM algorithm for estimating the parameter θ.

8. Let $X_1, \cdots, X_n$ be a random sample from Bernoulli distribution, i.e. $P(X_1 = 1) = p = 1 - P(X_1 = 0)$, where $p \in (0, 1)$ is unknown. Let $\theta = p^2$.

   (a) Find the Cramer-Rao lower bound for the variance of unbiased estimators for θ.

   (b) Find the MLE $\widehat{\theta}$ for the parameter θ.

   (c) Show that $E(\widehat{\theta}) \neq \theta$.

   (d) Outline a bootstrap procedure for estimating the bias of $\widehat{\theta}$.

9. Let $\mathbf{X} = (X_1, \cdots, X_n)'$ be a sample from distribution $N(\mu, \sigma^2)$. Let $\theta = (\mu, \sigma^2)'$. Find the Fisher information matrix $\mathcal{I}_{\mathbf{X}}(\theta)$.

   (**Hint**: Use $\theta_2 = \sigma^2$ in your calculation.)