Bios 661: 1 – 5; Bios 673: 2 – 6.

1. C&B 7.8

2. (C&B 7.7) Let $X_1, \ldots, X_n$ be a random sample from one of the two probability density functions, namely,

$$f(x|\theta = 0) = \begin{cases} 1, & \text{if } 0 < x < 1 \\ 0, & \text{otherwise,} \end{cases}$$

and

$$f(x|\theta = 1) = \begin{cases} 1/(2\sqrt{x}), & \text{if } 0 < x < 1 \\ 0, & \text{otherwise.} \end{cases}$$

   (a) Find the likelihood function of θ.

   (b) If $n = 10$ and $\sum_{i=1}^{n} \log(x_i) = -10.7$, find the maximum likelihood estimation (MLE) of θ.

3. C&B 7.10

4. (C&B 7.14) Let $(Y_1, Z_1), \ldots, (Y_n, Z_n)$ be iid random 2-vectors. Assume Y_i and Z_i are independent and follow an exponential distribution with mean λ and μ, respectively, for each i.

   (a) Find MLE for (λ, μ).

   (b) Suppose that we only observe $(X_1, \delta_1), \ldots, (X_n, \delta_n)$, where $X_i = \min(Y_i, Z_i)$ and

$$\delta_i = \begin{cases} 1, & \text{if } X_i = Y_i \\ 0, & \text{if } X_i = Z_i, \end{cases}$$

   for $i = 1, \ldots, n$. In fact, the joint pdf of (X, δ) can be written as

$$\left(\frac{1}{\lambda}\right)^{\delta} \left(\frac{1}{\mu}\right)^{(1-\delta)} e^{-\left(\frac{1}{\lambda}+\frac{1}{\mu}\right)x}.$$

   Find the MLE of (λ, μ).

   [Note: If you are interested in knowing how to derive the joint pdf of (X, δ), you can check the solution key of question 7 in homework 1.]

5. Suppose that n observations, $X_1, \ldots, X_n$, are taken from $N(\mu, 1)$ with an unknown μ. If one can only records the value when the observation is positive, find the MLE for μ.

(a) One possible approach is that one can still observe a complete data on $Y_1, \ldots, Y_n$, where $Y_i = I(X_i > 0)$, even though one did not observe a complete data. What is the distribution of Y?

(b) Given the distribution you identified in (a), find the MLE of the parameter in the distribution.

(c) We have learned that the MLE has a nice invariance property. Comment on how one can use the property to find the MLE of μ.

6. For a certain African village, available data strongly suggests that the expected number of new cases of AIDS developing in any particular year is directly proportional to the expected number of new AIDS cases that developed during the immediately preceding year. An important statistical goal is to estimate the value of this unknown proportionality constant θ ($\theta > 1$). To accomplish this goal, the following statistical model is to be used: For $j = 0, 1, \ldots, n$ consecutive years of data, let Y_j be the random variable denoting the number of new AIDS cases developing in year j. Further, suppose that the $(n + 1)$ random variables $Y_0, Y_1, \ldots, Y_n$ are such that the conditional distribution of Y_{j+1}, given $Y_k = y_k$ for $k = 0, 1, \ldots, j$, depends only on y_j and is Poisson with $E(Y_{j+1}|Y_j = y_j) = \theta y_j$, $j = 0, 1, \ldots, (n - 1)$. Further, assume that the distribution of the random variable Y_0 is Poisson with $E(Y_0) = \theta$, where $\theta > 1$.

   (a) Using all $(n+1)$ random variables $Y_0, Y_1, \ldots, Y_n$, develop an explicit expression for the MLE $\hat{\theta}$ of the unknown proportionality constant θ.

   (b) Find $E(Y_1)$, $E(Y_2)$, and a general expression of $E(Y_j)$, $j = 0, 1, \ldots, n$.

   (c) Suppose that the MLE $\hat{\theta}$ has a large sample property

   $$\sqrt{n}(\hat{\theta} - \theta) \to_d N(0, V(\theta)),$$

   where

   $$V(\theta)^{-1} = -E\{\partial^2 \ell(\theta)/\partial\theta^2\}.$$

   Find a 95% CI for θ when n is large.

7. (Bios 673/740 in class, C&B 7.37) Let $X_1, \ldots, X_{n+1}$ be iid Bernoulli(p), and define the function $h(p)$ by

$$h(p) = P\left(\sum_{i=1}^{n} X_i > X_{n+1}|p\right),$$

which is the probability that the first n observations exceed the $(n + 1)$st.

(a) Show that

$$T(X_1, \ldots, X_{n+1}) = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} X_i > X_{n+1} \\ 0 & \text{otherwise} \end{cases}$$

is an unbiased estimator of $h(p)$.

(b) Find the best unbiased estimator of $h(p)$.