Bios 661: 1 – 5; Bios 673: 2 – 6.

1. C&B 7.40
2. C&B 7.44
3. C&B 8.5(a)(b) [This is a two-parameter case in LRT, using the same principal.]
4. An epidemiologist gathers data (x_i, Y_i) on each of n randomly chosen noncontiguous cities in the United States, where x_i $(i = 1, \ldots, n)$ is the known population size (in millions of people) in city i, and where Y_i is the random variable denoting the number of people in city i with liver cancer. It is reasonable to assume that Y_i $(i = 1, \ldots, n)$ has a Poisson distribution with mean $E(Y_i) = \theta x_i$, where $\theta > 0$ is an unknown parameter, and that $Y_1, \ldots, Y_n$ constitute a set of mutually independent random variables.
   (a) Find the explicit expression for the MLE $\hat{\theta}$ of θ. Also, find the explicit expressions for $\mathrm{E}(\hat{\theta})$ and $\mathrm{Var}(\hat{\theta})$.
   (b) Show that $\hat{\theta}$ is the UMVUE of θ.
   (c) Find the explicit expression for the CRLB for the variance of any unbiased estimator of θ. Justify whether $\mathrm{Var}(\hat{\theta})$ achieves the lower bound.
5. Let $X_1, \cdots, X_n$ be a random sample from an exponential distribution with pdf

$$f(x|\theta) = \begin{cases} e^{-(x-\theta)} & x \geq \theta \\ 0 & x < \theta, \end{cases}$$

   where $-\infty < \theta < \infty$. Consider testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$, where θ_0 is a value specified by the researcher.
   (a) Using the definition of likelihood ratio test, find the test statistic $\lambda(\boldsymbol{x})$.
   (b) The rejection region of the likelihood ratio test is $R = \{\boldsymbol{x} : \lambda(\boldsymbol{x}) \leq c\}$ with some constant cutoff c. Show that this region is equivalent to $R^* = \{\boldsymbol{x} : x_{(1)} \geq c^* \text{ or } x_{(1)} < \theta_0\}$ with another cutoff constant c^*.
   (c) Find c^* specifically, using the definition of test size:

$$\alpha = \sup_{\theta \in \Theta_0} P(\boldsymbol{X} \in R^* | H_0).$$

   (d) Based on the rejection region R^*, draw the power function over the parameter space $-\infty < \theta < \infty$.

6. Let $X_1, \ldots, X_n$ be a random sample from $N(\mu_x, \sigma^2)$ and let $Y_1, \ldots, Y_m$ be a random sample from $N(\mu_y, \sigma^2)$. Assume that two samples are mutually independent and σ^2 is *unknown*. To test the hypothesis $H_0 : \mu_x = \mu_y$ versus $H_1 : \mu_x \neq \mu_y$: [This is a two-sample case in LRT, resulting in classic two-sample t-test.]

   (a) Derive the likelihood ratio test $\lambda(x, y)$.

   (b) Show that the rejection region $\lambda(x, y) \leq c$ is equivalent to $|t| \geq c^*$, where

   $$t = \frac{\bar{x} - \bar{y}}{\sqrt{(\frac{1}{m} + \frac{1}{n})s_p^2}} \quad \text{and} \quad s_p^2 = \frac{1}{m+n-2}\left\{\sum_{i=1}^{n}(x_i - \bar{x})^2 + \sum_{i=1}^{n}(y_i - \bar{y})^2\right\}.$$

   (c) Find the explicit c^* when $\alpha = 0.05$.

   (d) Given that $n = 14$, $m = 9$, $\bar{x} = 1249.9$, $\bar{y} = 1261.3$, $s_x^2 = n^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 549.1$, and $s_y^2 = 156.6$, should one reject the null hypothesis at $\alpha = 0.05$?

7. [Bios 673/740 class discussion, 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 $h(p)$.

8. [Bios 673 class discussion] Suppose $X_1, \ldots, X_n$ is a random sample from $N(\mu, \sigma^2)$.

   (a) If (μ, σ^2) is unknown, find the UMVUE of the 95th percentile.

   (b) If σ^2 is given but μ is unknown, find the UMVUE of $P(X_1 < 1)$.