Bios 661: 1 – 5; Bios 673: 1 – 5.

1. C&B 9.3
2. C&B 9.4
3. C&B 9.17
4. Let $X_1, \ldots, X_n$ be a random sample from a distribution with probability density function
$$f(x|\theta) = \left(\frac{a}{\theta}\right)\left(\frac{x}{\theta}\right)^{a-1}, \quad 0 < x < \theta,$$
where $a \geq 1$ is known and $\theta > 0$ is unknown.
   (a) Construct a confidence interval for θ with coverage probability $1 - \alpha$ by using the cumulative distribution function of the largest order statistic $X_{(n)}$.
   (b) Show that $(X_{(n)}/\theta)^{na}$ is a pivotal quantity and derive the $1 - \alpha$ confidence interval using the quantity.
   (c) Compare the intervals in (a) and (b) and comment on which one would you prefer if they are different.
5. [**2014 final exam**] The exponential distribution is often used to model survival times. This problem develops a simple model for comparing survival times in two groups of patients. Let $X_1, \ldots, X_m$ be a random sample from an exponential distribution with pdf
$$f(x|\mu_1) = \frac{1}{\mu_1} e^{-x/\mu_1}, \quad x > 0, \quad \mu_1 > 0,$$
and let $Y_1, \ldots, Y_n$ be a random sample from an exponential distribution with pdf
$$f(y|\mu_2) = \frac{1}{\mu_2} e^{-y/\mu_2}, \quad y > 0, \quad \mu_2 > 0.$$
Assume that X and Y are independent. Define $\psi = \mu_2/\mu_1$, and let $\bar{X} = m^{-1}\sum_{i=1}^m X_i$ and $\bar{Y} = n^{-1}\sum_{i=1}^n Y_i$ be the sample means.
   (a) Show that the **exact** likelihood ratio test statistic for the hypothesis $H_0 : \mu_1 - \mu_2 = 0$ against $H_1 : \mu_1 - \mu_2 \neq 0$ is
$$\lambda(\boldsymbol{x}, \boldsymbol{y}) = \frac{(m+n)^{m+n}}{m^m n^n} w^m (1-w)^n,$$
where $w = m/(m + nr)$ and $r = \bar{y}/\bar{x}$.

(b) Demonstrate that the rejection region $\{(\boldsymbol{x}, \boldsymbol{y}); \lambda(\boldsymbol{x}, \boldsymbol{y}) < c\}$ is equivalent to $\{r; r < c_1^*\} \cup \{r; r > c_2^*\}$. That means one may reject the null hypothesis by observing either $r < c_1^*$ or $r > c_2^*$. Given a type-I error rate α, find c_1^* and c_2^* using the fact that $\mu_1 \bar{Y}/\mu_2 \bar{X}$ follows $F_{2n,2m}$, which is F distribution with degree of freedoms $2n$ and $2m$.

(c) Explain why $\psi \bar{X}/\bar{Y}$ is a pivotal quantity. Use that pivot to derive an exact 95% confidence interval for ψ.

(d) [**We will discuss this in the review session, no need to return for homework**] Express the critical region of the Wald test for the hypothesis $H_0 : \mu_1 - \mu_2 = 0$ against $H_1 : \mu_1 - \mu_2 \neq 0$ given that the type-I error probability is α.

6. [**Bios 673 class**] Let (X_i, Y_i), $i = 1, \ldots, n$, be *paired* random variables with respective distributions

$$f_{X_i}(x) = (\theta\phi_i)^{-1} e^{-x_i/(\theta\phi_i)}, \quad x_i > 0,$$

and

$$f_{Y_i}(y) = \phi_i^{-1} e^{-y_i/\phi_i}, \quad y_i > 0,$$

where $\phi_i > 0$ is a parameter pertaining to characteristics of the ith pair, and $\theta > 0$ is the parameter reflecting any difference in average values between X and Y. Hence, it is of interest to test if $\theta = 1$ and to indicate if the distribution of X and Y are identical.

(a) Provide an explicit expression for the likelihood function of the random variables $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$, assuming that X_i and Y_i are mutually independent for $i = 1, \ldots, n$. Comment on how many parameters one needs to estimate by the method of maximum likelihood estimation.

(b) One points out that the only parameter of real interest is θ. She suggests that an alternative analysis, based just on the n ratios $R_i = X_i/Y_i$, $i = 1, \ldots, n$, should be favored in order to avoid the estimation of $\phi_1, \ldots, \phi_n$. Prove that this approach is feasible by showing

$$f_{R_i}(r_i) = \frac{\theta}{(\theta + r_i)^2}, \quad 0 < r_i < \infty, \quad i = 1, \ldots, n,$$

which is independent of $\phi_1, \ldots, \phi_n$.

(c) Using mutually independent $R_i = X_i/Y_i$, $i = 1, \ldots, n$, test $H_0 : \theta = 1$ versus $H_1 : \theta > 1$ by a likelihood-based large sample Wald-type test at the size $\alpha = 0.05$.