STA 261s2005 Assignment 10

Do this assignment in preparation for the quiz on Wednesday, March 30th. The questions are practice for the quiz, and are not to be handed in.

1. Show that when a critical region is based on the Neyman-Pearson Lemma, it will depend on $\mathbf{x}$ only through the value of a sufficient statistic.

2. Let C be a most powerful critical region of size α for testing the simple null hypothesis $H_0 : \theta = \theta_0$ against the simple alternative $H_1 : \theta = \theta_1$. Let $\theta_0 \in \Theta_0$, and $P_\theta(\mathbf{X} \in C) \leq P_{\theta_0}(\mathbf{X} \in C)$ for all $\theta \in \Theta_0$. Show that C is also most powerful for testing the *composite* null hypothesis $H_0 : \theta \in \Theta_0$ against the simple alternative.

3. In this unfamiliar but reasonable regression model, the effect of x is linear as usual, but each member of the population has his or her own individual slope. That makes the slope a random variable, because if you took another sample, you'd get another collection of slopes. Accordingly,

   Let $Y_i = x_i B_i$, for $i = 1, \ldots, n$, where

   - $x_1, \ldots, x_n$ are fixed, known constants
   - $B_1, \ldots, B_n$ are independent and identically distributed Normal(β, σ^2) random variables.
   - The parameters β and σ^2 are unknown.
   - The data consist of n pairs (x_i, Y_i). The slopes B_i are not given directly.
   - Yes, there is no error term. Don't worry.

   We wish to test $H_0 : \beta = \beta_0$ against $H_1 : \beta \neq \beta_0$.

   (a) What is Θ?

   (b) What is $\mathcal{X}$?

   (c) What is Θ_0? Is it simple or composite?

   (d) What is Θ_1? Is it simple or composite?

   (e) Find the distribution of $\frac{1}{n}\sum_{i=1}^n \frac{Y_i}{x_i}$.

   (f) Find the constant k so that the following test will be size α:

   $$C = \left\{ \mathbf{x} \in \mathcal{X} : \left| \frac{\sqrt{n(n-1)}[(\frac{1}{n}\sum_{i=1}^n \frac{y_i}{x_i}) - \beta_0]}{\sqrt{\sum_{i=1}^n [y_i/x_i - (\frac{1}{n}\sum_{j=1}^n \frac{y_j}{x_j})]^2}} \right| > k \right\}$$

   Hint: Let $W_i = Y_i/x_i$. What is the distribution of W_i? Write the critical region in terms of w_i values.

4. Now modify the last example so that that the slopes have an exponential distribution with parameter $\beta > 0$ (just to clarify, $E[B_i] = \beta$). This is supposed to model a situation where the random slopes are known to be positive.

(a) What is the length of $\mathbf{X}$?

(b) Find the distribution of $V = \frac{2}{\theta}\sum_{i=1}^{n} \frac{Y_i}{x_i}$.

(c) Use the Neyman-Pearson Lemma to find a size α test of $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1 < theta_0$.

(d) Why is your test uniformly most powerful for testing H_0 against $H_1 : \theta < \theta_0$?

(e) Find the power function of your test. Is it increasing or decreasing?

(f) Consider $H_0 : \theta \geq \theta_0$ against $H_1 : \theta = \theta_1 < \theta_0$. Draw a rough sketch of Θ, Θ_0, Θ_1 and $\pi(\theta)$. Why does your picture show that the test is size α for the composite null hypothesis?

(g) Denoting the critical region of your test by C, answer True or False: $P_\theta(\mathbf{X} \in D) \leq P_\theta(\mathbf{X} \in C)$ for all $\theta \in \Theta_1$, where D is any size α critical region for testing the *composite* null hypothesis.

5. Let $Y_i = x_i + \epsilon_i$, for $i = 1, \ldots, n$, where

- $x_1, \ldots, x_n$ are fixed, known constants
- $\epsilon_1, \ldots, \epsilon_n$ are independent and identically distributed Normal$(0,\sigma^2)$ random variables; the parameter σ^2 is unknown.
- The data consist of n pairs (x_i, Y_i). The error terms ϵ_i are not given directly.

We wish to test $H_0 : \sigma^2 = \sigma_0^2$ against $H_1 : \sigma^2 = \sigma_1^2 < \sigma_0^2$. Use the Neyman-Pearson Lemma to find the most powerful size α test.

6. Let $X_1, \ldots, X_{n_1}$ be a random sample from a distribution with density

$$f(x;\tau) = \sqrt{\frac{\tau}{2\pi}} e^{-\frac{\tau}{2}x^2}, \text{ where } \tau > 0.$$

(a) Let $Y = \tau \sum_{i=1}^{n} X_i^2$. What is the distribution of Y?

(b) Consider the null hypothesis $H_0 : \tau = \tau_0$ against $H_1 : \tau = \tau_1 > \tau_0$. Find the most powerful size α critical region. Call it C.

(c) Now consider $H_0 : \tau = \tau_0$ against $H_1 : \tau > \tau_0$. Why do you know that C is *uniformly* most powerful for this situation?

(d) Find the power function $\pi(\tau) = P_\tau(\mathbf{X} \in C)$.

(e) Is this function increasing, or is it decreasing? Prove it.

(f) Finally, consider $H_0 : \tau \leq \tau_0$ against $H_1 : \tau > \tau_0$. Draw a rough sketch of Θ, Θ_0, Θ_1 and $\pi(\theta)$. Why does your picture show that the test is size α for the composite null hypothesis?

(g) Let D be another size α test of this null versus this alternative. Show $P_\theta(\mathbf{X} \in D) \leq P_\theta(\mathbf{X} \in C)$ for all $\theta \in \Theta_1$.

7. Look again at Exercise 12.9, only this time the sample size is n.

(a) Show $\prod_{i=1}^{n} X_i$ is sufficient for θ.

(b) Show $\sum_{i=1}^{n} -\ln X_i$ is also sufficient for θ.

(c) Find the distribution of $-\ln X_i$. Show your work.

(d) What is the distribution of $-\sum_{i=1}^{n} \ln X_i$?

(e) What is the distribution of $-2\theta \sum_{i=1}^{n} \ln X_i$?

At this point we are on such familiar ground that we should stop.

8. Do Exercises 12.28, 12.30, 12.35, 12.38. For 12.38, Yes or No answers are enough.