Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (NOT YOUR NAME) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a “well known” theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

1. Let $X$ be one observation from Cauchy($\theta$) distribution.
   (a) Does this family have an MLR?
   (b) Show that the test
   \[
   \phi(x) = \begin{cases} 
   1 & \text{if } 1 < x < 3 \\
   0 & \text{otherwise.}
   \end{cases}
   \]
   is most powerful of its size for testing $H_0 : \theta = 0$ versus $H_1 : \theta = 1$. Calculate Type I and Type II Error probabilities.
   (c) Prove or disprove: The test in part (b) is UMP for testing $H_0 : \theta \leq 0$ versus $H_1 : \theta > 0$.

2. Let $X_1, X_2$ be iid Unif($\theta, \theta + 1$). For testing $H_0 : \theta = 0$ versus $H_1 : \theta > 0$ we have two competing tests:

   $\phi_1(X_1) :$ Reject $H_0$ if $X_1 > 0.95$
   $\phi_2(X_1, X_2) :$ Reject $H_0$ if $X_1 + X_2 > C$
(a) Find the constant $C$ so that $\phi_2$ has the same size as $\phi_1$.
(b) Calculate the power function of each test.
(c) Prove or disprove: $\phi_2$ is a more powerful test than $\phi_1$.

3. For a random variable $X$, $F(x) = P(X \leq x)$ and $F(x^−) = P(X < x)$.
The empirical CDF of iid $X_1, ..., X_n$ is given by

$$F_n(x) = \frac{\text{Number of } X_i \leq x}{n}.$$ 

Given iid $X_1, ..., X_n$ from CDF $F_0$, the non-parametric likelihood of the CDF $F$ is

$$L(F) = \prod_{i=1}^{n} [F(X_i) - F(X_i^−)].$$

Let $X_1, ..., X_n$ be iid from CDF $F_0$. Let $F_n$ be the empirical CDF and let $F$ be any CDF. Prove that if $F \neq F_n$ then $L(F) < L(F_n)$.

Hint: $\log(x) \leq x - 1$ for $x > 0$.

4. Let $X_1, ..., X_n$ be a random sample from $N(0, \sigma_x^2)$, and let $Y_1, ..., Y_m$ be a random sample from $N(0, \sigma_y^2)$, independent of the $X$’s. Define $\lambda = \sigma_y^2/\sigma_x^2$.

(a) Find a level $\alpha$ LRT of $H_0 : \lambda = \lambda_0$ versus $H_1 : \lambda \neq \lambda_0$.
(b) Express the rejection region of the LRT in (a) in terms of an $F$ random variable.
(c) Find a $1 - \alpha$ confidence interval for $\lambda$.

5. Suppose $X_1, ..., X_n$ are iid taking values in $\{x_0, x_1, ..., x_k\}$. Let $p(\theta) = (p(x_0, \theta), ..., p(x_k, \theta))^T$, where for $\theta \in \Theta$, $\Theta$ open $\subset R$,

$$p(x_j, \theta) = P_{\theta}(X_1 = x_j), \quad j = 0, 1, ..., k.$$ 

The problem is to estimate the one-dimensional $\theta$.

Let $\mathbf{N} = (N_0, N_1, ..., N_k)$ where $N_j = \sum_{i=1}^{n} I(X_i = x_j)$ is sufficient and $I(A) = 1$ if $A$ is true and is 0 otherwise. Define a function $h$ by

$$h(p(\theta)) = \theta, \quad \forall \theta \in \Theta.$$
Consider the plug-in estimator \( h \left( \frac{N}{n} \right) \). Under regularity conditions it is known that as \( n \to \infty \)

\[
\sqrt{n} \left( h \left( \frac{N}{n} \right) - \theta \right) \xrightarrow{L} N(0, \sigma^2(\theta, h)).
\]

Assume differentiability as needed.
(a) Provide an example of \( h \).
(b) Obtain \( p(X_1, \theta) \).
(c) Provide a lower bound for \( \sigma^2(\theta, h) \).

6. Suppose \( Y_n \) is a sequence of random variables such that

\[
\sqrt{n}(Y_n - \theta) \xrightarrow{L} N(0, \sigma^2).
\]

For a differentiable function \( g \) and a specific number \( \theta \), suppose \( g'(\theta) = 0 \), but \( g''(\theta) \neq 0 \).
(a) Obtain the asymptotic distribution of \( n[g(Y_n) - g(\theta)] \).
(b) Suppose \( \bar{X} \) is a random sample mean and \( E(X_1) = \mu \neq 0 \). Obtain the asymptotic distribution of \( \sqrt{n} \left( 1/\bar{X} - 1/\mu \right) \).