BASIC PHD WRITTEN EXAMINATION IN BIOSTATISTICS

THEORY, SECTION 1

(9:00 AM- 1:00 PM
Tuesday, August 10, 2010)

INSTRUCTIONS:

a) This is a **CLOSED-BOOK** examination.

b) The time limit for this Examination is four hours.

c) Answer any TWO (2) (BUT ONLY TWO) of the THREE (3) questions that follow.

d) Put the answers to different questions on separate sets of paper.

e) Put your code letter, **NOT YOUR NAME**, on each page. The same code will be used for Section 1 and Section 2 of the PhD Theory Exam. Please keep the code confidential and do not share this information with any students or faculty.

f) Return the examination with a signed statement of the UNC honor pledge, separately from your answers. The pledge statement is given on the last page of the exam handout.

g) In the questions to follow, you are required to answer only what is asked, and not to tell all you know about the topics involved.

1. Suppose that $(X_1, Y_1), \ldots, (X_n, Y_n)$ are i.i.d., where

$$\begin{pmatrix} X_i \\ Y_i \end{pmatrix} \sim N_2(\mu, \Sigma),$$

$i = 1, \ldots, n$, where $\mu = (\mu_1, \mu_2)'$ and

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}.$$

(a) Suppose that $\sigma_{12} = 0$ and all other parameters are *unknown*. Consider the hypothesis $H_0 : \frac{\sigma_2^2}{\sigma_1^2} = \Delta_0$ versus $H_1 : \frac{\sigma_2^2}{\sigma_1^2} \neq \Delta_0$, where $\Delta_0 > 0$ is a specified constant. Derive the UMPU size α test for this hypothesis and find the simplest possisble form of the test statistic and critical value for the test.

(b) Derive the simplest possible form of the size α likelihood ratio test corresponding to part (a), and compare it to the UMPU test.

(c) Now suppose that σ_{12} is *unknown* and all other parameters are also *unknown*. Let $\rho = \frac{\sigma_{12}}{\sqrt{\sigma_1^2 \sigma_2^2}}$ denote the population correlation coefficient. Suppose we wish to test $H_0 : \rho = 0$ versus $H_1 : \rho \neq 0$.

(i) Show that the size α likelihood ratio test rejects H_0 when $|R| > c$, where R denotes the sample correlation coefficient and c is chosen to make the test size α.

(ii) Derive the exact distribution of R under the null hypothesis and hence find an explicit expression of c for (i) above.

(iii) Derive the (appropriately normalized) asymptotic distribution of R assuming $\rho = 0$.

2. Suppose that $X_1, \ldots, X_n$ are i.i.d. from the uniform distribution $U(\theta, \theta+1)$, where θ is an unknown, finite, real-valued, scalar parameter.

   (a) Derive the maximum likelihood estimator (MLE) of θ.

   (b) Consider estimating θ under absolute error loss, that is, assume the loss function is given by $L(\theta, a) = \mid \theta - a \mid$. Suppose that the prior for θ is given by $\theta \sim N(\mu_0, \sigma_0^2)$, where (μ_0, σ_0^2) are specified hyperparameters. Derive the Bayes estimator for θ.

   (c) Under squared error loss, consider the class of estimators given by $d(X) = aX_{(1)} + bX_{(n)} + c$, where (a, b, c) are constants, $X = (X_1, \ldots, X_n)$, and $X_{(j)}$ is the jth order statistic. Within this class of estimators, derive an admissible estimator of θ.

   (d) Under squared error loss, obtain a minimax estimator for θ.

   (e) Derive the (appropriately normalized) asymptotic distribution of $R_n = X_{(n)} - X_{(1)}$.

3. Let $X_1, \ldots, X_n$ be an i.i.d. sample of real random variables with $EX_1 = 0$ and $0 < \operatorname{var}(X_1) = \sigma^2 < \infty$. Define $\overline{X}_n = n^{-1} \sum_{i=1}^n X_i$ and $S_n^2 = (n-1)^{-1} \sum_{i=1}^n \left(X_i - \overline{X}_n\right)^2$. Do the following:

(a) Show that for x close to zero, $e^x - 1 - x = x^2/2 + o(x^2)$.

(b) Show that $e^{\overline{X}_n} - 1 - \overline{X}_n \to 0$,

$$\frac{e^{\overline{X}_n} - \overline{X}_n - 1}{\overline{X}_n} \to 0 \text{ and } \frac{e^{\overline{X}_n} - \overline{X}_n - 1}{\overline{X}_n^2} \to \frac{1}{2}$$

in probability.

(c) Show that

$$\frac{2n}{S_n^2}\left(e^{\overline{X}_n} - 1 - \overline{X}_n\right)$$

converges in distribution to a χ^2 random variable with 1 degree of freedom.

(d) Show that

$$\frac{2\sqrt{n}}{S_n}\left(\frac{e^{\overline{X}_n} - 1 - \overline{X}_n}{\overline{X}_n}\right)$$

converges in distribution to a $N(0,1)$ random variable.

(e) Show that

$$\frac{2n}{S_n^2}\left(e^{\overline{X}_n} - 1 - \overline{X}_n\right)\tan \overline{X}_n \to 0$$

in probability, where tan denotes the tangent function.

(f) Show that

$$\frac{2\sqrt{n}\left(e^{\overline{X}_n} - 1 - \overline{X}_n\right)\tan \overline{X}_n}{S_n \overline{X}_n^2}$$

converges in distribution to a $N(0,1)$ random variable.

2010 PhD Theory Exam, Section 1

Statement of the UNC honor pledge:

"In recognition of and in the spirit of the honor code, I certify that I have neither given nor received aid on this examination and that I will report all Honor Code violations observed by me."

(Signed) ____

NAME

(Printed) ____

NAME