BASIC PHD WRITTEN EXAMINATION IN BIOSTATISTICS

THEORY, SECTION 1

(9:00 AM- 1:00 PM
Wednesday, August 10, 2011)

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 exam code, **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. Let $X_1, X_2, \ldots,$ be a sequence of i.i.d. real random variables with $EX_1 = 0$. Let N be a Poisson random variable with parameter $\lambda \geq 0$ and independent of $X_1, X_2, \ldots,$. For each integer $m \geq 0$, let $\overline{X}_m = m^{-1} \sum_{i=1}^m X_i$, where we define $\overline{X}_0 = 0$.

   (a) Assume $\sigma^2 = EX_1^2 < \infty$ and do the following:

      (i) Show that $\operatorname{var}(\overline{X}_N) \leq \sigma^2 \left[P(N < \lambda^{1/3}) + \frac{P(N \geq \lambda^{1/3})}{\lambda} \right]$.

      (ii) Show that $P(N < \lambda^{1/3}) \to 0$, as $\lambda \to \infty$. Hint: Use Chebyshev's inequality.

      (iii) Show that $\lim_{\lambda \to \infty} P(|\overline{X}_N| \geq \epsilon) = 0$ for every $\epsilon > 0$.

   (b) Let $\psi(t)$ be the characteristic function of a standard normal random variable, and define $Z_m = m^{1/2} \overline{X}_m / \sigma$. Continue to assume $\sigma^2 < \infty$. Do the following:

      (i) Show that for any real t,

$$\left| E(e^{itZ_N}) - \psi(t) \right| \leq 2P(N < \lambda^{1/3}) + \max_{m \geq \lambda^{1/3}} |E(e^{itZ_m}) - \psi(t)|.$$

      (ii) Show that for any real t, $|E(e^{itZ_N}) - \psi(t)| \to 0$, as $\lambda \to \infty$.

   (c) Now do not assume $\sigma^2 < \infty$. Do the following:

      (i) Show that for each $\epsilon > 0$,

$$P(|\overline{X}_N| \geq \epsilon) \leq P(N < \lambda^{1/3}) + P\left(\max_{m \geq \lambda^{1/3}} |\overline{X}_m| \geq \epsilon \right).$$

      (ii) Show that $\overline{X}_N \to 0$, in probability, as $\lambda \to \infty$. Hint: Use the strong law of large numbers.

2. (a) Let X be a random variable and let ν be a parameter of interest in the distribution of X. Suppose that $T(X)$ is an unbiased estimator of ν. Show that any unbiased estimator of ν is of the form $T(X) - U(X)$, where $E\{U(X)\} = 0$.

In the sequel, let X be a discrete random variable with $P(X = -1) = p, P(X = k) = (1-p)^2 p^k, k = 0, 1, 2, \ldots$, where $p \in (0, 1)$ is unknown.

(b) Show that $E\{U(X)\} = 0$ if and only if $U(k) = ak$ for all $k = -1, 0, 1, 2, \ldots$ and some a.

(c) Using the results in (a) and (b), show that $I(X = 0)$ is the unique admissible estimator under squared error loss amongst all unbiased estimators of $(1-p)^2$, where $I(\cdot)$ is the indicator function.

(d) Show that no unique admissible estimator exists for p under squared error loss amongst unbiased estimators for p.

(e) Prove whether there exist unbiased estimators of p^{-1}. If so, then determine whether a unique admissible estimator exists under squared error loss amongst unbiased estimators for p^{-1}.

3. Consider a sequence of numbers $x_1, x_2, \ldots$ and place vertical lines before x_1 and between x_j and x_{j+1} whenever $x_j > x_{j+1}$. We say that the runs are the segments between pairs of lines. Thus, each run is an increasing segment of the sequence $x_1, x_2, \ldots$

   Suppose that $X_1, X_2, \ldots$ are independent and indentically distributed uniform(0,1) random variables and that we are interested in the lengths of the successive runs. Let L_j denote the length of the jth run.

   (a) Compute $P(L_1 \geq m), m = 1, 2, \ldots$.

   (b) Suppose we know that the jth run starts with the value x. Compute $P(L_j \geq m|x)$.

   (c) Let I_j denote the initial value of the jth run. Show that $p_n(y|x)$, the probability density that the $n+1$st run has $I_{n+1} = y$ given that the nth run has just begun with $I_n = x$, equals e^{1-x} if $y < x$ and $e^{1-x} - e^{y-x}$ if $y > x$.

   (d) Demonstrate that $\pi(y)$, the probability density function for I_n as $n \to \infty$, satisfies $\pi(y) = 2(1-y), 0 < y < 1$. You may do this by verifying the continuous state equilibrium equations for discrete time Markov chains: $\pi(y) = \int_0^1 \pi(x)p(y|x)$.

   (e) Find $\lim_{n\to\infty} P(L_n \geq m)$.

   (f) What is the average length of a run as $n \to \infty$, that is, $\lim_{n\to\infty} E(L_n)$?

2011 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