BASIC PHD WRITTEN EXAMINATION

THEORY, SECTION 1

(9:00 AM–1:00 PM, July 23, 2019)

INSTRUCTIONS:

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

(b) The time limit for this examination is four hours.

(c) Answer both 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 is 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. Sharing your code with either students or faculty is viewed as a violation of the UNC honor code.

(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. (25 points) Let X denote a random variable from $N(0,1)$, and let Y be an outcome variable. The joint distribution of (X,Y) has a finite second moment and $E[X^2Y^2]<\infty$. Assume that we observe n i.i.d copies of (X,Y), denoted by $(X_1,Y_1),...,(X_n,Y_n)$. The goal is to obtain the best prediction of Y given X for a future subject.

   (a) One simple prediction is to consider a linear function, $\alpha+\beta X$, to minimize the following squared loss:
$$E\left[\{Y-(\alpha+\beta X)\}^2\right],$$
   where the expectation is with respect to the joint distribution of (Y,X). Show that the optimal solution for (α,β), denoted by (α^*,β^*), is given by
$$\begin{pmatrix}\alpha^*\\ \beta^*\end{pmatrix}=\begin{pmatrix}E[Y]\\ E[XY]\end{pmatrix}. \tag{1}$$

   (b) From (1), we estimate (α^*,β^*) as
$$\widehat{\alpha}=n^{-1}\sum_{i=1}^{n}Y_i,\quad \widehat{\beta}=n^{-1}\sum_{i=1}^{n}X_iY_i.$$
   Give the asymptotic distribution of the obtained estimator after a proper normalization.

Now suppose that we know the distribution of Y given X is from a log-normal family, i.e.,
$$\log Y=\gamma X+N(0,\sigma^2).$$

   (c) Obtain the maximum likelihood estimators for α^* and β^* given in (1) and derive their asymptotic distribution.

   (d) Calculate the asymptotic relative efficiency between the maximum likelihood estimator for β^* and $\widehat{\beta}$ given in (b).

   (e) If we allow the prediction function to be arbitrary, that is, we aim to find the best function, $g(X)$, to minimize
$$E\left[\{Y-g(X)\}^2\right],$$
   what is the optimal $g(X)$ in terms of (γ,σ^2)?
   Hint: consider minimization conditional on X.

Points: (a) 5; (b) 5; (c) 5; (d) 5; (e) 5.

2. (25 points) Let $X_1, \ldots, X_n$ be i.i.d samples from a distribution with density function

$$f(x) = \theta^{-1} e^{(a-x)/\theta} I(x > a), \text{ where } \theta > 0.$$

(a) When a is known, derive the uniformly most powerful test of size α for testing $H_0 : \theta \geq \theta_0$ versus $\theta < \theta_0$, where θ_0 is a known constant.

(b) When a is known, derive the asymptotic distribution of the maximum likelihood estimator of θ.

In the rest questions, we assume $a = \theta$, i.e. the density is $f(x) = \theta^{-1} e^{(\theta - x)/\theta} I(x > \theta)$.

(c) Prove that both $\bar{X}/\theta$ and $X_{(1)}/\theta$ are pivotal quantities, where $\bar{X}$ is the sample mean and $X_{(1)}$ is the smallest order statistic.

(d) Obtain two confidence intervals with confidence coefficient $1 - \alpha$ for θ, based on two pivotal quantities in (c).

(e) When n is sufficiently large, which of the two confidence intervals has shorter length? Justify your answer.

Points: (a) 5; (b) 5; (c) 5; (d) 5; (e) 5.

2019 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