PhD Qualifying Exam
Department of Statistics
University of Florida
August 2009

Instructions:

1. You have exactly four hours to answer questions in this examination.
2. There are 8 problems of which you must answer 6. You must do at least one problem from each of the four categories of Linear Models, Generalized Linear Models, Probability, and Inference. If there is doubt as to what area a particular problem covers, then ask.
3. Only your first 6 problems will be graded.
4. Write your chosen identifying number on every page.
5. Do not write your name anywhere on your exam.
6. Write only on one side of each sheet of paper. For each problem you do, start the problem on a new page. At the end of the exam, for each problem, staple together all pages for that problem.
7. Clearly label each part of each question with the question number and the part.
8. You must show your work to receive credit.
9. While the eight questions are equally weighted, within a given question, the parts may have different weights.
10. Do not write near the upper left corner of the page where the pages will be stapled together.

1. Consider the simple linear regression model

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \qquad i = 1, \ldots, n,$$

with the ϵ_i's iid with mean 0 and variance σ^2, and β_0, β_1, σ are unknown, and the x_i's are fixed known constants, to be determined by the experimenter, and n is even. Suppose also that the x_i's are constrained to lie in the interval $[-1, 1]$.

(A) For what choice of $x_1, \ldots, x_n$ is $\text{Var}(\hat{\beta}_1)$ minimized?
(B) For what choice of $x_1, \ldots, x_n$ is $\text{Var}(\hat{\beta}_0)$ minimized?
(C) Consider now the model

$$Y_i = \beta_0 + \beta_1 x_i + \beta_2 z_i + \epsilon_i, \qquad i = 1, \ldots, n$$

with the same stipulations on the ϵ_i's, except that we now assume additionally that their common distribution is normal. What is the F-test for testing

$$H_0\colon\ \beta_2 = 0 \quad \text{vs.} \quad H_1\colon\ \beta_2 \neq 0?$$

2. Consider a standard linear model $Y \sim N_n(\mu, \sigma^2 I_{n\times n})$ where μ is known to lie in a subspace V_f of $\mathbb{R}^n$. Let $V_r \subset V_f$, and consider the F-test for testing

$$H_0\colon\ \mu \in V_r \quad \text{vs.} \quad H_1\colon\ \mu \notin V_r.$$

For $\mu \in V_f$, let $p(\mu)$ denote the power of the F test when μ is the true value of $E(Y)$. Let P_f and P_r be the projections of Y onto the spaces V_f and V_r, respectively, and express μ uniquely as

$$\mu = \mu_1 + \mu_2, \quad \text{where } \mu_1 \in V_r,\ \mu_2 \in V_f \cap V_r^{\perp}.$$

Show that $p(\mu)$ is an increasing function of $\|\mu_2\|^2 = \mu'(P_f - P_r)\mu$. You may want to use the fact that for fixed ν, the family of distributions $\chi^2_\nu(\gamma)$ is stochastically increasing with γ. In this case, you must prove that fact.

3. Confidence interval on the selected population mean.

Suppose that we observe $X_i \sim N(\theta_i, 1)$, $i = 1, 2$ with the goal of putting a $1-\alpha$ confidence interval on the θ_i corresponding to the larger of X_1 and X_2. That is, if $X_1 > X_2$ we assert that $\theta_1 \in X_1 \pm c$ with probability $1-\alpha$, and if $X_2 > X_1$ we assert that $\theta_2 \in X_2 \pm c$ with probability $1-\alpha$. Show that if c is the upper $\alpha/2$ cutoff from the standard normal, this procedure is, in fact, a $1-\alpha$ confidence procedure.

4. Stein estimation with unequal shrinkage.

   Let $X_{p\times 1}$ be multivariate normal, $N(\theta, I)$, and consider estimation of θ under squared error loss, $L(\delta, \theta) = |\delta - \theta|^2$. Consider the Stein estimator given componentwise by

$$\delta_i(X) = \left(1 - \frac{c_i}{|X|^2}\right) X_i,$$

   where $c_1, \ldots, c_p$ are constants.

   (A) Find the unbiased estimate of the risk of $\delta(X)$.

   (B) How do you know that the unbiased estimate of the risk is unique?

   (C) We know that if all of the c_i are equal, then $\delta(X)$ is minimax as long as $c \leq 2(p-2)$. With unequal c_i, under what conditions on $c_1, \ldots, c_p$ will $\delta(X)$ remain minimax? Specifically, how large can $\max\{c_i\} - \min\{c_i\}$ be? What is the upper bound on c_i?

5. Consider the setting of toxicology studies where we want to model the probability of a response (here, death) to different doses of a toxin. Suppose each subject has a tolerance T for a dose x. That is, if $T \leq x$, then the subject dies. Define Y to be the indicator of death ($Y = 1$ corresponds to the subject dying).

   Assume the distribution of tolerances in the population follows an extreme value (or Gumbel) distribution with cdf $F(t) = \exp\big(-\exp\{-(t-a)/b\}\big)$ with mean $a + .577b$ and standard deviation $\pi b/\sqrt{6}$.

   (A) For a given dose x, derive the probability a randomly selected subject dies, i.e., $\pi(x) = P(Y = 1|x)$ as a function of the mean and variance of the tolerance distribution.

   (B) Derive the link function g for the regression of the binary response Y on the covariate dose (x), such that $g(\pi(x))$ has the following form: $g(\pi(x)) = \alpha + \beta x$. How do the parameters α and β relate to a and b?

   (C) Does $\pi(x)$ approach one at the same rate that it approaches zero? Explain.

   (D) Derive the LD50 defined as the dose x_0 such that $\pi(x_0) = .5$ as a function of the regression parameters, (α, β).

   For parts (E)–(G) below, assume that we have independent data Y_i, $i = 1, \ldots, m$ from Binomial$(n_i, \pi(x_i))$ distributions with $\pi(x_i)$ defined in part (A).

   (E) Construct a large sample 95% confidence interval for the LD50 assuming the mle of (α, β) is approximately normal. Provide details.

   (F) In a large sample size, how might you assess graphically if the extreme value tolerance distribution is reasonable? Explain.

   (G) In a small sample size, would you expect to be able to determine whether the assumption that the distribution of tolerances follows a logistic distribution versus a normal distribution? Explain.

6. Suppose that Y_1 and Y_2 are independent Poisson random variables with means μ_1 and μ_2 and that our only interest is inference on the ratio $\psi = \mu_1/\mu_2$.

   (A) Derive the conditional likelihood for conducting inference on ψ. Please justify all your steps.

   (B) Derive the conditional maximum likelihood estimator (i.e., the mle based on the conditional likelihood) $\hat{\psi}_c$ in closed form.

   (C) Derive the Fisher information based on the conditional likelihood.

   (D) Derive $\hat{\psi}$, the unconditional mle of ψ, and its large sample variance.

   (E) Compare the large sample variance of $\hat{\psi}$ to that of $\hat{\psi}_c$. Comment.

   (F) Derive a $100(1-\alpha)\%$ confidence interval for $\hat{\psi}_c$ without resorting to large sample theory.

   (G) What is the typical way to "eliminate" nuisance parameters in the setting of Bayesian inference? Is there a way to do this in closed form here? Explain and provide details.

7. Let $\{X_n, n \geq 1\}$ be a sequence of random variables.

   (A) Prove that if
$$E\left(\sum_{n=1}^{\infty} \frac{|X_n|}{|X_n|+1}\right) < \infty$$
   then
$$\sum_{n=1}^{\infty} |X_n| < \infty \quad \text{a.c.}$$

   (B) Prove that if $\{X_n,\, n \geq 1\}$ are independent and
$$\sum_{n=1}^{\infty} |X_n| < \infty \quad \text{a.c.},$$
   then
$$E\left(\sum_{n=1}^{\infty} \frac{|X_n|}{|X_n|+1}\right) < \infty.$$

   (C) Prove that if
$$E\left(\sum_{n=1}^{\infty} \frac{|X_n|}{|X_n|+b_n}\right) < \infty$$
   where $\{b_n,\, n \geq 1\}$ is a sequence of positive constants with $b_n \uparrow \infty$, then
$$\lim_{n\to\infty} \frac{\sum_{j=1}^{n} X_j}{b_n} = 0 \quad \text{a.c.}$$

8. Let $S_n = \sum_{j=1}^{n} X_j$, $n \geq 1$ where $\{X_n, n \geq 1\}$ is a sequence of independent random variables and let $\{b_n, n \geq 1\}$ be a nondecreasing sequence of constants in $(0, \infty)$ with $\lim_{n \to \infty} b_n = \infty$. Suppose that

$$\frac{S_n}{b_n} \xrightarrow{P} 0.$$

Can it be concluded that

$$\frac{\max_{1 \leq k \leq n} |S_k|}{b_n} \xrightarrow{P} 0?$$

Give a proof or else present (with verification) a suitable counterexample.