Ph. D. Qualifying Exam (Part II of Written Exam)
Tuesday, January 5, 1999

You have four hours. Do as many problems as you can. No one is expected to answer all the problems correctly. Partial credit will be given. All problems are worth an equal amount of credit.

Put your solution to each problem on a separate sheet of paper.

Problem 1. Let $F_n, n \geq 1$ and F be distribution functions.

(a) Suppose $F_n(x) \to F(x)$ for all x in some dense subset of the real line. Show that F_n converges in distribution to F.

(b) Is it possible for F_n to converge in distribution to F but $F_n(x) \not\to F(x)$ for all x in some dense subset of the real line?

Problem 2. Let Ω_n be the the space of all counting measures on $[0,1]$ of total mass n:

$$\Omega_n = \Big\{ \sum_{i=1}^{n} \delta_{x_i} \colon x_i \in [0,1],\ i = 1, \ldots, n \Big\},$$

where δ_{x_i} denotes the Dirac measure with mass 1 at x_i. Define the projection $\pi_n \colon [0,1]^n \to \Omega_n$ by $\pi_n(x_1, \ldots, x_n) = \sum_{i=1}^{n} \delta_{x_i}$.

(a) Show that $\mathcal{F}_n = \{\pi_n(B) \colon B \in \mathcal{B}^n\}$ is a σ-field of subsets of Ω_n for each $n \geq 1$, where $\mathcal{B}^n$ is the σ-field of n-dimensional Borel subsets of $[0,1]^n$.

(b) How would you define Ω_0 and $\mathcal{F}_0$?

(c) Let $\mathcal{F}$ be the smallest σ-field of subsets of $\Omega = \cup_{n=0}^{\infty} \Omega_n$ containing $\mathcal{F}_0, \mathcal{F}_1, \ldots$. Show that all sets in $\mathcal{F}$ can be explicitly presented in terms of sets in $\mathcal{F}_0, \mathcal{F}_1, \ldots$.

(d) Define

$$P(A) = \frac{1}{e} \left(1\{A \cap \Omega_0 \neq \emptyset\} + \sum_{n=1}^{\infty} \frac{\lambda^n(\pi_n^{-1}(A \cap \Omega_n))}{n!} \right)$$

for $A \in \mathcal{F}$, where λ^n is Lebesgue measure on $[0,1]^n$ and $1\{\cdot\}$ denotes the indicator function. Show that P is a probability measure on $(\Omega, \mathcal{F})$.

Problem 3. Let $X_1, X_2, \ldots, X_n$ be i.i.d. Bernoulli(p).

(a) What is the MLE of p^2?

(b) What is the best unbiased estimator of p^2?

(c) What is the best unbiased estimator of p^n?

Hint for (b) and (c): Use the Rao-Blackwell procedure.

(d) Do the estimators in (b) and (c) achieve the Cramer-Rao lower bound? (This can be answered with little or no calculation.)

Problem 4. Suppose that you are helping a client to carry out an experimental study to determine the effects of 6 factors on the yield of a chemical reaction. Called the 6 factors A, B, C, D, E, and F.

(a) Construct a 2^{6-2} fractional factorial design with as high a resolution as possible.

(b) What are the generators of your design? What is the definition relation of your design?

(c) Which effects are confounded with the main effect of factor A in your design? Which effects are confounded with the two-factor interaction BC?

Problem 5. Consider the following regression model:

$$\boldsymbol{Y} = X\boldsymbol{\beta} + Z\boldsymbol{\gamma} + \boldsymbol{\epsilon},$$

where X and Z are known $n \times p$ and $n \times q$ full-rank matrices and $X'Z = 0$, and $\boldsymbol{\epsilon}$ is normally distributed with $\mathrm{E}\boldsymbol{\epsilon} = \mathbf{0}$ and $\mathrm{Var}(\boldsymbol{\epsilon}) = \sigma^2 I$.

(a) Find the maximum likelihood estimates of $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, and σ^2. Denote these estimates by $\hat{\boldsymbol{\beta}}$, $\hat{\gamma}$, and $\hat{\sigma}^2$.

(b) Find the expectations and covariance matrices of $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$.

(c) Show that $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$ are independent. Show that $(\hat{\boldsymbol{\beta}}, \hat{\gamma})$ is independent of $\hat{\sigma}^2$.

Problem 6. Suppose that, under P_θ, the random variables $X_1, X_2, \ldots, X_n$ are i.i.d. with distribution $N(\theta, 1)$.

(a) Write down the asymptotic distribution of $\bar{X}_n = n^{-1}\sum_{i=1}^n X_i$ under P_θ.

(b) Define

$$Y_n = \begin{cases} 1 & \text{if} \quad |\bar{X}_n| \le n^{-\frac{1}{4}}, \\ 0 & \text{if} \quad |\bar{X}_n| > n^{-\frac{1}{4}}. \end{cases}$$

For each θ, show that

$$Y_n \xrightarrow{p} \begin{cases} 1 & \text{if} \ \ \theta = 0 \\ 0 & \text{if} \ \ \theta \neq 0 \end{cases}$$

under P_θ, where $\xrightarrow{p}$ stands for convergence in probability.

Problem 7. There are five workers selling food from little carts at five locations on our campus. The owner keeps getting complaints from her workers that they could sell more if they were at some other location which (they claim) is a better one than theirs.

(a) Set up an experiment so that, in one week (MTuWThF) she can look at the daily sales in such a way that

-each worker works one day in each location;
-she can find out whether some locations are better than others.

Defend your choice of design.

(b) What other questions can the owner answer? Give a model, and use a table format to help describe your answers, but please do not include any formulas in the table.

(c) The owner wants to know whether (taking all workers as a group) workers who work at their customary location get more sales than when they are working at a different location. How would you answer this question? Give a model and a description.

Problem 8. Let X be an $n \times p$ (with $p \le n$) matrix having rank $r < p$, and assume the $n \times 1$ random vector Y follows the linear model $Y \sim N(X\beta, \sigma^2 I)$.

(a) Let L be an $m \times p$ matrix. Consider the hypothesis $H_0 : L\beta = 0$ versus $H_1 : L\beta \neq 0$. State conditions on L which guarantee that this hypothesis is *testable*, that is, state the definition of a *testable* hypothesis.

(b) Let $V = \text{range}(X)$ and let W be any subspace of V. Define $\mu = X\beta$. Show that the hypothesis $H_0 : \mu \in W$ versus $H_1 : \mu \notin W$ (but $\in V$) can be formulated as a testable hypothesis.

Problem 9. Suppose that $Y_1, \cdots, Y_n$ are independent Bernoulli random variables with the corresponding covariates $x_1, \cdots, x_n$ such that

$$\log\left\{\frac{P(Y_i = 1)}{P(Y_i = 0)}\right\} = \gamma + \beta x_i$$

(a) Find a minimal sufficient statistic for (γ, β).
[Recall that a complete sufficient statistic is automatically minimal.]

(b) Let (R, T) be the minimal sufficient statistic for (γ, β) found in part (a). Derive the conditional distribution of T given R.

(c) Treating γ as a nuisance parameter, develop a size α test for testing the null hypothesis $\beta = 0$ (no dependence on the covariates) versus the one-sided alternative $\beta > 0$.