Ph. D. Qualifying Exam (Part II of Written Exam)
Saturday, January 4, 1997

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 X be a random variable such that $0 \le X \le 1$, and let $X_1, X_2, \ldots$ be i.i.d. copies of X. Denote $\mu = EX$ and let $\psi(t) = Ee^{tX}$ be the moment generating function of X.

(a) Show that

$$\psi(t) \le 1 - \mu + \mu e^t \quad \text{for all } t.$$

(b) Show that

$$P\left(\sum_{i=1}^{n} X_i \ge n(\mu + \epsilon)\right) \le \left\{e^{-t(\mu+\epsilon)}\left(1 - \mu + \mu e^t\right)\right\}^n$$

for all $t \ge 0$ and $\epsilon > 0$.

(c) Use (b) to conclude that

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i \le \mu$$

almost surely. HINT: Regard the expression inside the braces as a function of t and consider its behavior close to $t = 0$.

Problem 2. Let $X_1, X_2, \ldots$ be i.i.d. with common distribution $N(1, 1)$. Let $V_n = \frac{1}{n}\sum_1^n X_i$ and $M_n =$ median of $\{X_1, \ldots, X_n\}$ be the mean and median based on the first n random variables.

(a) Write down the limiting distributions of $\sqrt{n}(V_n - 1)$ and $\sqrt{n}(M_n - 1)$.

(b) Let

$$T_n = \begin{cases} V_n & \text{if } |V_n| \le n^{-\frac{1}{4}} \\ M_n & \text{if } |V_n| > n^{-\frac{1}{4}}. \end{cases}$$

Obtain the limiting distribution of $\sqrt{n}(T_n - 1)$. *Hint: First compute the limit of* $P(|V_n| \le n^{-\frac{1}{4}})$ *and consider the two cases of the definition of* T_n *separately.*

Problem 3. Let N be a positive integer-valued random variable and let $X_1, X_2, \ldots$ be random variables with finite expectations.

(a) Define $X_N(\omega) = X_{N(\omega)}(\omega)$. Show that X_N is a random variable.

(b) Suppose that N is bounded and let $\{N_n,\ n = 1, 2, \ldots\}$ be a sequence of positive integer-valued random variables such that $N_n \nearrow N$ almost surely. Show that

$$\lim_{n\to\infty} EX_{N_n} = EX_N.$$

Problem 4. Suppose that $\{X_i, i = 1, \ldots, n\}$ are independent random variables with $\mathrm{E}(X_i) = \mu_i$ and $\mathrm{Var}(X_i) = \sigma_i^2$. In ANOVA and regression, Box-Cox power transformations are often used to stabilize the variances of $\{X_i\}$.

(a) If $\sigma_i = \alpha\mu_i^\beta$ and the transformation $Y_i = X_i^p$ is used to stabilize the variances, find an approximate relationship between β and p using the Taylor expansion.

(b) If $\{X_i, i = 1, \ldots, n\}$ are independent Poisson random variables, which transformation will stabilize the variances of $\{X_i\}$?

(c) In practice, β is usually unknown. If $\hat{\sigma}_i^2$ and $\hat{\mu}_i$ are available, give an empirical way to estimate the power p.

Problem 5. Consider the following linear regression model:

$$y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi} + \epsilon_i, \quad i = 1, \ldots, n,$$

where the ϵ_i's are iid normally distributed random variables with mean zero and variance σ^2.

(a) If $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)'$ has a multivariate normal distribution with mean zero and non-singular covariance matrix $V = \mathrm{Cov}(\boldsymbol{\xi})$, and if A is a $n \times n$ symmetric matrix with rank r, show that $\boldsymbol{\xi}' A \boldsymbol{\xi}$ has a χ^2 distribution with $df = r$ if AV is idempotent. (Hint: If AV is idempotent and $V = C'C$, then $B = CAC'$ is idempotent with rank r).

(b) Let $\hat{\boldsymbol{\beta}}$ be the least squares estimate of $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)'$ and let $\hat{\boldsymbol{\epsilon}}$ be the residuals from the the fitted model. Assuming that the rank of X is $p + 1$, show that the statistic

$$F = \frac{(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'(X'X)(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})/(p+1)}{\hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}}/(n-p-1)}$$

has an F-distribution.

Problem 6. Questionnaires about NAFTA, containing 20 multiple-choice questions, were sent to a random sample of executives in each of the three countries involved in this important free-trade agreement. Each questionnaire was categorized as "high" or "low" based on the number of questions correctly answered ($Y \geq 12$ or $Y < 12$). The results are listed in the following table:

	USA	Canada	Mexico
High	22	33	31
Low	29	19	19

(a) A χ^2 test of association for the table yielded a value of 5.333 with the critical value of $\chi^2_2(0.05) = 5.99$. In words, what hypothesis is being tested? What is your conclusion based on the test?

(b) After examining the table, do you think your conclusion in **(a)** is reasonable? If not, what further analysis can you perform on the table? What are the conclusions from your analysis? ($\chi^2_1(0.05) = 3.84$).

(c) What simple alternative analysis would be appropriate if the number of correct answers (Y) for each questionnaire is used directly? Include the analysis suggested in **(b)**. Which form of analysis would you prefer, and why?

Problem 7. Let $X_1, X_2, \ldots, X_n$ be i.i.d. with density

$$f(x|\theta) = \frac{1}{2}\theta^3 x^2 e^{-\theta x}, \quad x > 0.$$

(a) What is the Cramer-Rao lower bound for the variance of an unbiased estimator of θ?

(b) Find a simple unbiased estimator of θ. (Hint: evaluate EX^k for arbitrary real values k.)

(c) Does the estimator you found in the previous part achieve the Cramer-Rao lower bound? Is it asymptotically efficient?

(d) Does there exist an unbiased estimator of θ which exactly achieves the Cramer-Rao lower bound for finite n? (Justify your answer.)

Problem 8. Let $X_1, X_2, \dots, X_n$ be independent with probability density functions given by

$$f(x_i|\theta) = \begin{cases} \frac{1}{2i\theta} & \text{for } -i\theta < x_i < i\theta \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find a one-dimensional sufficient statistic for θ. (Your statistic should **also** be minimal, but you are not required to show this.)

(b) Find the minimum variance unbiased estimator of θ. (Prove that your estimator has this property.)

Problem 9. Suppose that $X_1, \dots, X_n$ are i.i.d. with the Poisson distribution,

$$P[X_i = x|\lambda] = \lambda^x e^{-\lambda}/x!$$

for $x = 0, 1, \dots$ and for $\lambda > 0$.

(a) Show that a conjugate family of prior distributions is given by the Gamma distribution, with probability density:

$$f(\lambda, \alpha, \beta) = \Gamma(\alpha)^{-1}\beta^{-\alpha}\lambda^{\alpha-1}e^{-\lambda/\beta}I_{(0,\infty)}(\lambda),$$

and show that the posterior distribution of $\lambda|X_1, \dots, X_n$ is

$$\text{Gamma}(\sum_i X_i + \alpha, (n + 1/\beta)^{-1}).$$

(b) Show that the mean of the posterior distribution is

$$\hat{\lambda} = (\sum_i X_i + \alpha)/(n + 1/\beta).$$

(c) Find weights, so that $\hat{\lambda}$ is a weighted average of $\bar{X}$ and the mean of the prior distribution.

(d) Show that in the limit as $n \to \infty$, with $n^{-1}\sum_i X_i = \lambda^* + O(n^{-1})$ (some given value), if Λ is a random variable with the posterior distribution, then

$$n^{1/2}(\Lambda - \lambda^*) \to N(0, \lambda^*)$$

in distribution. (Hint: recall that the Gamma(α, β) distribution has characteristic function $(1 - \beta it)^{-\alpha}$).