Ph.D. Comprehensive Exam

August 2019

Instructions: *The exam has 6 problems, most with multiple parts. All of the problems will be graded. Write your code words on each of your answer sheets. Do not put your name on any of the sheets. Be clear, concise, and complete. All solutions should be rigorously explained.*

Problem 1. Recall that for real-valued sequences, $a_n = o(b_n)$ means that $a_n/b_n \to 0$. For each of the following statements, show that it is true or give a counterexample to show that it is false:

(a) $b_n = a_n + o(a_n) \Rightarrow b_n = a_n + o(1)$

(b) $b_n = a_n + o(1) \Rightarrow b_n = a_n + o(a_n)$

Problem 2. Let X_1, X_2, X_3 be exchangeable Bernoulli random variables with $P(X_i = 1) = 1/2$, $i = 1, 2, 3$. Is it possible that $E[X_i X_j] = 1/10$ for all distinct $i, j \in \{1, 2, 3\}$? If so, give an example joint distribution that satisfies these properties. Otherwise, prove that it is impossible for three Bernoulli random variables to satisfy these properties.

Problem 3. For $i \in \{1, ..., 12\}$, let $X_i \sim \mathrm{N}\left(\mu_x, {}^1/_\tau\right)$ and let $Y_i \sim \mathrm{N}\left(\mu_y, {}^1/_\tau\right)$. We are interested in doing inference on the difference in means, $\mu_x - \mu_y$.

(a) Give the form of a frequentist test statistic to test if $\mu_x = \mu_y$ when τ is assumed to be known. Derive the distribution of this frequentist test statistic.

(b) Give the form of a frequentist test statistic for the difference in means when τ is unknown and must be estimated. Derive the distribution of this test statistic. You may use the well-known fact that if $Z \sim \mathrm{N}(0, 1)$ and $W \sim \chi^2(n)$ with $Z \perp\!\!\!\perp W$, then $\frac{Z}{\sqrt{W/n}} \sim \mathrm{t}(n)$.

(c) Consider a conjugate Bayesian approach where μ_x and μ_y both have $\mathrm{N}\left(\mu_0, \frac{1}{\omega_0 \tau}\right)$ priors and τ has a $\mathrm{Gamma}\left(\frac{a}{2}, \frac{b}{2}\right)$ prior. Derive the posterior distributions for τ, $\mu_x|\tau$, and $\mu_y|\tau$. You may use the completing-the-square formula:

$$r(\mu - u)^2 + s(\mu - v)^2 = (r + s)(\mu - \hat{\mu})^2 + \frac{rs}{r + s}(u - v)^2, \quad \hat{\mu} = \frac{r}{r + s}u + \frac{s}{r + s}v.$$

(d) Consider the Bayesian test statistic $\frac{(\mu_x - \mu_y)|\tau}{\tau}$. Show that the posterior distribution of this test statistic is a generalized t distribution—that is, a t distribution which admits non-integer parameter values.

(e) Explain how to choose μ_0, ω_0, a, and b to give a Bayesian test statistic that approximates the standard frequentist test statistic for the difference in means.

Problem 4. Let $X_1, \dots, X_n$ be i.i.d. $N(\theta, 1)$.

(a) Suppose the parameter space is restricted to $\theta \in \Theta = \{0, 1\}$. What is the MLE for θ?

(b) In the same setting as (a), show that the MLE, $\widehat{\theta}$, is unbiased.

(c) Also in the same setting as (a), what is the distribution of $\widehat{\theta}$?

(d) Now suppose the parameter space is $\Theta = \{0, 1, 2, \dots\}$ (i.e., the nonnegative integers). Find the MLE for θ.

Problem 5. The purpose of this problem is to show that it is possible, in some sense, to get better estimates using an incorrect reduced linear model than you can get by using a correct full model. Suppose the standard linear model

$$Y = X\beta + e, \quad E(e) = 0, \quad Cov(e) = \sigma^2 I, \tag{1}$$

is correct and consider fitting a reduced model

$$Y = X_0\gamma + e, \quad \text{with } C(X_0) \subset C(X). \tag{2}$$

Under the standard linear model (1), the best fitted values one could ever have are $X\beta$ but we don't know β. For estimated fitted values, say, $F(Y)$, their quality can be measured by looking at

$$\mathrm{E}\left[(F(Y) - X\beta)' \, (F(Y) - X\beta)\right].$$

(a) Assuming that the full model is correct, for least squares estimates $X\hat{\beta} = MY$, show that

$$\mathrm{E}[(MY - X\beta)'(MY - X\beta)] = \mathrm{E}[(Y - X\beta)'M(Y - X\beta)] = \mathrm{tr}\left(M\sigma^2 I\right) = \sigma^2 r(X).$$

Now consider the reduced model (2) with M_0 the ppo onto $C(X_0)$. Estimate the fitted values from the reduced model with $X_0\hat{\gamma} = M_0 Y$.

(b) Show that when the full model is true

$$\mathrm{E}[(M_0Y - X\beta)'(M_0Y - X\beta)] = \sigma^2 r(X_0) + \|X\beta - M_0X\beta\|^2.$$

(c) Under what mathematical conditions does the reduced model give better estimates?

(d) Say something intelligent about when this might happen in practice.

Problem 6. Let $X_1, \dots, X_n$ be i.i.d. Bernoulli random variables with probability p, and let $Y_1, \dots, Y_n$ be i.i.d. Bernoulli with probability $\sqrt{p}$. The X_is are independent of the Y_is. Suppose there is interest in estimating p when p is small. Consider two methods for estimating p.

Method 1. Use $\widehat{p} = \overline{x}$.

Method 2. Use $\widetilde{p} = (\overline{y})^2$.

Give approximate large sample confidence intervals for p using both methods and give a condition for which Method 2 is better than Method 1. Justify your answer.