Ph.D. Comprehensive Exam

January 2020

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 $a_n/b_n \to 0$. Further recall that for real-valued sequences $\{a_n\}$ and $\{b_n\}$, $a_n \sim b_n$ if $\lim_{n\to\infty} a_n/b_n = 1$. For each of the following statements, show that it is true or give a counterexample to show that it is false:

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

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

Problem 2. Consider the probability field $(\Omega, \mathcal{F}, \mathcal{P})$ and let $X_1(\omega), X_2(\omega), \ldots : \Omega \to \mathbb{R}$ be random variables with $X_n(\omega)$ taking the values $-1, \frac{-1}{n+1}, \frac{1}{n+1}, 1$ each with probability 1/4. Further, for each $\omega \in \Omega$, assume that one of the following must hold for all $n \in \mathbb{N}$: $X_n(\omega) = -1$, $X_n(\omega) = 1$, or $-1 < X_n(\omega) < 1$. Let $X(\omega) : \Omega \to \mathbb{R}$ be defined such that when $X_1(\omega) = -1$ or $X_1(\omega) = 1$, $X(\omega) = X_1(\omega)$, and otherwise $X(\omega) = 0$.

(a) Show that $X_n(\omega) \xrightarrow{\text{a.s.}} X(\omega)$.

(b) Show that $X_n(\omega) \xrightarrow{\mathscr{L}} -X(\omega)$ but $X_n(\omega) \overset{\text{a.s.}}{\nrightarrow} -X(\omega)$. That is, X_n converges to $-X$ in distribution, but not almost surely.

Problem 3. Show that for any events, $A_1, \ldots, A_n$, $P[A_1 \cap A_2] \geq P(A_1) + P(A_2) - 1$, and more generally, that $P[\bigcap_{i=1}^n A_i] \geq \sum_{i=1}^n P(A_i) - (n-1)$. Discuss how this is related to Bonferroni adjustments when doing multiple testing of statistical hypotheses. Hint: Let A_i be the event that the ith decision is correct when making n hypothesis tests.

Problem 4. Consider a linear model

$$Y = X\beta + e, \qquad \mathrm{E}[e] = 0, \qquad \mathrm{Cov}\,[e] = \sigma^2 V,$$

where V is known and positive definite.

(a) If X is $n \times p$ with $r(X) = r$, what is the rank of the null space of X?

(b) Show that the null space of X equals the null space of VX.

(c) What is $r(VX)$?

(d) If $C(VX) \subset C(X)$, show that $C(VX) = C(X)$.

(e) Show that $C(V^{-1}X) = C(X)$.

Define the oblique projection operator onto $C(X)$, $A = X(X'V^{-1}X)^{-}X'V^{-1}$.

(f) Show that for any vector w, $w \perp C(X)$ implies $Aw = 0$.

(g) Show that A is the perpendicular projection operator onto $C(X)$.

(h) Explain why knowing $C(VX) = C(X)$ implies that least squares estimates are BLUEs for this model.

Problem 5. Let $X_1, \ldots, X_n$ be i.i.d. $N(\alpha, \alpha\beta)$. where $\alpha > 0$ and $\beta > 0$.

(a) What is the Fisher information matrix for $\theta = (\alpha, \beta)$?

(b) Find the MLE for β

(c) Give a confidence interval for β

(d) Give a large sample confidence interval for $\sqrt{\beta}$.

Problem 6. In Bayesian inference, Box's marginal p-value can be used to assess the appropriateness of a Bayesian model (e.g. a choice of parametric data model and prior distribution) for a given set of data. Let $X_1, ..., X_{20}$ be random variables each taking the value 0 or 1. Consider the data model $X_1, ..., X_{20} \overset{\text{iid}}{\sim} \text{Bern}\,(p)$ and the prior $p \sim \text{Beta}\,(1, 1)$.

(a) Find the marginal distribution for $Y = \sum_{i=1}^{20} X_i$. Simplify as much as possible. *(Hint:* $\Gamma(n+1) = n!$*)*

(b) Identify the mode of the marginal distribution for Y.

(c) Box's marginal p-value can be defined as the marginal probability (given the data model and the prior) of seeing data as unlikely as, or more unlikely than, the data that you obtained. Find Box's p-value when $Y = 20$.

(d) In 2-4 sentences, discuss what Box's p-value says about the appropriateness of Binomial data models with $\text{Beta}\,(1, 1)$ priors.