ST 705 FINAL EXAM

May 3, 2020

NAME:

- You have **3.5 hours** to complete this exam.
- Communication with other individuals is **not** permitted during this exam.

1. (3 points) Let X be an $n \times p$ matrix with $\text{rank}(X) = r$, and C be a $(p-r) \times p$ matrix with $\text{rank}(C) = p-r$, such that $\text{col}(X') \cap \text{col}(C') = \{0\}$. Show that

$$\text{rank}\begin{pmatrix} X \\ C \end{pmatrix} = p.$$

2. The problem of least squares regression can be understood as a special case of the more general problem of ridge regression. For an n-dimensional column vector y and an $n \times p$ design matrix X, the problem of ridge regression is to solve for the parameter vector b that minimizes

$$a\|b\|_2^2 + \|y - Xb\|_2^2,$$

where $a \geq 0$ is fixed.

   (a) (3 points) Derive a closed-form expression of the ridge regression solution.

   (b) (3 points) Assume that X has full column rank, and suppose that y is an observed instance of the random vector $Y = X\beta + U$, where $\beta \in \mathbb{R}^p$ is fixed and U satisfies the Gauss-Markov assumptions. Under what condition(s) is the ridge regression solution the best linear unbiased estimator (BLUE) for any β?

3. Suppose that $Y_i \sim \text{Binomial}(p, n_i)$ for $i \in \{1, \ldots, N\}$, and assume that $Y_1, \ldots, Y_N$ are independent.

   (a) (3 points) Write this as a linear model.

   (b) (3 points) Find the BLUE of p.

   (c) (3 points) Find the MLE of p. How does the variance of the MLE compare to the variance of the BLUE?

4. (3 points) Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables with $X_n \sim \chi_p^2(\phi_n)$ for all $n \geq 1$. For fixed p, show that if $\phi_n \to \infty$ as $n \to \infty$, then X_n converges in distribution to a normal distribution as $n \to \infty$, after centering and rescaling X_n.