Statistics comprehensive exam. January 2022

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

Problem 1. Consider a linear model

$$Y = X\beta + e, \qquad \mathrm{E}[e] = 0$$

with the (not necessarily estimable) linear constraint $\Lambda'\beta = d$.

(a) Characterize the reduced model associated with this constraint (hypothesis).

(b) Consider two solutions to the constraint, b_1 and b_2, so that $\Lambda' b_k = d$, $k = 1, 2$. Define appropriate least squares fitted values $\hat{Y}_k$.from the corresponding reduced models.

(c) Show that $\hat{Y}_1 = \hat{Y}_2$. Hints: Show that $(I - M_0)X(b_1 - b_2) = 0$. In what space does $(b_1 - b_2)$ lie?

Problem 2. For each n, let y_{ni}, $i = 1, \ldots, n$, be independent with mean 0 and variance σ^2_{ni}. Let $z_n \equiv \sum_{i=1}^n y_{ni}$ and $B_n^2 \equiv \mathrm{Var}[z_n] = \sum_{i=1}^n \sigma^2_{ni}$. We are going to look at how the Lindeberg Central Limit Theorem applies to exponentially weighted moving averages (EWMAs) of iid random variables x_i with $\mathrm{E}[x_i] = 0$, $\mathrm{Var}[x_i] = \sigma^2$. Define the EWMA as

$$\hat{\mu}_n \equiv (\alpha x_n + \alpha^2 x_{n-1} + \alpha^3 x_{n-2} + \cdots + \alpha^n x_1) \Big/ \sum_{i=1}^n \alpha^i = \frac{\sum_{i=1}^n \alpha^i x_{n-i+1}}{\sum_{i=1}^n \alpha^i},$$

for $0 < \alpha < 1$. Some algebra applied to power series establishes that

$$\mathrm{Var}[\hat{\mu}_n] = \sigma^2 \left(\frac{1+\alpha^n}{1-\alpha^n}\right)\left(\frac{1-\alpha}{1+\alpha}\right)$$

Relative to Lindeberg, define

$$z_n \equiv \hat{\mu}_n.$$

(a) State the Lindeberg Central Limit Theorem.

(b) What is y_{ni} in terms of the x_is? (WLOG assume $\mathrm{Var}[y_{ni}] < \mathrm{Var}[y_{nj}]$ when $i < j$.)

(c) What is B_n^2 and what does it converge to?

(d) Does $\hat{\mu}_n$ converge in probability to $\mathrm{E}[x_i]$?

(e) Does $\mathrm{E}\left[|y_{nn}|^2 \mathcal{I}_{[\epsilon B_n, \infty)}(|y_{nn}|)\right]$ converge to 0?

(f) Does the Lindeberg condition hold?

Problem 3. Let $X_1, \ldots, X_n$ be a sample from

$$f(x|b,g) = \frac{1}{b^g\,\Gamma(g)} x^{g-1} e^{-x/b}.$$

Show that there exists a UMP test for $H_0 : b \leq b_0$ versus $b > b_0$ when g is known. Find the form of the rejection region.

Problem 4. Let $y_1, ..., y_n$ be a random sample modeled by a Bern (p) distribution.

(a) Consider the case where p is unknown. Find the score function for p.

(b) Find the Fisher information for p.

(c) Find the Jeffreys prior for p.

(d) The Jeffreys prior is always a flat prior on some parameterization of the model it comes from. In the case where p is unknown, find the parameterization for the model that gives rise to a flat Jeffreys prior over $\mathbb{R}$.

Problem 5. A set of n counts $X = (X_1, X_2, \ldots, X_n)$ are modeled as

$$Pr(X_i = 0 \mid \gamma_i = 0, \pi, \lambda) = 1,$$

$$X_i \mid (\gamma_i = 1, \pi, \lambda) \sim Poisson(\lambda), \text{(indep. across i)}$$

where $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ is a set of n latent binary counts, $\pi \in [0, 1]$ and $\lambda > 0$ are assigned as what is known as an "hierarchical prior":

$$\gamma_i|(\pi, \lambda) \sim \text{iid } Bernoulli(\pi),\ \ i = 1, \ldots, n;$$

$$\pi|\lambda \sim Beta(1, c\lambda);$$

$$\lambda \sim Gamma(a, b)$$

for some positive constants a, b and c. ($Beta(1, r)$ has pdf $r[1 - \pi]^{r-1}, 0 \leq \pi \leq 1$)

(a) Show that the conditional prior pdf of λ given π is $Gamma(a + 1, b - c\ log\ [1 - \pi])$.

(b) Write down the posterior conditional probability distributions of $\pi|(\gamma, \lambda, x)$, $\lambda|(\gamma, \pi, x)$, and $\gamma|(\pi, \lambda, x)$ given data $x = (x_1, x_2, \ldots, x_n)$ on X. Answer in terms of conditional distributions with explicit formulas for their parameters and with appropriate use of conditional independence.

Problem 6. Let X_1, X_2, X_3 be exchangeable Bernoulli random variables with

$$P(X_i = 1) = 1/2, \quad i = 1, 2, 3.$$

Under this setup, it is not possible that $E[X_iX_j] = 1/10$ for all distinct $i, j \in \{1, 2, 3\}$. Show that this is impossible. (There are *at least* four arguments we would accept for why this is impossible.)