Statistics comprehensive exam. August 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 Bayesian model for 8 conditionally independent Bernoulli trials, where the probability of success on a trial is $\theta \sim \text{Beta}(1,1)$ and the total number of successes on 8 trials is $X|\theta \sim \text{Bin}(8,\theta)$. (For this problem, remember that $\Gamma(a) = (a-1)!$ for all $a \in \mathbb{N}$.)

(a) Find the marginal probability that X, the number of successes on 8 trials, is even: $\Pr[X \in \{0,2,4,6,8\}]$.

(b) Assume that $X = 2$ successes are observed—and then a further 4 trials are performed, exchangeable with the first 8 trials. Let Y be the total number of successes on these 4 new trials. Find the predictive probability that $Y = 0$.

Problem 2. Consider a linear model

$$Y = X\beta + e, \quad \text{E}[e] = 0, \quad \text{Cov}[e] = \sigma^2 I.$$

Let M be the perpendicular projection operator onto $C(X)$. Consider the parameter $\lambda'\beta$ where λ is known. For known vectors a and ρ, suppose $a'Y$ and $\rho'Y$ are unbiased estimates of $\lambda'\beta$.

(a) Show that $\lambda'\beta$ is estimable.

(b) Show that $a'X = \rho'X$.

(c) Show that $\text{Var}[a'Y] = \text{Var}[a'Y - \rho'MY] + \text{Var}[\rho'MY]$.

(d) State and prove the Gauss-Markov Theorem.

Problem 3. Let $X_1, X_2, \ldots$ be independent with X_n taking the values $\pm\sqrt{n-1}$ each with probability 1/2. Use a (not "the") Central Limit Theorem and other asymptotic results to show that the sample mean converges in distribution to a $\text{N}(0, 0.5)$.
Hints: $\sum_{i=1}^{r} i = r(r+1)/2$; $\sum_{i=1}^{r} i^2 = r(r+1)(2r+1)/6$.

Problem 4. Consider a random n-vector Y with $\mathrm{E}[Y] = \mu J$ and $\mathrm{Cov}\,[Y] = \sigma^2\,[(1-\rho)I + \rho JJ']$. Here J is a vector of 1s. Suppose ρ is known. Recall

$$\overline{y} = \frac{1}{n}J'Y \quad \text{and} \quad s^2 = \frac{1}{n-1}Y'[I - (1/n)JJ']Y.$$

(a) Show that $\mathrm{E}[s^2] = \sigma^2(1-\rho)$.

(b) Show that $\mathrm{Var}[\overline{y}] = (\sigma^2/n)[(1-\rho) + n\rho]$

(c) For sampling w/o replacement from a population of size N, $\mathrm{Cov}\,[y_i, y_j] = -\sigma^2/(N-1)$. Show that

$$\left(1 - \frac{n}{N}\right)\frac{s^2}{n}$$

is an unbiased estimate of $\mathrm{Var}[\overline{y}]$.

Problem 5. Let $U_1 \sim \mathrm{N}(\mu_1, 0.5)$, $U_2 \sim \mathrm{N}(\mu_2, 0.5)$, and $V \sim \mathrm{Bern}\,(p)$. The mixture random variable X is defined as

$$X = U_1^V U_2^{1-V}.$$

(a) Derive the cdf of X using conditional probabilities.

(b) Find the density of X.

(c) Prove that when $p = 0.5$ and $|\mu_1 - \mu_2| > 1$, X does not have a unimodal distribution.

Problem 6. 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 $C(V^{-1}X) = C(X)$.

Now 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.