Student ID (NOT your name):

<u>Final Examination: QUESTION BOOKLET</u>

Prof. Will Fithian

Fall 2024

- Do *NOT* open this question booklet until you are told to do so.
- Write your Student ID number (**NOT** your name) at the top of this page.
- Write your solutions in this booklet.
- No electronic devices are allowed during the exam.
- Be neat! If we can't read it, we can't grade it.
- You can treat any results from lecture or homework as "known," and use them in your work without rederiving them, but do make clear what result you're using. You do not need to explicitly check regularity conditions for the theorems from class that required them.
- For a multi-part problem, you may treat the results of previous parts as given (if you don't prove the result for part (a), you can still use it to solve part (b)).
- I have starred some parts which I believe are the most difficult, and which I expect most students won't necessarily be able to solve in the time allotted. They are generally not worth more points than the less difficult parts, so don't waste too much time on them until you're happy with your answers to the latter.
- Be careful to justify your reasoning and answers. We are primarily interested in your understanding of concepts, so show us what you know.

Good luck!

1. Six Gaussians (20 points, 5 points / part).

Some useful facts for this problem:

- Recall that the Gaussian density function for $Z \sim N(\theta, \sigma^2)$ is

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\theta)^2}{2\sigma^2}\right\}$$

Assume that we observe Gaussian random variables $X_1, \ldots, X_6$ where $X_i \sim N(\theta_i, \sigma^2)$, independently. Different parts of the question will assume $\sigma^2 > 0$ is known or unknown.

(a) Assume it is known that $\sigma^2 = 1$. Suppose we want to test the null hypothesis:

$$H_0 : \ \theta_2 = \theta_3 \text{ and } \theta_4 = \theta_5 = \theta_6,$$

against the alternative that θ is any other vector in $\mathbb{R}^6$. Suggest a χ^2 test statistic and specify the degrees of freedom.

(b) Continue to assume $\sigma^2 = 1$ and consider the following estimator for θ:

$$\delta(X) = \gamma \cdot \left(X_1, \overline{X}_{23}, \overline{X}_{23}, \overline{X}_{456}, \overline{X}_{456}, \overline{X}_{456}\right),$$

where $\gamma \in [0, 1]$ is a fixed constant, $\overline{X}_{23} = \frac{X_2+X_3}{2}$, and $\overline{X}_{456} = \frac{X_4+X_5+X_6}{3}$. Give an unbiased estimator for the MSE of $\delta(X)$.

(c) Now, assume that σ^2 is unknown, but it *is* known that $\theta_2 = \theta_3$ and $\theta_4 = \theta_5 = \theta_6$. That is, what in part (a) was a null hypothesis to be tested is now a modeling assumption. Suggest a confidence interval based on the Student's t-distribution for the parameter $g(\theta) = \theta_4 - \theta_3$. Specify the degrees of freedom.

(d) Under the same assumptions as part (b), now suppose that we want to test the null hypothesis $H_0 : \ \theta_1 = \theta_2 = \cdots = \theta_6$ against the alternative that θ is any other vector in $\mathbb{R}^6$ with $\theta_2 = \theta_3$ and $\theta_4 = \theta_5 = \theta_6$. Suggest an F test statistic and specify the degrees of freedom.

Problem 1 answers continued (1):

Problem 1 answers continued (2):

Problem 1 answers continued (3):

2. Inverse gamma prior (20 points, 5 points / part). Some useful facts for this problem:

- A χ^2_d random variable has mean d and variance $2d$.
- If Y is a Gamma(α, β) random variable (in its "rate parameterization") then it has density
$$\frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} \exp\{-\beta y\},$$
on $(0, \infty)$. Y has mean α/β and variance α/β^2. This distribution is defined for $\alpha, \beta > 0$.
- The inverse-gamma distribution (denoted $IG(\alpha, \beta)$) is the distribution of $W = 1/Y$ where $Y \sim$ Gamma(α, β). Then $W \in (0, \infty)$ has the density
$$\frac{\beta^\alpha}{\Gamma(\alpha)} w^{-\alpha-1} \exp\{-\beta/w\}.$$
Note that β is a scale parameter for W. W has mean $\frac{\beta}{\alpha-1}$ provided $\alpha > 1$, and variance $\frac{\beta-1}{(\alpha-1)^2(\alpha-2)}$ provided $\alpha > 2$. This distribution is likewise defined for $\alpha, \beta > 0$.
- Define the *squared relative error* loss function
$$L_{\text{rel}}(d, \theta) = \left(\frac{d-\theta}{\theta}\right)^2 = \left(\frac{d}{\theta} - 1\right)^2,$$
and define the corresponding risk function $R_{\text{rel}}(\delta(\cdot), \theta) = \mathbb{E}_\theta[L_{\text{rel}}(\delta(X), \theta)]$.

Consider the Bayesian model with
$$\theta \sim IG(\alpha, \beta),$$
$$X_1, \ldots, X_n \mid \theta \overset{\text{i.i.d.}}{\sim} N(0, \theta)$$

Note that the variance is θ, not θ^2, and assume $n \geq 2$.

(a) Find the posterior distribution of θ given $X = (X_1, \ldots, X_n)$ and the Bayes estimator for θ under the standard (not relative) squared error loss.

(b) Find the Bayes estimator for θ under the squared relative error loss L_{rel}.

(c) (*) For the estimator in part (b), find the risk function $R_{\text{rel}}(\delta(\cdot), \theta)$ as a function of θ and show that the Bayes risk is $\frac{2}{n+2(\alpha+1)}$.

(d) For the relative squared error risk, find a linear estimator of the form $\delta(X) = a \sum_{i=1}^n X_i^2$ that is minimax, and prove it is minimax.

Problem 2 answers continued (1):

Problem 2 answers continued (2):

Problem 2 answers continued (3):

3. Social network model (20 points, 5 points / part).

Some useful facts you may assume are true for this problem:

- An *undirected graph* is a set of vertices V and a set of edges E connecting pairs of vertices. Assume (without loss of generality) that the edges are labeled 1 to m (so $V = \{1, \ldots, m\}$) and each edge is represented by a pair of vertices (i, j) with $1 \le i < j \le m$; that is, vertices i and j are connected to each other if $(i, j) \in E$, in which case we say the edge (i, j) is present.
- The binomial distribution Binom(n, θ) with parameter $\theta \in (0, 1)$ has probability mass function

$$p_\theta(x) = \mathbb{P}_\theta(X = x) = \binom{n}{x} \theta^x (1 - \theta)^{n-x}, \quad \text{for } x \in \{0, 1, \ldots, n\}.$$

  Its mean and variance are

$$\mathbb{E}_\theta[X] = n\theta, \quad \text{Var}_\theta(X) = n\theta(1 - \theta).$$

  The binomial distribution arises when a coin lands heads with probability θ, and we flip it n times. Then, the number of heads we see is a Binom(n, θ) random variable.

We will consider a model where the set V of vertices is fixed but the set E of edges is random, governed by parameters that we are interested in. Define $X_{i,j} \in \{0, 1\}$ as a binary indicator that (i, j) is present. Assume that for each pair (i, j), $X_{i,j} \sim \text{Bern}(\pi_{i,j})$, for $\pi_{i,j} \in (0, 1)$, and the $X_{i,j}$ values are independent.

Assume this model represents a social network, where each vertex represents an individual student in a school and an edge represents a friendship relation between two students; students i and j are friends if $(i, j) \in E$.

In addition, assume each student belongs to a group $g(i) \in \{1, \ldots, K\}$. Students in the same group may have a higher chance of forming friendships than students in different groups. We assume that it is known which group each student belongs to.

(a) Suppose that $\pi_{i,j} = \alpha + \beta 1\{g(i) = g(j)\}$, for unknown parameters $\alpha \in [0, 1]$ and $\beta \in [0, 1 - \alpha]$ (note β is non-negative). Let

$$T_w = \sum_{i<j:\ g(i)=g(j)} X_{i,j},$$

the total number of friendships between pairs of students within the same group, and let

$$T_b = \sum_{i<j:\, g(i)\neq g(j)} X_{i,j},$$

the total number of friendships between pairs of students in different groups. Show that (T_w, T_b) is a complete sufficient statistic for this model.

(b) Find the maximum likelihood estimator $\hat{\alpha}, \hat{\beta}$.

(c) (*) Now (for this part only) assume $\alpha \in (0, 1)$ is known. Does there exist an admissible unbiased estimator for β, for the squared error loss?

(d) (*) Now assume that the β parameter possibly varies by group. That is, $\pi_{i,j} = \alpha + \sum_{k=1}^{K} \beta_k 1\{g(i) = g(j) = k\}$, where $\alpha \in [0, 1]$ and $\beta_1, \ldots, \beta_K \in [0, 1-\alpha]$ are all unknown.

   Assume we want to test the hypothesis $H_0 : \ \beta_1 = \ldots = \beta_K = 0$ against the alternative $H_1 : \ \max_k \beta_k > 0$. Assume we have an estimator $\hat{\beta}^*(X)$ for the parameter $\beta^* = \max_k \beta_k$, and we want to will use a test that rejects for large values of $\hat{\beta}^*(X)$. How could we carry out an exact (finite-sample) test using $\hat{\beta}^*$ as the test statistic? You do not need to give an explicit formula for the threshold, but explain how you would calculate it either in words or pseudocode.

Problem 3 answers continued (1):

Problem 3 answers continued (2):

Problem 3 answers continued (3):

4. Estimation in the Geometric model (25 points, 5 points / part).

Some useful facts for this problem:

- The geometric distribution Geom(θ) with parameter $\theta \in (0,1)$ has probability mass function

$$p_\theta(x) = \mathbb{P}_\theta(X = x) = (1-\theta)^x \theta, \quad \text{for } x \in \{0, 1, 2, \ldots\}.$$

  Its mean and variance are

$$\mathbb{E}_\theta[X] = \frac{1-\theta}{\theta}, \quad \text{Var}_\theta(X) = \frac{1-\theta}{\theta^2}.$$

  The geometric distribution arises when a coin lands heads with probability θ, and we flip it repeatedly until it lands heads. Then, the number of tails we see before the first heads is a Geom(θ) random variable.

Assume throughout this problem that we observe $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \text{Geom}(\theta)$.

(a) Find a minimal sufficient statistic for the distribution. Is it complete?

(b) Let $\hat{\theta}(X)$ denote the maximum likelihood estimator for θ, where $X = (X_1, \ldots, X_n)$. Give an explicit formula.

(c) Show that $\hat{\theta}(X)$ is consistent and use it to construct a Wald confidence interval for θ. Give an explicit formula.

(d) Define the tail probability $\tau_k(\theta) = (1-\theta)^k$ to be the probability that a single observation is at least k. That is,

$$\tau_k(\theta) = \mathbb{P}_\theta(X_1 \geq k) = (1-\theta)^k.$$

Give the asymptotic distribution for the maximum likelihood estimator $\hat{\tau}_k(X) = \tau_k(\hat{\theta}(X))$ when k is fixed and $n \to \infty$ (appropriately centered and scaled).

(e) (*) Assume that $n = 4$, and $(X_1, X_2, X_3, X_4) = (0, 3, 4, 2)$. Evaluate the UMVU estimator for $\tau_{10}(\theta)$ on the given data set. Your answer should be a number, and you should justify how you calculated it.

Problem 4 answers continued (1):

Problem 4 answers continued (2):

Problem 4 answers continued (3):