UNL Statistics PhD Qualifying Exam - January 2025

Print Your Qualifying Exam ID: ______________________

- **Date: January 16th (8:00am – 1:00pm, Thursday) and January 17th, 2025**
- **Location: Room 142 Hardin Hall**
- **Please put your phones on vibrate/silent mode and don't use during the exam.**
- The exam is a written exam over the MS core courses that assesses preparedness for the PhD program. Students are allowed to take the exam if they have a GPA of at least 3.5 in their MS and PhD core courses taken.
- You will be assigned an ID number.
- Closed book and note. You will not have access to computers or the Internet. You may not use your phones for any reason. Bring a calculator.
- Your answers will be hand-written. We will provide you with paper. Please ensure that your answers are written as clearly as possible.
  - The assigned ID number must be written on each page of the answerscript in place of your name.
  - Start each problem on a new sheet of paper (you do not need to start each part of each problem with a new sheet).
    - (Ex.) Problem 1, parts 1-5 should be numbered: 1.1, 1.2, 1.3 . . . etc.
    - Example of problems with lettered parts and subsections: 6.a, 6.b...6.e(a), 6.e(b), 6.e(c) . . . etc.
  - When you are finished, number **all** pages of work in the bottom right corner 1 – x pages.
  - All answers must be written only on one side of the paper.

Day 1

1. (100 points) $X_1, X_2, \cdots, X_n$ are independent and identically distributed (i.i.d.) random variables from $N(\mu, \sigma^2)$.

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{and} \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

denote the sample mean and sample variance. Prove $\bar{X}$ and S^2 are independent.

2. (100 points) Consider an i.i.d. random sample $X_1, \ldots, X_n$ from Uniform$[0, \theta]$.

   (a) (20 points) Find the maximum likelihood estimator (MLE) $\hat{\theta}$ of θ.

   (b) (30 points) Is the MLE an unbiased estimator of θ ?

   (c) (25 points) Obtain the mean squared error (MSE) of $\hat{\theta}$.

   (d) (25 points) Use the method of pivoting a continuous cumulative distribution function to obtain a 95% confidence interval for θ.

3. (100 points) Let $X_1, X_2, \ldots, X_n$ be i.i.d. Poisson(λ) random variables with $\lambda > 0$.

   (a) (20 points) Find a sufficient statistic $T(X_1, \ldots, X_n)$ for λ.

   (b) (20 points) Show that the family of distribution of the above sufficient statistic T is complete.

   (c) (30 points) Using (a) and (b), find the UMVUE for $1 - e^{-\lambda}$.

   (d) (30 points) Find the asymptotic distribution of $1 - e^{-\lambda}$.

4. (100 points) Let $X_1, \ldots, X_n$ be a random sample from the Uniform$(\theta, \theta+1)$ distribution. To test $H_0 : \theta = 0$ versus $H_1 : \theta > 0$, we use the test

$$\text{reject } H_0 \text{ if } Y_n \geq 1 \text{ or } Y_1 \geq k,$$

where k is a constant, $Y_1 = \min\{X_1, \ldots, X_n\}, Y_n = \max\{X_1, \ldots, X_n\}$.

(a) (20 points) Find the joint distribution of Y_1 and Y_n.

(b) (20 points) Determine k so that the test will have size α.

(c) (30 points) Find an expression for the power function of the test in part (b).

(d) (30 points) Prove that the test is UMP size α.

5. (100 points) Consider the $p-$dimensional random vector $\boldsymbol{X}_i^{p\times 1} \sim MVN_p(\boldsymbol{\mu}_i, \Sigma),\ i = 1, 2, 3$. Suppose $\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3$ are mutually independent. Define $\bar{\boldsymbol{X}} = \frac{\boldsymbol{X}_1+\boldsymbol{X}_2+\boldsymbol{X}_3}{3}$. Let $\boldsymbol{Y}_1 = \boldsymbol{X}_1 - \bar{\boldsymbol{X}}$ and $\boldsymbol{Y}_2 = \bar{\boldsymbol{X}}$.

   (a) Find the distribution of $\boldsymbol{Y}_1$.

   (b) Find the joint distribution of $(\boldsymbol{Y}_1, \boldsymbol{Y}_2)$. Is $\boldsymbol{Y}_1$ independent of $\boldsymbol{Y}_2$? Why, or why not?

6. (100 points) Consider n iid samples from a $p-$variate normal distribution, that is, $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n \overset{iid}{\sim} MVN_p(\boldsymbol{\mu}, \Sigma)$. For pre-specified $\boldsymbol{\mu}_0$ and Σ_0, we wish to test the hypotheses

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0,\ \Sigma = \Sigma_0$$
$$vs.$$
$$H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0,\ \Sigma \neq \Sigma_0.$$

Obtain the likelihood ratio test statistic for this test.

7. Assume data from a block design are to be analyzed using the linear mixed model $y_{ij} = \mu + \tau_i + u_j + \varepsilon_{ij}$, where $u_j \sim \text{NID}(0, \sigma_u^2)$ and $\varepsilon_{ij} \sim \text{NID}(0, \sigma^2)$.

   (a) Prove that the covariance of two y_{ij} values within block is $\text{cov}(y_{ij}; y_{i'j}) = \sigma_u^2$.

   (b) Assume the τ_i represents 3 randomized treatments (i.e. $i = 1, 2, 3$). Give the 3×3 covariance matrix of the responses in the jth block, $y_i' = (y_{1j} y_{2j} y_{3j})$, i.e., give $\text{V}(y_j)$.

   (c) Show how the covariance matrix above will be incorporated into a generalized least squares estimator of the treatment means $\mu + \tau_i$.

   (d) Use the model and conditional expectation to illustrate the concept of a BLUP for $\mu + \tau_i + u_j$.

   (e) Now assume the τ_i represents 3 repeated measurements (i.e. $i = 1, 2, 3$) and that the residuals follow a heterogeneous AR(1) process within a block. Give the 3×3 covariance matrix of the residuals for the jth block, $\underline{\varepsilon}_j' = (\varepsilon_{1j} \varepsilon_{2j} \varepsilon_{3j})$, that is, give $\text{V}(\varepsilon_j)$.

8. Assume that a binary response, $y_i = 0$ or 1 where $P(Y_i = 1) = p_i$ for the i th observation, is to be modeled as a function of $\underline{x}\beta$.

   (a) Prove that the expectation and variance of Y_i are p_i and $p_i(1 - p_i)$.

   (b) Assume you will model Y_i using a generalized linear model with the logit link, such that is $\log_e(p_i/(1 - p_i)) = \underline{x}\underline{\beta}$. Show that $p_i = \exp(\underline{x}_i\beta)/(1 + \exp(\underline{x}_i\beta))$.

   (c) With a binary (Bernoulli) random variable, what happens to the model parameters if the data are over dispersed?

   (d) Give two different ways to identify overdispersion.

   (e) Give two different ways to account for overdispersion.