Quiz 01

Drew Davison

02-09-2025

Question 1

Which member of the teaching team has office hours Tuesdays 10:30 - 11:30am?

According to the course syllabus, no member of the teaching team has office hours Tuesdays 10:30 - 11:30 AM. The office hours per the syllabus for the teaching team are as follows:

Professor Rebecca Steorts - MW 1:00 PM - 2:00 PM

Suchismita Roy - TuTh 4:00 PM – 5:00 PM

Wenxin Guo MW 4:30 PM – 5:30 PM

Question 2

If you have a question about the course content, how should you ask it? Select all that apply.

a. Ask in office hours b. Post in Slack c. Ask in Slack

According to the course syllabus, the only correct answer is option **"a."** to ask in office hours. There is no Slack for the course. However, Ed Discussion is an alternative option to ask questions about course content.

Question 3

If a homework assignment is due on Wednesday at 11:59pm, when is the latest it can be turned in for credit?

The latest it can be turned in for credit is the due date Wednesday at 11:59pm. Per the course syllabus, there are no late work extensions, as the lowest grades are dropped.

Question 4

The course policy on the use of artificial intelligence is based on two guiding principles: (1) Cognitive dimension and (2) Ethical dimension. Briefly describe what each dimension means.

   1. The course policy on the Cognitive dimension of artificial intelligence states that:

"Working with AI should not reduce your ability to think clearly. We will practice using AI to facilitate— rather than hinder—learning."

This means that when using AI a student should be learning and understanding the material. Simply copying and pasting AI produced material does not facilitate learning.

2. The course policy on the Ethical dimension of artificial intelligence states that:

"Students using AI should be transparent about their use and make sure it aligns with academic integrity"

Practically, students are required to cite any use of AI in the production of code, and students are not permitted to use AI to produce narrative. Students should also comply with all Duke Acadmeic Policies regarding academic integrity.

Question 5

Suppose data from 100 randomly selected elementary school students are used to fit a regression model to predict height (in inches) based on age (in months). Which of the following best describes the normality assumption for this model?

The residuals, being the differences between the observed and predicted heights (in inches), are normally distributed in the model.

Question 6

Let $Y \sim \text{Binomial}(n, \theta)$. The likelihood function is given by:

$$L(\theta) = P(Y = y \mid \theta) = \binom{n}{y} \theta^y (1 - \theta)^{n-y}$$

where n is known and θ is the unknown parameter.

Part 2 of Question 6:

Since the binomial coefficient $\binom{n}{y}$ does not depend on θ, it can be ignored in maximization. Taking the natural logarithm of the likelihood function (log-likelihood):

$$\ell(\theta) = \log L(\theta) = \log \binom{n}{y} + y \log \theta + (n - y) \log(1 - \theta)$$

To find the maximum likelihood estimator $\hat{\theta}_{MLE}$, we differentiate the log-likelihood function with respect to θ:

$$\frac{d}{d\theta} \ell(\theta) = \frac{y}{\theta} - \frac{n - y}{1 - \theta}$$

Setting the derivative equal to zero:

$$\frac{y}{\theta} - \frac{n - y}{1 - \theta} = 0$$

Solving for θ:

$$\frac{y}{\theta} = \frac{n - y}{1 - \theta}$$

Cross multiplying:

$$y(1 - \theta) = (n - y)\theta$$

$$y - y\theta = n\theta - y\theta$$

$$y = n\theta$$

$$\hat{\theta}_{MLE} = \frac{y}{n}$$

Thus, the maximum likelihood estimator of θ is:

$$\hat{\theta}_{MLE} = \frac{Y}{n}$$

Question 7

Provide the **three** assumptions of a generalized linear model that may deviate from a linear model. State the **one** assumption they share in common.

The three assumptions that may deviate from a linear model are:

Relationship between response and predictor(s) can be nonlinear.

Response variable can be non-normal.

Variance in response can differ at each level of predictor(s).

The assumption that they share in common is:

The independence assumption still must hold.

Question 8

Describe, in your own words, the importance of exploratory data analysis (in less than two sentences).

Exploratory data analysis is important because it provides context to the variables, observations, and their relationships, as well as identifying anomalies (outliers or data errors), and it aides in the model selection process.