

Quiz 1

Sonya Eason

Question 1

There is not a member of the teaching team with office hours at that time.

Question 2

a) ask in office hours.

Question 3

Wednesday at 11:59 PM. There are no late extensions.

Question 4

- 1) cognitive dimension: The purpose of assignments are to foster learning, so if a student isn't actually utilizing these assignments to learn, they are essentially pointless. In the context of AI use, this means a student can use it as a resource but they need to ensure they are still taking on cognitive tasks involved in assignment completion. This can look like not just blindly copying the code but also asking why this method was done, why it works, how it relates to class content. In addition, prompts shouldn't be the assignment prompts copy and pasted into the tool, but rather guiding questions that actually address the student's content weaknesses.
- 2) ethical dimension: Because using AI involves ideas or guidance from another source, it's important to cite the use. Otherwise, it appears that the student completed the task without this outside help. For this reason, being transparent or open about the use of AI is very important.

Question 5

The normality assumption would mean that at each level of student age, the height of students follows a normal distribution. Specifically, at 80 months, the heights of students would be normally distributed, and at 85 months, the heights of students would also be normally distributed.

Question 6

$$Y \sim \text{Binomial}(n, \theta)$$
$$L(\theta) = \binom{n}{y} \theta^y (1 - \theta)^{n-y}$$

The value of θ at which the likelihood function is the same value of θ at which the log likelihood function is maximized. This is due to the monotonicity of the log function.

For the sake of numerical ease, we will look at the log likelihood function.

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

To look for extrema, we will take the derivative of our function with respect to our parameter of interest, θ . Then, we will set it to 0.

$$\frac{\partial}{\partial \theta} \log(L(\theta)) = \frac{y}{\theta} - \frac{n - y}{1 - \theta} = 0$$

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

After cross multiplication and algebraic simplification, we can solve to find $\hat{\theta}_{MLE}$

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

Knowing that this method solves for extrema, and not taking into account the shape of the log likelihood function, we can verify this is a value where our log likelihood is maximized (and not minimized) by checking the second derivative.

$$\frac{\partial^2}{\partial \theta^2} \log(L(\theta)) = \frac{-y}{\theta^2} - \frac{(n - y)}{(1 - \theta)^2}$$

Because we know y is positive and $n - y$ must be ≥ 0 , it's clear that the second derivative of the log likelihood with respect to θ will be negative. This is confirmation that the value we found is in fact one where our likelihood is maximized.

Question 7

For three assumptions, GLMs deviate from linear regression. For example, linearity assumption, normality assumption, and equal variance assumption differ. The linearity assumption which says there should be a linear relationship between mean response and explanatory variable doesn't necessarily hold for GLMs. This is really easy to see in the case of a Poisson regression, where instead we view the log of the rate parameter as having linearity with its response. Additionally, GLMs doesn't requires that the response be normally distributed for each level of the predictors. The constant variance condition, or the requirement that variance of responses be equal for each level of X , is also no longer required. In fact, if we think of Poisson regression, our variance will continue to increase as the mean increases, so this is clearly violated.

One assumption that GLMs and linear models have in common is the independence assumption, the assumption that each observation occurred independently of the others. For example, a violation of this could be looking at student scores when students come from different schools.

Question 8

Exploratory data analysis enables statisticians to identify how potential predictor variables relate to the response variable, which serve informative in developing statistical models that match these relationships. Exploratory data analysis can also be helpful in determining whether model conditions are met.

Question 9

The maximum likelihood estimator is the value of a parameter where seeing the fixed data is most likely. This is where the likelihood is maximized.

While the MLE is better than some arbitrary value, it's not necessarily the best estimate. For example, in the case where information beyond fixed data is informative like if you have some outside knowledge beyond the observed data, then a Bayesian estimator may be preferable.

Question 10

The likelihood is a function that tells us probability of seeing the fixed data under different parameters. Besides talking about seeing the fixed data based on variable parameter values, We can not really make probability statements with the likelihood because we do not know the true parameter and the output would just tell us the likelihood of our fixed data, not the probability of different outcomes that we would use a probability function to describe.