

# Math 341 / 641 Fall 2025

## Midterm Examination One

Professor Adam Kapelner

Sept 30, 2025

Full Name \_\_\_\_\_

### Code of Academic Integrity

Since the college is an academic community, its fundamental purpose is the pursuit of knowledge. Essential to the success of this educational mission is a commitment to the principles of academic integrity. Every member of the college community is responsible for upholding the highest standards of honesty at all times. Students, as members of the community, are also responsible for adhering to the principles and spirit of the following Code of Academic Integrity.

Activities that have the effect or intention of interfering with education, pursuit of knowledge, or fair evaluation of a student's performance are prohibited. Examples of such activities include but are not limited to the following definitions:

**Cheating** Using or attempting to use unauthorized assistance, material, or study aids in examinations or other academic work or preventing, or attempting to prevent, another from using authorized assistance, material, or study aids. Example: using an unauthorized cheat sheet in a quiz or exam, altering a graded exam and resubmitting it for a better grade, etc.

I acknowledge and agree to uphold this Code of Academic Integrity.

\_\_\_\_\_  
signature

\_\_\_\_\_  
date

### Instructions

This exam is 110 minutes (variable time per question) and closed-book. You are allowed **one** page (front and back) of a “cheat sheet”, blank scrap paper (provided by the proctor) and a graphing calculator (which is not your smartphone). Please read the questions carefully. Within each problem, I recommend considering the questions that are easy first and then circling back to evaluate the harder ones. No food is allowed, only drinks.

**Problem 1** Consider the following rv and some of its properties:

$$X \sim \text{Gamma}(\theta_1, \theta_2) := \frac{\theta_2^{\theta_1}}{\Gamma(\theta_1)} x^{\theta_1-1} e^{-\theta_2 x} \mathbb{1}_{x>0}, \quad \theta_1 > 0, \quad \theta_2 > 0, \quad \mathbb{E}[X] = \frac{\theta_1}{\theta_2}, \quad \text{Var}[X] = \frac{\theta_1}{\theta_2^2}$$

We realize  $n = 100$  iid realizations from this DGP and calculate the following quantities rounded to two decimals:

$$\begin{aligned} \bar{x} &:= \frac{1}{n} \sum_{i=1}^n x_i &= 3.51 \\ \frac{1}{n} \sum_{i=1}^n x_i^2 &= 7.89 \\ \frac{1}{n} \sum_{i=1}^n x_i^3 &= 29.62 \\ \frac{1}{n} \sum_{i=1}^n x_i^4 &= 127.21 \\ \frac{1}{n} \sum_{i=1}^n x_i^5 &= 602.66 \end{aligned}$$

- (a) [3 pt / 3 pts] Observing  $n = 100$  iid realizations from this DGP is equivalent to sampling an SRS of  $n = 100$  from a population with which property(ies)?

**Infinitely large**

- (b) [3 pt / 6 pts] Find the method of moments (MM) estimator for  $\mathbb{E}[X]$ .

**$\bar{X}$**

- (c) [3 pt / 9 pts] What is the MM estimate for  $\mathbb{E}[X]$ ? Round to the nearest two decimals.

**$\bar{x} = 3.51$**

(d) [6 pt / 15 pts] Find the MM estimator for  $\theta_2$ .

$$\begin{aligned}\mathbb{E}[X] &= \hat{\mu}_1 = \frac{\theta_1}{\theta_2}, \quad \mathbb{V}\text{ar}[X] = \hat{\mu}_2 - \hat{\mu}_1^2 = \frac{\theta_1}{\theta_2^2} \\ \Rightarrow \quad \theta_1 &= \hat{\mu}_1 \theta_2 \quad \Rightarrow \quad \hat{\mu}_2 - \hat{\mu}_1^2 = \frac{\hat{\mu}_1 \theta_2}{\theta_2^2} = \frac{\hat{\mu}_1}{\theta_2} \quad \Rightarrow \quad \hat{\theta}_2^{\text{MM}} = \frac{\hat{\mu}_1}{\hat{\mu}_2 - \hat{\mu}_1^2}\end{aligned}$$

(e) [3 pt / 18 pts] What is the MM estimate for  $\theta_2$ ? Round to the nearest two decimals.

$$\hat{\theta}_2^{\text{MM}} = \frac{\hat{\mu}_1}{\hat{\mu}_2 - \hat{\mu}_1^2} = \frac{3.51}{7.89 - 3.51^2} = -0.79$$

(f) [4 pt / 22 pts] Does the MM estimate for  $\theta_2$  make sense? Why or why not?

No, it is negative and the parameter space for  $\theta_2$  is positive.

(g) [5 pt / 27 pts] Find a MM estimator for  $\theta_1$ .

We use intermediate results and our result found in (d)

$$\theta_1 = \hat{\mu}_1 \theta_2, \quad \hat{\theta}_2^{\text{MM}} = \frac{\hat{\mu}_1}{\hat{\mu}_2 - \hat{\mu}_1^2} \quad \Rightarrow \quad \hat{\theta}_1^{\text{MM}} = \hat{\mu}_1 \left( \frac{\hat{\mu}_1}{\hat{\mu}_2 - \hat{\mu}_1^2} \right) = \frac{\hat{\mu}_1^2}{\hat{\mu}_2 - \hat{\mu}_1^2}$$

Now consider the scenario where we know that the value of  $\theta_1 = 4$ . The constant in the denominator can be calculated to be 6. Now, there is only one parameter, so we simplify the expression and just denote it  $\theta$ . Thus,

$$X \sim \text{Gamma}(4, \theta) := \frac{\theta^4}{6} x^3 e^{-\theta x} \mathbf{1}_{x>0}$$

- (h) [7 pt / 34 pts] Show that  $\hat{\theta}^{\text{MLE}} = \frac{4}{\bar{X}}$  for general sample of size  $n$ . All data is properly realized and are positive values and thus you can drop the indicator term during your calculation.

$$\mathcal{L}(\theta; \mathbf{X}) = \prod_{i=1}^n \frac{\theta^4}{6} X_i^3 e^{-\theta X_i} = \frac{\theta^{4n}}{6^n} X_i^{3n} \exp \left( -\theta \sum_{i=1}^n X_i \right)$$

$$\ell(\theta; \mathbf{X}) = 4n \ln(\theta) - n \ln(6) - \theta \sum_{i=1}^n X_i$$

$$\ell'(\theta; \mathbf{X}) = \frac{4n}{\theta} - \sum_{i=1}^n X_i \stackrel{\text{set}}{=} 0 \Rightarrow \frac{4n}{\theta} = \sum_{i=1}^n X_i \Rightarrow \hat{\theta}^{\text{MLE}} = \frac{4n}{\sum_{i=1}^n X_i} = \frac{4}{\bar{X}}$$

- (i) [3 pt / 37 pts] Show that this MLE is indeed a maximum of the log-likelihood by using a second derivative check. Explain your reasoning.

$$\ell''(\theta; \mathbf{X}) = -\frac{4n}{\theta^2} < 0 \quad \text{since } n \in \mathbb{N}, \text{ the sample size and } \theta > 0, \text{ the parameter space}$$

(j) [6 pt / 43 pts] Derive the CRLB for  $\theta$  in the DGP  $X \sim \text{Gamma}(4, \theta)$ .

$$I_n(\theta) = \mathbb{E}[-\ell''(\theta; \mathbf{x})] = \mathbb{E}\left[-\left(-\frac{4n}{\theta^2}\right)\right] = \frac{4n}{\theta^2}, \quad \mathbb{V}\text{ar}\left[\hat{\theta}\right] \geq \frac{1}{I_n(\theta)} = \frac{\theta^2}{4n}$$

(k) [1 pt / 44 pts] It can be shown that

$$\mathbb{E}\left[\hat{\theta}^{\text{MLE}}\right] = \frac{4n}{4n-1}\theta.$$

Is  $\hat{\theta}^{\text{MLE}}$  unbiased? Yes or no.

No

(l) [6 pt / 50 pts] It can be shown that

$$\mathbb{V}\text{ar} \left[ \hat{\theta}^{\text{MLE}} \right] = \frac{(4n)^2}{(4n-1)^2(4n-2)} \theta^2.$$

Calculate  $\text{MSE} \left[ \hat{\theta}^{\text{MLE}} \right]$  as a function of  $n$  and  $\theta$ . Simplify as much as you can.

$$\begin{aligned} \text{MSE} \left[ \hat{\theta}^{\text{MLE}} \right] &= \mathbb{B}\text{ias} \left[ \hat{\theta}^{\text{MLE}} \right]^2 + \mathbb{V}\text{ar} \left[ \hat{\theta}^{\text{MLE}} \right] \\ &= \left( \frac{4n}{4n-1} \theta - \theta \right)^2 + \frac{(4n)^2}{(4n-1)^2(4n-2)} \theta^2 \\ &= \left( \left( \frac{4n}{4n-1} - 1 \right)^2 + \frac{(4n)^2}{(4n-1)^2(4n-2)} \right) \theta^2 \\ &= \left( \left( \frac{1}{4n-1} \right)^2 + \frac{(4n)^2}{(4n-1)^2(4n-2)} \right) \theta^2 \\ &= \left( \frac{(4n-2)}{(4n-1)^2(4n-2)} + \frac{(4n)^2}{(4n-1)^2(4n-2)} \right) \theta^2 \\ &= \frac{(4n)^2 + 4n - 2}{(4n-1)^2(4n-2)} \theta^2 \end{aligned}$$

(m) [5 pt / 55 pts] Under squared error loss, compute  $\sup_{\theta \in \Theta} \left\{ R(\hat{\theta}^{\text{MLE}}, \theta) \right\}$ .

$\infty$

**Problem 2** Consider a survey of iPhoneness.  $n = 27$  students were chosen outside of Kiely Hall on Tuesday afternoon, Sept 30, 2025 were asked if they had an iPhone ( $x_i = 1$ ) or not ( $x_i = 0$ ). In total, 23 students had iPhones. We wish to infer about the population of all QC undergraduate students.

- (a) [4 pt / 59 pts] Is this sample an SRS from the population of interest? Why or why not?

You can say yes but you must argue that students at CUNY take classes at uniformly random times in uniformly random locations. Likely the answer is no because if you are on campus on Tuesday afternoon you are not a weekend student (those of which may be different socioeconomically). Also, Kiely Hall has more STEM classes than other buildings so those students may not be representative of all undergraduates either.

Regardless of what you wrote in (a), consider this sample to be an SRS from the population of interest for the rest of this question.

- (b) [9 pt / 68 pts] The US mean countrywide proportion of iPhones is estimated to be 62%. Prove via a hypothesis test at  $\alpha = 5\%$  using the one-proportion Z-test that the QC population has higher iPhone proportion than the countrywide mean. Your answer must include a clear setup of the two hypotheses, a retainment region and a decision. Do not interpret the decision nor find Fisher's  $p_{val}$ . Those will be asked in subsequent questions.

$H_0 : \theta \leq .62$ ,  $H_a : \theta > .62$ . We can then build the retainment region for  $\alpha = 0.05$  with the relevant  $z_{1-\alpha} = 1.65$  as it's one sided. So

$$RET = \left( -\infty, \theta_0 + z_{1-\alpha} \sqrt{\frac{\theta_0(1-\theta_0)}{n}} \right] = \left( -\infty, .62 + 1.65 \sqrt{\frac{.62(1-.62)}{27}} \right] = (-\infty, .774]$$

Now we calculate  $\bar{x} = 23/27 = .852 \notin RET$  hence  $H_0$  is rejected.

- (c) [4 pt / 72 pts] Interpret your decision from the previous question.

There is statistically significant evidence to conclude that the QC population has a higher iPhone proportion than the countrywide mean.

- (d) [2 pt / 74 pts] If your decision was an error, what type of error was it?

Type I

- (e) [6 pt / 80 pts] Represent Fisher's  $p_{val}$  by using the  $\Phi$  notation (i.e. the CDF of the standard normal).

$$\begin{aligned} p_{val} &= \mathbb{P}(\hat{\theta} > \hat{\theta} \mid H_0) \\ &= \mathbb{P}\left(\hat{\theta} > .852 \mid \hat{\theta} \sim \mathcal{N}\left(.62, \frac{.62(1-0.62)}{27}\right)\right) \\ &= \mathbb{P}\left(\frac{\hat{\theta} - .62}{\sqrt{\frac{.62(1-0.62)}{27}}} > \frac{.852 - .62}{\sqrt{\frac{.62(1-0.62)}{27}}}\right) \\ &= \mathbb{P}(Z > 2.48) \\ &= 1 - \Phi(2.48) \end{aligned}$$

- (f) [3 pt / 83 pts] Why is the  $p_{val}$  from the previous question approximate?

We used the standard normal as the result of the CLT which is approximate for any finite sample size.



- (g) [1 pt / 84 pts] If you wanted an exact test, what is the name of the test would you use?

Binomial Exact Test

- (h) [3 pt / 87 pts] In order to calculate the power of this test, what information would you need?

The true distribution of  $\hat{\theta}$  (which is equivalent to knowing  $\theta$  in the iid Bernoulli DGP).

- (i) [3 pt / 90 pts] If you ran the test with  $\alpha = 6\%$  instead of the original 5%, would power increase or decrease?

increase

- (j) [5 pt / 95 pts] If the null hypothesis were true, as the sample size approaches infinity, what would the decision of the test be?

You would still fail to reject. You are only guaranteed to reject if the null is false (even by a hairbreadth  $\delta$ ).

- (k) [5 pt / 100 pts] In order to calculate a 95% CI for  $\theta$ , what specific information do you need? Note: this is a difficult question.

When inverting the test, you'll see that you need the actual value of  $\theta$  as the standard error of  $\hat{\theta}$  is  $\sqrt{\theta(1-\theta)/n}$ . (This is impossible as  $\theta$  is unknown and hence why we did not yet cover the CI for the proportion).