ECON 0150 | MiniExam 4 | Fall 2025

This MiniExam will take 8 minutes with a quick break to follow. MiniExams are designed to both test your knowledge and challenge you to apply familiar concepts in new environments. Treat it as if you're trying to show me that you understand the material. Answer clearly, completely, and concisely.

Academic Conduct Code

The following academic conduct code is designed to protect the integrity of your work. Print your name/initials beside the three academic honesty agreements. I pledge to my fellow students, the university, and the instructor, that:

_____ I will complete this MiniExam solely using my own work.
_____ I will not use any digital resources unless explicitly allowed by the instructor.
_____ I will not communicate directly or indirectly with others during the MiniExam.

Q1. A healthcare economist studies whether a telemedicine program reduces emergency room visits. In a sample of 120 patients, those with telemedicine access had 1.8 fewer ER visits per year than those without access. The standard error is 0.65 visits.

a) If the null hypothesis is true (no effect of telemedicine), the sampling distribution of the difference in means would be:

□ A normal distribution with mean = -1.8 and SE = 0.65
□ A normal distribution with mean = 0 and SE = 0.65
□ A t-distribution with mean = -1.8 and SE = 0.65
□ A t-distribution with mean = 0 and SE = 0.65

b) Which statement correctly interprets a 95% confidence interval of (-3.08, -0.52) for this study?

□ We are 95% certain that the true effect of telemedicine is between -3.08 and -0.52 visits
□ 95% of patients in the study had a reduction between 0.52 and 3.08 visits
□ If we repeated this study many times, about 95% of the resulting confidence intervals would contain the true effect
□ The p-value for this result is exactly 0.05

Q2. A population has a true mean of μ = 75 and standard deviation σ = 20. Researchers take samples of size n = 64.

a) Which statement about the standard error of the sampling distribution is true?

□ If the sample size doubles, the standard error is cut in half
□ If the sample size doubles, the standard error doubles
□ If the sample size is quadrupled, the standard error is cut in half
□ If the sample size is quadrupled, the standard error doubles

b) If we reject the null hypothesis that $\mu = 70$ with $\alpha = 0.05$, we can conclude that:

□ The probability that $\mu = 70$ is less than 5%
□ If the true mean were 70, we'd observe our sample mean or more extreme less than 5% of the time
□ There is a 95% chance that $\mu = 75$
□ We've proven that $\mu = 75$

Q3. A social policy researcher studied the relationship between public library funding (dollars per capita) and high school graduation rates (%), with the following regression results:

	coef	std err	t	P>|t|	[0.025	0.975]
intercept	74.892	3.247	23.064	0.000	68.459	81.325
library_funding	0.897	0.234	3.833	0.003	0.434	1.360

a) What is the best interpretation of the coefficient on library_funding?

□ For each additional dollar of library funding per capita, the graduation rate is predicted to be 0.897% higher
□ Increasing library funding by $1 causes the graduation rate to increase by 0.897%
□ Cities that increase library funding see graduation rates increase 0.897 times
□ Graduation rates and library funding have a correlation of 0.897

b) The t-statistic of 3.833 for library_funding indicates:

□ Library funding is 3.833 times more important than other factors affecting graduation rates
□ The coefficient is 3.833 standard errors away from zero
□ We can be 3.833 times more confident in this result than a typical finding
□ There is a 3.833% chance that the relationship occurred by random chance

c) Based on the regression results, what is the predicted graduation rate for a city with library funding of $0 per capita?

□ 74.892%
□ 75.789%
□ 101.802%
□ Cannot be determined from the information provided

1. Below are two residual plots from separate regression analyses:

a) Which regression assumption is violated in Plot A?

□ Linearity
□ Normality
□ Homoskedasticity
□ Independence

b) The most appropriate approach to address the issue in Plot A would be:

□ Transform the dependent variable (e.g., take the log)
□ Add a squared term of the independent variable to the model
□ Use robust standard errors
□ Remove outliers from the dataset

2. An economist studies the relationship between advertising spending, product quality ratings, and monthly sales. Here are the regression results:

```
                        OLS Regression Results
===========================================================================
                  coef    std err         t     P>|t|      [0.025      0.975]
---------------------------------------------------------------------------
Intercept       23.846      2.514     9.485     0.000      18.819      28.873
ad_spending      0.385      0.067     5.746     0.000       0.251       0.519
quality_rating 5.927        0.894     6.630     0.000       4.139       7.715
===========================================================================
```

a) Write the estimated regression equation for this model.

□ sales = 23.846 + 0.385 × ad_spending + 5.927 × quality_rating
□ sales = 23.846 + 5.927 × ad_spending + 0.385 × quality_rating
□ ad_spending = 23.846 + 0.385 × sales + 5.927 × quality_rating
□ sales = 23.846 × (0.385 × ad_spending) × (5.927 × quality_rating)

b) If we had instead run a simple regression of sales on quality_rating only (without including ad_spending), the coefficient on quality_rating would likely be:

□ The same as in the multiple regression (5.927)
□ Lower than in the multiple regression (<5.927)
□ Higher than in the multiple regression (>5.927)
□ Cannot be determined from the information provided

3. A health economist wants to study whether patient recovery time differs between hospitals with different levels of nurse-to-patient ratios. They have data on recovery time and nurse-to-patient ratios for 85 hospitals.

a) What regression model would be most appropriate to answer this question?

□ RecoveryTime = β_0 + β_1 × NurseRatio + ε
□ RecoveryTime = β_0 + β_1 × I(NurseRatio > threshold) + ε
□ NurseRatio = β_0 + β_1 × RecoveryTime + ε
□ log(RecoveryTime) = β_0 + β_1 × log(NurseRatio) + ε

b) The researcher runs an appropriate model and finds a significant negative relationship between nurse-to-patient ratios and recovery time ($p = 0.003$). Which of the following is the most accurate interpretation of this result?

☐ Hospitals with higher nurse-to-patient ratios have significantly lower patient recovery times
☐ Increasing nurse-to-patient ratios causes shorter recovery times
☐ The data shows a negative correlation between nurse staffing levels and how long patients stay in the hospital
☐ Hospitals should hire more nurses to reduce patient recovery times