

ECN 102: Analysis of Economics Data

Midterm Exam FQ 2024

Remy Beauregard, Department of Economics, UC Davis

This exam consists of 5 short-answer questions (with sub-parts) and 5 multiple choice questions. You will have a maximum of 50 minutes to complete this exam. No additional time may be taken without prior accommodation. Make sure to read every question thoroughly and carefully. Make sure to answer all parts of all questions completely and fully explain when asked. Show all work to receive full credit.

You may use only the calculators provided by the instructor. Graphing calculators, phone calculators, and other non-standard calculators may not be used. For questions requiring computation, it suffices to express your final answer with 2 significant figures.

There is a formula sheet and scratch paper provided at the end of this exam. You may remove these pages and discard them after the exam. If you plan to use any of these extra pages for your final answers, you must write your name and student ID at the top of each scratch page to ensure it is not lost.

This exam is worth 70 points.

Name:

Student ID:

Section:

Question 1: Sample Means [10 points]

Suppose we know $Q \sim (16, 16)$ and we draw 50 samples of size 121 and obtain 50 sample means.

- a) How would we properly denote/write the sample mean of Q ? (You do not need to explain) [2 points]

For the distribution of **sample means** of Q :

- b) What would you expect the population mean to be equal to? [2 points]
- c) What would you expect the population standard deviation to be equal to? [3 points]
- d) What range would you expect 68% of the sample means of Q to fall under? Explain. [3 points]

Question 2: Summation Notation [6 points]

Please calculate the following sums:

a) [3 points]

$$\frac{1}{3} \sum_{i=3}^5 (i + 2)$$

b) [3 points]

$$\sum_{i=1}^3 (2^i)$$

Question 3: Summary Statistics [10 points]

Suppose you obtain the following summary for a sample of data with `summarize x, detail:`

x				

	Percentiles	Smallest		
1%	1	1		
5%	1	1		
10%	1	3	Obs	11
25%	3	3	Sum of wgt.	11
50%	4		Mean	5
		Largest	Std. dev.	3.162278
75%	8	7		

90%	9	8	Variance	10
95%	10	9	Skewness	.238797
99%	10	10	Kurtosis	1.694

a) Is the median of this data higher, lower, or equal to the mean? What does this mean for our data? [2 points]

b) Interpret the skewness and kurtosis of this data. What do each mean in words? [4 points]

c) Is this data distributed normally? Explain. [4 points]

Question 4: Univariate Inference [11 points]

You are given the following from Stata:

```

invttail(64,0.05)=1.669013
invttail(64,0.10)=1.2949198
invttail(63,0.05)=1.6694022

```

```

invttail(63,0.10)=1.2951343
invttail(62,0.05)=1.6698042
invttail(62,0.10)=1.2953558

```

Suppose you then use the `summarize` command to obtain the following results for the three variables below. The variables *price* and *weight* have been transformed to represent units of **thousands**.

(1978 automobile data)

Variable	Obs	Mean	Std. dev.	Min	Max
price	64	6.266484	3.065138	3.291	15.906
weight	64	2.980156	.8023428	1.76	4.84
mpg	64	21.57813	6.054789	12	41

a) Which of the three variables has the greatest dispersion? Which has the least? [3 points]

b) Provide a 90% CI for the mean of *mpg*. Interpret your answer. [3 points]

c) The claim is made that the average weight of a car in 1978 was above 2.8 thousand lbs. Using your sample data, test this claim at the 5% significance level. Clearly state your null and alternate hypotheses, your test statistic, and your conclusion. [5 points]

Question 5: Bivariate Data [23 points]

a) Briefly, what are the four key assumptions about our population bivariate regression model? What do they each mean? [4 points]

b) Assuming the four assumptions above hold and our sample is large enough, complete the statement about the distribution of b_2 : $b_2 \sim \text{---}(\text{---}, \text{---})$. Please use Greek letters when appropriate. [3 points]

Suppose we run the following regression and obtain these results:

Linear regression	Number of obs	=	64
	F(1, 62)	=	93.84
	Prob > F	=	0.0000
	R-squared	=	0.6350
	Root MSE	=	3.6874

		Robust				
mpg		Coefficient	std. err.	t	P> t	[95% conf. interval]

weight		-6.013482	.6207753	-9.69	0.000	-7.254394 -4.77257
_cons		39.49924	2.099907	18.81	0.000	35.30159 43.6969

c) Write down the estimated linear relationship between *weight* and *mpg*. [3 points]

- d) What does **robust** mean in this regression command? Why might we have added this? [3 points]

Interpret the t-statistic and p-value for the OLS slope coefficient from the regression output:

- e) What test is this referring to? What are its null and alternate hypotheses? [3 points]

- f) What is the conclusion of this test? [2 points]

- g) How many degrees of freedom will this test have? Why? [2 points]

- h) What would we expect the mpg to be for a car with a weight of 3.5 thousand lbs? [3 points]

Multiple Choice [2 points each]

Only one answer is correct for each question. Choose the best possible answer.

MC 1

A local polling firm plans to conduct phone surveys on issues of concern for voters. They plan to call 45 new households in the county each week and ask respondents how much money they spent on groceries the previous week. This will be repeated for 6 weeks.

This data will be:

- a) Experimental, categorical, panel
- b) Experimental, numerical, time-series
- c) Observational, numerical, repeated cross-section
- d) Observational, categorical, cross-section
- e) None of the above

MC 2

The OLS estimator:

- a) minimizes the sum of horizontal deviations of actual data points from the sample regression line
- b) minimizes the sum of horizontal deviations of actual data points from the population regression line
- c) minimizes the sum of vertical deviations of actual data points from the sample regression line
- d) minimizes the sum of vertical deviations of actual data points from the population regression line

MC 3

For a r.v. X with population mean μ and population variance σ^2 , the *standard error* is:

- a) An estimator for the standard deviation of μ
- b) An estimator for the standard deviation of $\bar{X}$
- c) An estimator for the standard deviation of X
- d) Always equal to the sample standard deviation of X
- e) None of the above

MC 4

Suppose a r.v. $Z \sim (3, 25)$. We would expect the transformed r.v. $\tilde{Z} = (Z - 3)/5$ to be:

- a) A z-score
- b) Distributed identically to Z
- c) Normally distributed as $n \rightarrow \infty$
- d) All of the above
- e) None of the above

MC 5

Suppose we regress y on x and obtain the result $\hat{y} = 0.33x + 1$ with $R^2 = 0.64$. If we then regress x on y , we would expect:

- a) A slope coefficient of 0.33
- b) A slope coefficient of $0.33^{-1} \approx 3$
- c) $r_{yx} = 0.8$
- d) $R^2 \neq 0.64$
- e) None of the above