

**Instructor:** Prof. Murdock

**Duration:** 120 minutes. You must stay in the test room the entire time.

**Part 1 (60 points total):** 20 multiple-choice questions with answers recorded on SCANTRON form.

**Part 2 (40 points total):** Written questions with answers recorded on test papers.

**Allowed aids:** A non-programmable calculator (and attached aid sheets, which you may detach)

**PART 1 INSTRUCTIONS:**

Do NOT write your answers on these test papers; You MAY do scratch work on these pages  
ONLY those answers correctly marked on the SCANTRON form can earn positive marks

Correct answers are worth: 3.00 points

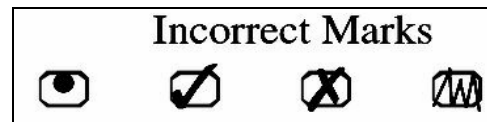
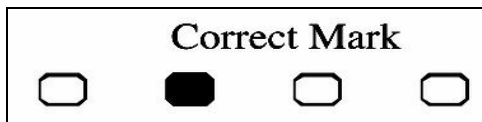
Incorrect answers are worth: -0.60 points

Alternative (E) "Don't know" are worth: 0.75 points

- Use only a pencil or blue or black ball point pen



- **Pencil strongly recommended**, it can be erased if a mistake is made
- Make dark solid marks that fill the bubble completely



- Select the one best alternative
- Erase completely any marks you want to change
  - Crossing out a marked box is not acceptable and is incorrect

**1<sup>st</sup>:** Print your **LAST NAME** and **INITIALS** in boxes provided

- Use exact name you are officially registered under
- Darken each letter in the corresponding bracket below each box

**2<sup>nd</sup>:** Print your 9 digit **STUDENT NUMBER** in the boxes provided

- Fill in zeros in front of the number if less than 9 digits
- Darken each number in the corresponding bracket below each box

**3<sup>rd</sup>:** Print 2 digit **FORM** number in the boxes provided

- Your FORM number is **01**
- Darken each number in the corresponding bracket below each box

**4<sup>th</sup>:** Sign your name in the **SIGNATURE** box

**For the 20 questions, choose the best answer and mark it on the SCANTRON form. Choice (E) "Don't know" is never the correct answer, but you obtain partial credit for choosing it.**

**(1)** Which of the following is a continuous probability distribution?

- (A)** Bernoulli
- (B)** Binomial
- (C)** Poisson
- (D)** Uniform
- (E)** Don't know

**(2)** In measuring the strength of a positive linear relationship between X and Y, which of the following affect the magnitude of the coefficient of correlation?

I.	How scattered the points are about the least squares line
II.	How steep the least squares line is
III.	How X and Y are measured: their units of measurement

- (A)** I.
- (B)** II.
- (C)** I. and II.
- (D)** I., II. and III.
- (E)** Don't know

**(3)** Which kind of data often causes an endogeneity bias?

- (A)** Experimental
- (B)** Natural experiment
- (C)** Observational
- (D)** Cross sectional
- (E)** Don't know

**(4)** For a sample taken from negatively skewed population, which of the following statements about the sample mean would be expected?

- (A)** The sample mean is larger than the sample median
- (B)** The sample mean is larger than the sample 45<sup>th</sup> percentile
- (C)** The sample mean is smaller than the sample 10<sup>th</sup> percentile
- (D)** The sample mean is smaller than the sample 55<sup>th</sup> percentile
- (E)** Don't know

(5) For a sample of size 100 taken from a Poisson population with  $\lambda = 1$ , what is the distribution of the sample mean?

- (A) Poisson with  $\lambda = 1$
- (B) Poisson with  $\lambda = 0.01$
- (C) Normal with  $\mu = 1$  and  $\sigma^2 = 1$
- (D) Normal with  $\mu = 1$  and  $\sigma^2 = 0.01$
- (E) Don't know

(6) Considering a bell shaped population with a mean of \$41,201 and a standard deviation of \$10,779, what is the population 51<sup>st</sup> percentile?

- (A) \$41,470
- (B) \$43,950
- (C) \$66,316
- (D) \$68,957
- (E) Don't know

(7) Which of the following describe observational data?

I.	To study the effect that product placement on grocery store shelves has on customer purchasing behavior, a grocery store rotates product placements and records changes in customer behavior.
II.	To study the link between herbal supplements and heart disease, a researcher selects a random sample of adults and asks about their intake of herbal supplements and measures their heart health.

- (A) Only I.
- (B) Only II.
- (C) Both I. and II.
- (D) Neither I. nor II.
- (E) Don't know

(8) X measures the negotiated selling price in dollars for a new Toyota Camry in Toronto. The population distribution of X is normal with a mean of \$23,000. What is the probability that for a randomly selected deal the buyer pays a price more than one standard deviation above average?

- (A) 0.1587
- (B) 0.1915
- (C) 0.2301
- (D) 0.3413
- (E) Don't know

► For Questions (9) – (10): A sample of size 16 is taken from a normal population. Assuming the population variance is known, the 92% interval estimator of the population mean is (26.16, 34.04).

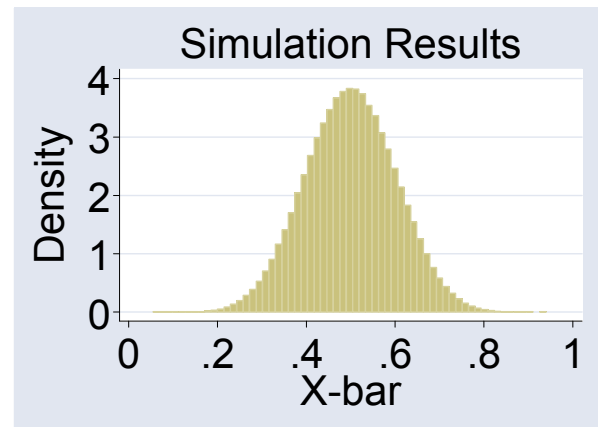
(9) What is the point estimate of the population mean?

- (A) 30.0
- (B) 30.1
- (C) 30.2
- (D) 30.3
- (E) Don't know

(10) What is the population standard deviation?

- (A) 7
- (B) 8
- (C) 9
- (D) 10
- (E) Don't know

► For Questions (11) – (13): A population is Uniformly distributed with  $a = 0$  and  $b = 1$ . For 1,000,000 simulation draws, a Monte Carlo simulation calculates the sample mean for a random sample of size 8. The following results are obtained:



-----				
	Percentiles	Smallest		
1%	.2665249	.0694433		
5%	.331973	.0834262		
10%	.3681694	.0920229	Obs	1000000
25%	.4300038	.0974796	Sum of Wgt.	1000000
50%	.5000456		Mean	.5000387
		Largest	Std. Dev.	<input type="text"/>
75%	.5698275	.898767		
90%	.6320018	.9000767	Variance	
95%	.6680485	.9079162	Skewness	
99%	.7339593	.9251378	Kurtosis	

(11) For the empty box next to “Std. Dev.” what is the expected value of what belongs in that box?

- (A) 0.100
- (B) 0.101
- (C) 0.102
- (D) 0.103
- (E) Don't know

(12) Which of the following would be expected to decrease the value of “Std. Dev.”?

I.	Increasing the number of simulation draws
II.	Increasing the sample size

- (A) I.
- (B) II.
- (C) Both I. and II.
- (D) Neither I. nor II.
- (E) Don't know

(13) These Monte Carlo simulation results support which conclusion?

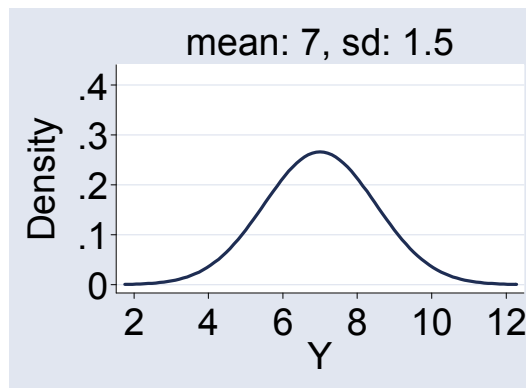
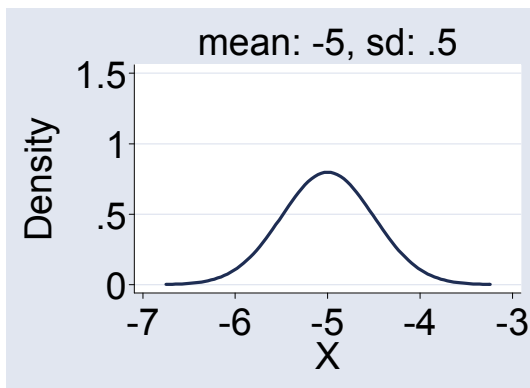
- (A) In this case a sample size of 8 is “sufficiently large” for the purposes of applying the CLT
- (B) The sample mean is an upwardly biased measure of the population median
- (C) The sample mean is a downwardly biased measure of the population median
- (D) The sampling distribution of the sample mean is discrete
- (E) Don't know

(14) In words, the meaning of  $E[\bar{X}_1 - \bar{X}_2] = \mu_1 - \mu_2$  is that the difference between the sample means is \_\_\_\_\_ the difference between the population means.

I.	equal to
II.	an unbiased estimator of
III.	a consistent estimator of

- (A) I.  
 (B) II.  
 (C) II. and III.  
 (D) I., II. and III.  
 (E) Don't know

► For Question (15): Consider the following graphs.



(15) Which linear transformation of X yields Y?

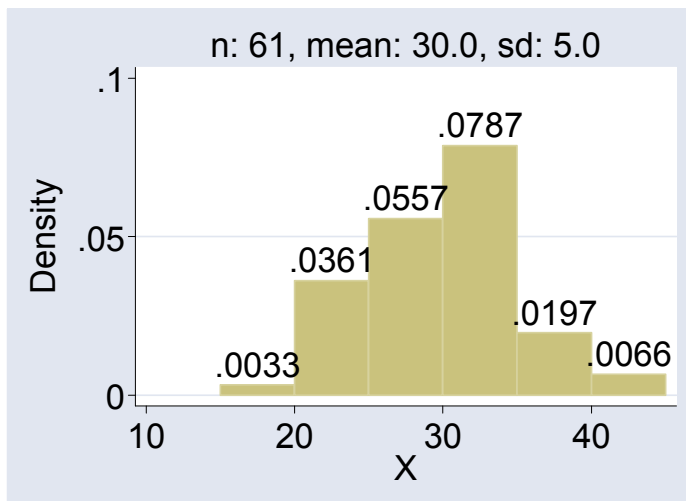
- (A)  $Y = 2 + (-1) \cdot X$   
 (B)  $Y = 12 + (1) \cdot X$   
 (C)  $Y = 12 + (3) \cdot X$   
 (D)  $Y = 22 + (3) \cdot X$   
 (E) Don't know

**(16)** The width of the confidence interval estimator of the population mean will increase with an increase in which of the following?

I.	Sampling error
II.	Non-sampling error

- (A) I.  
 (B) II.  
 (C) Both I. and II.  
 (D) Neither I. nor II.  
 (E) Don't know

► For Questions **(17) – (18)**: Consider the following histogram of a random sample.



**(17)** What percent of the sample is within one standard deviation of the mean?

- (A) 67.2  
 (B) 67.7  
 (C) 68.3  
 (D) 68.6  
 (E) Don't know

**(18)** Which is the approximate 95% confidence interval estimator of the population mean?

- (A) (20, 40)  
 (B) (21, 39)  
 (C) (27, 33)  
 (D) (29, 31)  
 (E) Don't know

► For Questions **(19)** – **(20)**: Consider the following STATA summary of a random sample taken from a normal population.

Variable	Obs	Mean	Std. Dev.	Min	Max
x	6	4.702875	.7994208	3.558295	5.968183

**(19)** What is the 95% confidence interval estimator of the population mean?

- (A)** (3.864, 5.542)
- (B)** (3.988, 5.418)
- (C)** (4.045, 5.360)
- (D)** (4.063, 5.343)
- (E)** Don't know

**(20)** Supposing that the sample standard deviation did not change, doubling the sample size would do what to the overall width of the 95% confidence interval estimator?

- (A)** Reduce its width by 19%
- (B)** Reduce its width by 29%
- (C)** Reduce its width by 39%
- (D)** Reduce its width by 49%
- (E)** Don't know



**Instructor:** Prof. Murdock

**Duration:** 120 minutes. You must stay in the test room the entire time.

**Part 1 (60 points total):** 20 multiple-choice questions with answers recorded on SCANTRON form.

**Part 2 (40 points total):** Four written questions with answers recorded on test papers.

**Allowed aids:** A non-programmable calculator (and the aid sheets attached to Part 1)

**PART 2 INSTRUCTIONS:**

Answer four written questions with point values next to each for a total of 40 points. If you run out of room you may continue your answers on pages 7 and 8, but clearly indicate you have done so (for example: "See page 7 for the rest of my answer...") and clearly label your additional responses (for example: "Question (4) (b) continued:").

**Last Name:**

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

**First Name:**

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

**Student ID #:**

--	--	--	--	--	--	--	--	--	--

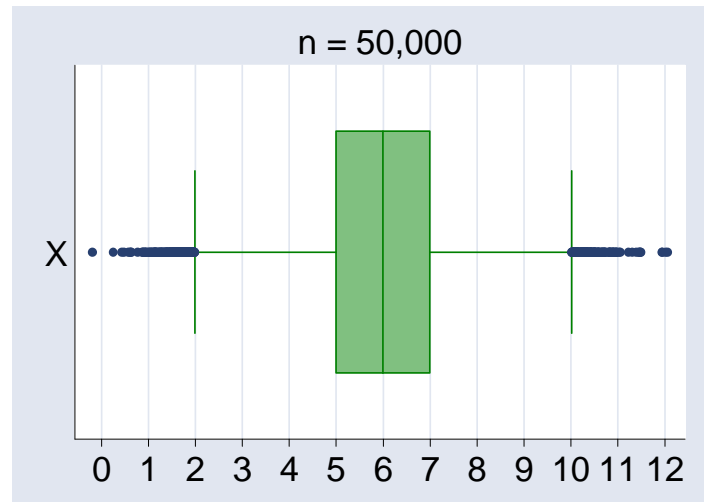
	Q1	Q2	Q3	Q4	Part 2	Part 1	Total
<i>Point Value</i>	8	6	8	18	40	60	100
Points Earned							

**Problems:** Write answers clearly, concisely, and completely on these test papers. Show your work.

**(1) [8 points]** Consider this tabulation. Is it reasonable to infer that this sample is from a Poisson population? If not, show why not. If so, show why and give reasonable values for the Poisson parameter(s).

X	Freq.	Percent	Cum.
0	67	23.59	23.59
1	91	32.04	55.63
2	64	22.54	78.17
3	41	14.44	92.61
4	13	4.58	97.18
5	7	2.46	99.65
8	1	0.35	100.00
Total	284	100.00	

**(2) [6 points]** Consider this box plot of a random sample drawn from a normal population. Compute the best approximate value of the sample standard deviation (to the nearest hundredth).



**(3) [8 points]** Worker compensation is positively skewed and  $\sigma^2 = 100$ . The Public Relations person for the corporation claims that the mean worker compensation is 20. The employee union randomly samples 35 employees and obtains a sample mean of only 18. You are hired by the union to do a statistical analysis that will help it determine if the corporation is lying. Do that analysis and write an explanation of the results for the union boss who does not have any training in statistics, but wants to understand exactly what the results mean.

**(4) [18 points]** A population is Bernoulli distributed and the probability that  $X$  is equal to one is 0.35 and the probability that  $X$  is equal to zero is 0.65.

**(a) [9 points]** For a sample size of 3, find the sampling distribution of the mean. Graph it. Use it to find the probability that the sample mean is greater than 0.25.

**(b) [9 points]** For a sample size of 300, find the sampling distribution of the mean. Graph it. Use it to find the probability that the sample mean is greater than 0.25.

**Extra Space:** If you use this space, clearly indicate for which question(s).

**Extra Space:** If you use this space, clearly indicate for which question(s).