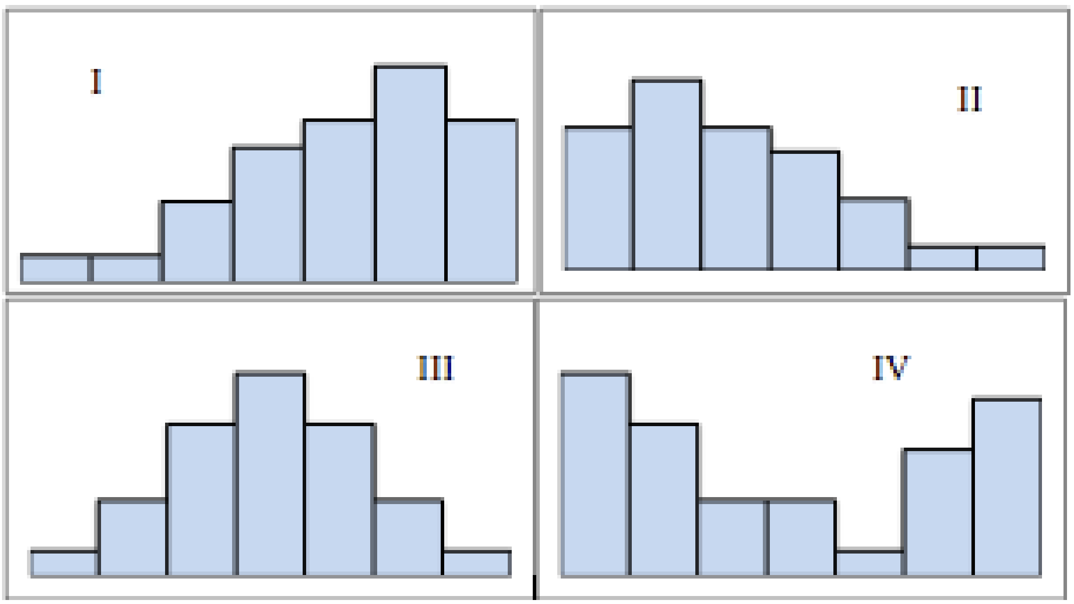
Name: __________________________ **Section: <u>MAT098/181C-</u>**

MAT098/181C FINAL EXAM (FORM B Key)

*A scientific or graphing calculator is permitted. **<u>Cellphones may not be used as calculators and must be off or on vibrate during the exam</u>**. Show all work on the test*

1. For each description, choose the histogram (I, II, III, IV) that matches the description. (8 pts)



a. The distribution of quiz scores on an easy quiz.
   I

b. The distribution of annual income for NBA basketball players where only a few are high-paid superstars.
   II

c. The distribution of hours that students studied for an exam. Many students studied a lot. A similar number of students did not study very much.
   IV

d. The distribution of length measurements at birth for 10,000 babies.
   III

2. A program selected a group of outstanding participants and recorded their gender and majors. A student was randomly selected from the program. (10 pts)

	Engineering	**Business**	**Biology**	**English**	**Total**
Male	42	22	14	22	100
Female	45	21	13	21	100
Total	87	43	27	43	200

a. What is the probability that the student is a biology major?

$$P(BIO) = \frac{27}{200}$$

b. Given that the student is a business major, what is the probability that the student is a female?

$$P(F|BUS) = \frac{21}{43}$$

c. Given that the student is a female, what is the probability that her major is engineering?

$$P(EGG|F) = \frac{45}{100}$$

3. Jan earns a salary of \$86,000 as a saleswoman, while her sister Amy earns \$82,000 as a YouTuber making traveling videos. Assume that sales have a mean salary of \$80,000 with standard deviation \$4,000. Assume that YouTuber have a mean salary of \$70,000 with standard deviation \$6,000. (10 pts)
   a. Convert each salary to a standard z score.

$$z = \frac{86 - 80}{4} = 1.5$$

$$z = \frac{82 - 70}{6} = 2$$

   b. Who does better compare to other workers doing the same job?

Amy

4. In a study, the average time for students (part-time and full-time) to get their bachelor's degree was $\mu = 6$ years and standard deviation $\sigma = 1.2$ years. This distribution is a normal distribution. A student is selected at random. Find the probability that the time for that student to get his or her bachelor's degree is between 4 years and 5 years. (10 pts)

$$P(4 < x < 5) = P(-1.67 < z < -0.83) = 0.2033 - 0.0475 = 0.1558$$

5. To keep up with the competition, employee in Google tends to be on a rigid work schedule. The employee work on an average of 50 hours per week with a standard deviation of 5.25 hours. A newly hired engineer hopes that it's shorter. She asks 9 engineering colleagues about their mean work week. What is the probability that the ***sample mean*** $\bar{x}$ of mean work week is less than 45 hours? (10 pts)

$$P(x < 45) = P\left(z < \frac{45 - 50}{5.25/\sqrt{9}}\right) = P(z < -2.86) = 0.0021$$

6. To ensure efficient usage of a server, it is necessary to estimate the mean number of concurrent users. According to records, the number of concurrent users at 49 randomly selected times has a sample mean $\bar{x} = 37.7$ and sample standard deviation $s = 9.2$. Construct a 95% confidence interval for the mean number of concurrent users. Please round to 2 decimal places. (12 pts)

$$E = t_c \cdot \frac{s}{\sqrt{n}} = 2.01 \cdot \frac{9.2}{\sqrt{49}} = 2.64$$

$$\bar{x} \pm E = (35.06, 40.34)$$

7. A recent study showed that in Cambridge, the average single-family homeowner lived at one address 8.0 years before moving. In Middlesex County, a random sample of 56 single-family homeowners lived at one address an average of 7.3 years. Similar studies of home ownership use $\sigma = 2.8$ yrs. Test the claim that single-family homeowners in Middlesex County live in one place shorter than the state average at a 5% significance level. (20 pts)

   a. What is the level of significance? State the null and alternate hypotheses. Will you use a left-tailed, right-tailed or two-tailed test?

      $H_0: \mu = 8$
      $H_a: \mu < 8$ (left tailed)
      $\alpha = 0.05$

   b. Check the conditions. Identify the sampling distribution you will use: the standard normal or the Student's t. Explain the rationale for your choice.

      $$n > 30, \sigma\ known$$

   c. What is the value of the sample test statistic? Draw a picture and find (or estimate) the P-value.

      $$z = \frac{x - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{7.3 - 8}{\frac{2.8}{\sqrt{56}}} \approx -1.87$$

      P-value = 0.0307

   d. Based on your answers for parts (a)-(c), will you reject or fail to reject the null hypothesis?

      Reject null hypothesis.

   e. Interpret your decision in the context of the application.

      At 5% significant level, there is enough evidence to support the alternative hypothesis that that single-family homeowners in Middlesex County live in one place shorter than 8 years.

8. A marketing analyst is studying the relationship between x = amount spent on television advertising and y = increase in sales. The data are reported in thousands of dollars. The following data represent a random sample from the study. Use this information for parts a) to c). Round the answers to 2 decimal places. (20 pts)

x (advertising)	15	28	19	47	10	92
y (sales increase)	340	260	152	413	130	855

a. Use your calculator to find the correlation coefficient r. What does r tell you about the association between x and y, as far as the ***direction*** and ***strength***?

$$r = 0.954$$

Positive, strong.

b. Use your calculator to find
$\bar{x} =$ 35.17 $\bar{y} =$ 358.33 $s_x =$ 30.73 $s_y =$ 266.23

c. Find the equation for the regression line, $y = a + bx$ using the following:

$$b = r\frac{s_y}{s_x} \quad \text{and} \quad a = \bar{y} - b\bar{x}$$

$$b = 0.954 \cdot \frac{266.23}{30.73} = 8.26$$

$$a = 358.33 - 8.26 * 35.17 = 67.83$$

$$y = 67.83 + 8.26x$$

d. A lurking variable is a variable that is not measured in the study. It is a third variable that is neither the explanatory nor the response variable, but it affects your interpretation of the relationship between the explanatory and response variable. Identify one lurking variable in this study.

The quality of the product. The strength of the economy. Weather it is a holiday sale season. etc…

<u>*EXTRA CREDIT</u>: (3 pts) Based upon past experience, 8% of all customers at Miller's pizza restaurant pay for their purchases with a visa credit card. If Miller makes 219 sales last month, what is the probability that less than 20 customers paid with a visa credit card?

$$\mu = 219 * 0.08 = 17.52$$

$$\sigma = \sqrt{219 * 0.08 * 0.92} = 4.01$$

$$P(x_B < 20) = P(x_N < 19.5) = P(z < 0.49) = 0.6879$$

<u>*EXTRA CREDIT</u>: (3 pts) A Nissan Motor Corporation advertisement read, "The average man's I.Q. is 107. The average brown trout's I.Q. is 4. So why can't man catch brown trout?" Suppose you believe that the brown trout's mean I.Q. is greater than 4. You catch 12 brown trout. A fish psychologist determines the I.Q.s as follows:

5	4	7	3	6	4	5	3	6	3	8	5

Assume normal distribution. Conduct a hypothesis test of your belief. State:

1) Null, Alternate Hypothesis, type of test & level of significance
2) Check the conditions.
3) Compute the sample test statistic, draw a picture and find the P-value.
4) State the conclusion about the Null Hypothesis.
5) Interpret the conclusion.

$$n = 12, \qquad \bar{x} = 4.92, \qquad s = 1.62, \qquad \alpha = 0.05.$$

$$H_0: \mu = 4$$

$$H_a: \mu > 4$$

Data is approximately normal

$$t = \frac{4.92 - 4}{1.62/\sqrt{12}} = 1.967$$

P-value is 0.037 $< \alpha$. Reject H_0

There is enough evidence to support the belief that the brown trout's mean I.Q. is greater than 4.

z-score

$$z = \frac{x - \mu}{\sigma}$$

Central Limit Theorem

Mean of the sample mean is $\mu_{\bar{x}} = \mu$

Standard deviation of the sample mean is $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

z-score for sample mean

$$z = \frac{x - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

Binomial Distribution

Mean: $\mu = np$

Standard Deviation: $\sigma = \sqrt{np(1-p)}$

Sampling Distribution of Sample Proportion

Mean: p

Standard Deviation: $\sigma = \sqrt{\frac{p(1-p)}{n}}$

Confidence Interval for Population Parameters

Concept	Population Proportion p	Population Mean μ	
confidence interval formula	$\hat{p} \pm Z_c \cdot \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$	σ known $\bar{x} \pm Z_c \cdot \dfrac{\sigma}{\sqrt{n}}$	σ unknown $\text{df} = n - 1$ $\bar{x} \pm T_c \cdot \dfrac{s}{\sqrt{n}}$
sample size formula	$\hat{p} = \frac{x}{n}$ known $n = \hat{p} \cdot \hat{q} \cdot \left(\dfrac{Z_c}{E}\right)^2$ $\hat{p}$ unknown $n = \dfrac{1}{4} \cdot \left(\dfrac{Z_c}{E}\right)^2$	$n = \left(\dfrac{Z_c \cdot \sigma}{E}\right)^2$	

- 90% confidence interval: $Z_c \approx 1.645$
- 95% confidence interval: $Z_c \approx 1.960$
- 99% confidence interval: $Z_c \approx 2.576$

Hypothesis Testing

Concept	Population Proportion p	Population Mean μ	
test statistics	$z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{p(1-p)}{n}}}$	σ known $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$	σ unknown $\text{df} = n - 1$ $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$

- If the P-value $< \alpha$, we reject the null hypothesis.
- If the P-value $\geq \alpha$, we fail to reject the null hypothesis.