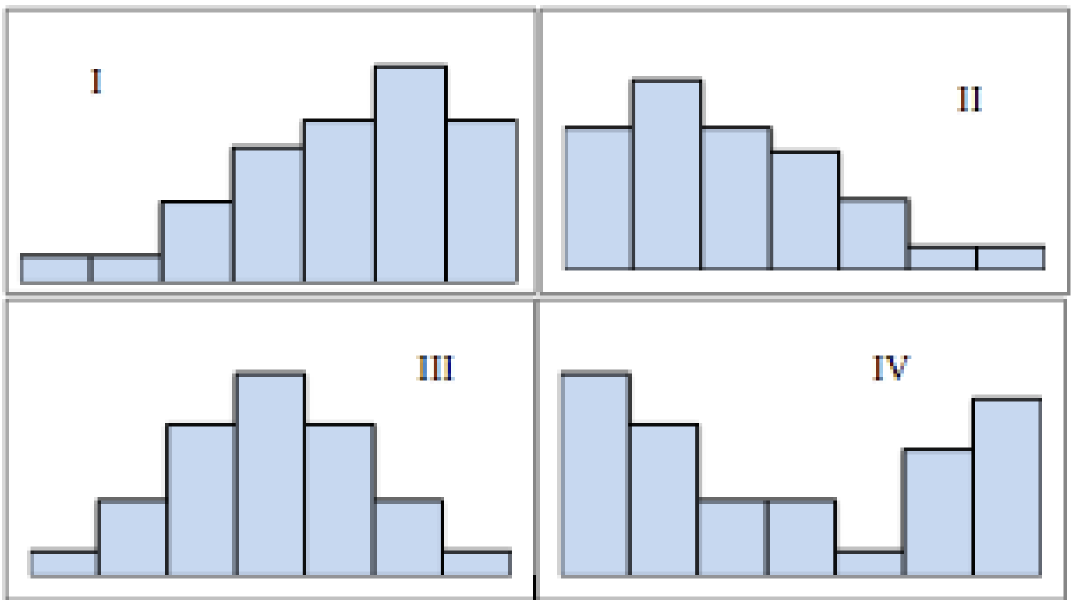
Name: ________________________ **Section: MAT098/181C-**

MAT098/181C FINAL EXAM (FORM A Key)

A scientific or graphing calculator is permitted. ***Cellphones may not be used as calculators and must be off or on vibrate during the exam****. 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 length measurements at birth for 10,000 babies.
   III

   b. The distribution of annual income for school employees where a high percentage of employees are entry-level teachers and only a few are high-paid administrators.
   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 quiz scores on an easy quiz.
   I

2. The manager of a used car lot took inventory of the automobiles on his lot and constructed the following table based on the age of his car and its make (foreign or domestic). A car was randomly selected from the lot. (10 pts)

	Age of Car (in years)				
Make	**0 – 2 yrs**	**3 – 5 yrs**	**4 – 6 yrs**	**6 – 10 yrs**	**Total**
Foreign	42	22	14	22	100
Domestic	45	21	13	21	100
Total	87	43	27	43	200

a. What is the probability that a car is between 3-5 yrs?

$$P(3-5) = \frac{43}{200}$$

b. Given that a car is between 6-10 yrs old, what is the probability that it is a domestic car?

$$P(d|6-10) = \frac{21}{43}$$

c. Given that the car selected was a foreign car, what is the probability that it was older than 2 years?

$$P(>2|f) = \frac{22+14+22}{100} = \frac{58}{100}$$

3. Jan earned 86 on her political science midterm and 82 on her chemistry midterm. In the political science class the mean score was 80 with standard deviation 4. In the chemistry class the mean score was 70 with standard deviation 6. (10 pts)
   a. Convert each midterm score to a standard z score.

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

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

   b. On which test did she do better compared to the rest of the class?

   chemistry midterm

4. The snow pack on the summit of Wolf Creek Pass, Colorado on March 1 has been measured for many years. It is normally distributed with mean $\mu = 78.1$ inches and standard deviation $\sigma = 10.4$ inches. A year is selected at random. Find the probability that the snowpack is between 60 inches and 85 inches. (10 pts)

$$P(60 < x < 85) = P\left(\frac{60 - 78.1}{10.4} < z < \frac{85 - 78.1}{10.4}\right) = P(-1.74 < z < 0.66)$$
$$= 0.7454 - 0.0409 = 0.7066$$

5. The Internal Revenue Service has studied millions of tax returns. They have determined that the mean deduction taken for charitable contributions by people with incomes between twenty-five and thirty thousand dollars per year is about \$800. The standard deviation is estimated to be about \$170. What is the probability that the ***sample mean*** $\bar{x}$ deduction for charitable contributions based on a random sample of 100 returns is more than \$830. (10 pts)

$$z = \frac{x - \mu}{\sigma/\sqrt{n}} = \frac{830 - 800}{170/\sqrt{100}} = 17.64$$

$$P(z > 17.64) = 0$$

6. As part of an environmental studies class project, students measured the circumferences of a random sample of 45 blue spruce trees near Brainard Lake, Colorado. The sample mean circumference was $\bar{x}$ = 29.8 inches. Assume that σ is known to be 7.2 inches. Find a 95% confidence interval for the population mean circumference of all blue spruce trees near this lake. Please round to 2 decimal places. (12 pts)

$$E = z_c \cdot \frac{\sigma}{\sqrt{n}} = 1.96 \cdot \frac{7.2}{\sqrt{45}} = 2.10$$

$$\bar{x} \pm E = (27.7, 31.9)$$

7. A recent study showed that in California, the average single-family homeowner lived at one address 8.0 years before moving. In Sonoma County, a random sample of 56 single-family homeowners lived at one address an average of 8.7 years. Similar studies of home ownership use $\sigma = 2.8$ yrs. Test the claim that single-family homeowners in Sonoma County live in one place longer than the state average at a 1% 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?

      $H_0: \mu = 8$
      $H_1: \mu > 8$ (right tailed)
      $\alpha = 0.01$

   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.

      $$z = \frac{x - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{8.7 - 8}{\frac{2.8}{\sqrt{56}}} \approx 1.87$$

      P-value = 0.0307

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

      Fail to reject null hypothesis.

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

      At 5% significant level, there is not 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). Please round 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 $\quad\bar{y} =$ 358.33 $\quad s_x =$ 30.73 $\quad s_y =$ 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) Peter is a door-to-door sales representative who makes a sale at only 8% of his house calls. If Peter makes 219 house calls today, what is the probability that the makes a sale at 20 or fewer houses?

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

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

$$P(x_B \leq 20) = P(x_N < 20.5) = P(z < 0.74) = 0.7704$$

<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 four. 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.

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

$$H_0: \mu = 4$$

$$H_a: \mu > 4$$

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.