

22S:008 Exam 2, Part I, April 9, 1999 Name: _____

Please circle one. Section: 11:30, 12:30, or 2:30

Please enter all of your answers on these exam pages. There are 19 questions. The Defective Question Report Formula sheet and Normal tables will be handed out separately.

1. (2 pts) In a designed experiment, a sham or fake treatment disguised as a real treatment it is called a **placebo**.
A) True B) False

2. (2 pts) An **ecological variable** is a variable that plausibly explains a relationship between two other variables.
A) True B) False

3. (3 pts) Suppose y is a response variable of annual salary, x is a continuous predictor variable of years of experience, and z is a binary indicator variable indicating gender ($z=1$ means female, $z=0$ means male). Consider the regression model with fitted equation: $\hat{y} = b_0 + b_1x + b_2z$. What coefficient or combination of coefficients represents the predicted annual salary for a male with one year of experience?
A) b_0
B) b_1
C) b_2
D) b_0+b_1
E) b_0+b_2

4. (3 pts) In a taste test of three pizza recipes, several different tasters evaluated each of the three recipes. This is an example of
A) replication
B) blinding
C) placebo
D) randomization
E) control

5. (2 pts.) In the preceding question there are _____ factor(s) each with _____ levels.

6. (3 pts) A professor asked his class to guess his age in years. Half the class were given the hint that his oldest child was 35 years old. The other half received no hint. The half that received no hint is called the
A) blind group
B) control group
C) factor group
D) treatment group
E) None of the above.

7. (3 pts) There is a strong positive correlation between the number of fire trucks sent to a fire and the amount of damage done at the fire. This is an example of:
- A) a lurking variable
 - B) ecological correlation
 - C) cause and effect
 - D) consistency
 - E) mechanism

8. (2 pts) A straight-line model was fit to several data points using least squares. *Some* of the results are shown in the table at the right. What is the fitted value when $x = 12$ (to the nearest tenth)?

x	y	fitted value	residual
7	510	504.9	5.1
?	503	505.2	?
...
12	504	?	-2.4
13	512	506.7	?

- A) -2.4
- B) 12
- C) 504
- D) 506.4
- E) Cannot be determined from the information given.

9. (2 pts) Use the information in question 8. What is the residual for the last case where $x = 13$? Show your work.

residual = _____

10. (3 pts) Use the information in question 8. What is the slope of the least squares regression line? Show your work.

slope = _____

11. Two regression models are developed to relate monthly Sales (in \$1000's) to monthly advertising expenditures (Ads in \$1000's). The first model is a straight-line model while the second uses Ads and Ads squared in a quadratic model. Here are some of the results:

Straight-Line Model			Quadratic Model		
Source	SS	DF	Source	SS	DF
Regression	858	1	Regression	889	2
Residual Error	1230	10	Residual Error	1199	9
Total	2088	11	Total	2088	11

a) (2 pts) Find the R -squared value for the quadratic model. Show your work.

$$R^2 = \underline{\hspace{2cm}}$$

b) (2 pts) Find the residual standard deviation of the straight-line model. Show your work.

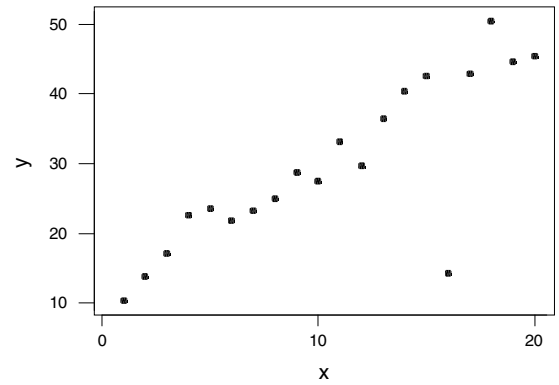
$$\text{residual standard deviation} = \underline{\hspace{2cm}}$$

12. (4 pts) An economist obtains a regression model that relates a person's annual Savings (in \$1000's), a person's annual Income (in \$1000's), and an indicator variable, House, where House = 1 if the person owns a house and House = 0 if not a homeowner. The fitted equation is $\text{Savings} = -0.32 + 0.0675 \text{ Income} + 0.827 \text{ House}$. How much more (or less) Savings does this model predict for a non-homeowner with annual income of \$30,000 compared to a homeowner with annual income of \$20,000. Show your work.

$$\text{answer} = \underline{\hspace{2cm}}$$

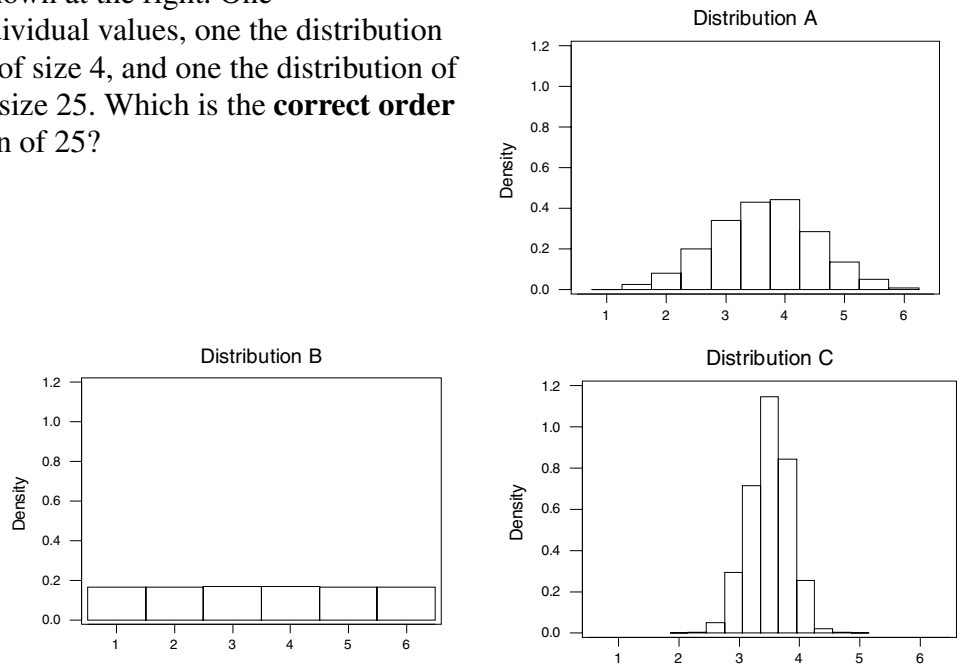
13. (2 pts) Consider the scatterplot shown at the right. If the y value for the outlier were changed to a value of $y = 40$ the correlation in the scatterplot would **increase**.

A) True B) False



14. (3 pts) Three distributions are shown at the right. One represents the distribution of individual values, one the distribution of means from random samples of size 4, and one the distribution of means from random samples of size 25. Which is the **correct order** for: individuals, mean of 4, mean of 25?

- A) A, B, C
 B) C, B, A
 C) C, A, B
 D) B, C, A
 E) B, A, C



15. (2 pts) Which **one** of the following statements is **false** regarding control charts?

- A) Control charts can help avoid tampering with a stable process.
 B) Control charts can help managers decide appropriate action to take concerning their processes.
 C) Control charts help managers detect special causes of variation in their processes.
 D) Control charts can be applied to service businesses as well as manufacturing.
 E) Control charts were invented by the Japanese within the last 25 years.

16. (3 pts) The times necessary to complete service for a class of bank customers is described by a normal distribution with mean 9 minutes and standard deviation 2.2 minutes. Service times are considered excessive if they exceed 15 minutes. Over the long run, what percent of customers will experience excessive service times?

Answer = _____

17. (4 pts) A random sample of size 100 is selected from 10,000 accounts. In the 10,000 accounts the mean account size is \$150 and the standard deviation is \$75. The shape of the distribution of account sizes is not known. What is the chance that the **mean** amount in the 100 accounts sampled is greater than \$165?

Answer = _____

18. (3 pts) In lecture we used a “managed” dice-roll process where we adjusted the current dice total based on the last dice total in order to roll more sevens. This example illustrates:
- A) improving an out-of-control process.
 - B) improving an in-control process.
 - C) tampering with an in-control process.
 - D) how to reduce the variation in an in-control process.
 - E) how to find special causes.
19. (3 pts) An analyst is using regression to model the Market value of houses using the size of the house, Sq. ft, and a categorical variable, Grade, that has values Grade=-1 for Low grade houses, Grade=0 for Medium grade houses, and Grade=+1 for High grade houses. The model is $\text{Market} = b_0 + b_1 \text{Sq. ft} + b_2 \text{Grade}$. Which of the following statements best describes this model?
- A) This model is a parallel-planes model.
 - B) This model is one plane.
 - C) This model is a parallel-lines model with arbitrary vertical distance between the lines.
 - D) This model is a parallel-lines model with equal vertical distance between the lines.
 - E) None of the above.

Defective Question Report

Name: _____

Section: _____

If you believe that a test question is defective in some way, please list your complaint here. All complaints will be considered in our interpretation of the test results.

Question number: _____ Your answer: _____

Your complaint:

Question number: _____ Your answer: _____

Your complaint:

Question number: _____ Your answer: _____

Your complaint: