Name: _____
Student #: _____

This is a 1 hour 50 minute closed book exam.

In order to receive credit, you must show your work or give brief explanations, where appropriate. You will be graded only on the work you show. Partial credit may be given when you show the process used to solve a problem, even if your answer is incorrect. You may also use a calculator, <u>although be sure to show your work—do not just put down an answer.</u> All numerical answers should be accurate to at least two decimal places (nearest hundredth). **You should report probabilities to four decimal places, if applicable. Don't forget units where applicable.**

The exam consists of 8 problems worth a total of 105 points, but the exam will be scored out of 100 points. Point values for each part of a question are designated in parentheses at the beginning of the problem. Take a moment to look through the exam before you begin in order to budget your time. If you don't know how to do a particular problem, move on to a new problem. Return to any unfinished problems after completing the ones that are easier for you.

GOOD LUCK!

Problem 1 (3 points each): Some summary statistics for the final exam scores from an introductory statistics class are shown in the table below, in addition to a histogram of the scores. The exam was worth 100 points.

Minimum	Q1	Median	Mean	Q3	Maximum
5	65	85	77	95	99
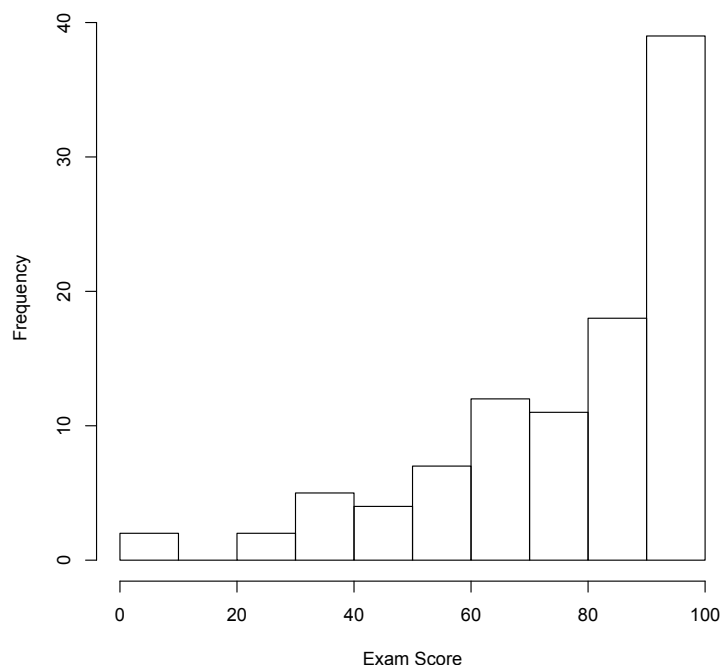
a) The distribution of the data can be described as

   i)   bimodal, right-skewed
   ii)   unimodal, left-skewed
   iii) unimodal, approximately symmetric
   iv) bimodal, left-skewed
   v)   unimodal, right-skewed

b) The lower bound for a flagged outlier based on the 1.5IQR criteria is

   i)   35
   ii)   20
   iii) 45
   iv) 40

c) Based on the given information, we can conclude

   i)   There is one outlier
   ii)   There are exactly two outliers
   iii) There are at least two outliers

d) List two other types of graphs that you could have used to visually display these data (no need to draw them, just list their names).

Problem 2 (2 points each): A candy dish contains five blue and three red candies. A child reaches up and selects three candies without looking.

a) What is the probability that there are two blue and one red candies in the selection?

b) What is the probability that the candies are all blue?

Problem 3 (3 points each): A college student frequents one of two coffee houses near campus, choosing Starbucks 70% of the time and Peet's 30% of the time. Regardless of where she goes, she buys a café mocha on 60% of her visits. Use the following events to answer a) and b).

> Let S = student chooses Starbucks
> Let P = student chooses Peet's
> Let C = student orders a café mocha

a) The next time the student goes into a coffee house near campus, what is the probability that she goes to Starbucks and orders a café mocha?

b) If she goes into a coffee house and orders a café mocha, what is the probability that she is at Peet's?

Problem 4 (3 points each): A company has five applicants for two positions: two women and three men. The five applicants are equally qualified and there was no preference in choosing either gender. Let x = the number women chosen to fill the two positions. Then $p(x)$ is given in the following table.

X	0	1	2
$p(x)$	0.30		0.10

a) What is $p(1)$?

b) What is the expected value of x?

c) The standard deviation of x is

    i)   0.1680 women

    ii)   0.4099 women

    iii)   1.0000 woman

Problem 5 (3 points each): A test consists of 20 multiple choice questions with four choices for each question. As an experiment, you GUESS on each and every answer without even reading the questions. Define X to be the number of questions you get correct based on guessing and assume that X follows a binomial distribution.

a) In order to minimally pass the test, you need to correctly guess 12 answers. The <u>exact</u> probability that you guess correctly on exactly 12 questions is

    iv)   0.00075

    v)   0.000000006

    vi)   0.0004

b) What is the <u>mean</u> and <u>standard deviation</u> for the number of correct guesses on the test?

c) Using normal approximation **with continuity correction** the estimated probability that you correctly guess the right answer for at least 12 of the questions is

    vii)   0.0002

    viii)   0.0004

    ix)   0.9996

Problem 6 (3 points/part): A gasoline tank for a certain model car is designed to hold 12 gallons of gas. Suppose that the actual capacity of the gas tank in cars of this type is well approximated by a normal distribution with $\mu = 12.0$ gallons and $\sigma = 0.2$ gallons.

a) Find the gallons of gas, x, that separates the highest 80% of gas tank capacities from the rest.

b) If you randomly select one car of this model, what is the probability that the car's gas tank holds at most 11.7 gallons?

c) What is the probability that the gas tank capacity of a randomly selected car is exactly 12.2 gallons?

Problem 7 (3 points each): Indicate whether the statement is true or false. **If false, give a careful explanation as to why the statement is false**. If true, just indicate true.

_____ A study is an observational study if the investigator observes the behavior of a response variable when one or more factors are manipulated.

_____ As long as the sample size is small relative to the population, there is little practical difference between sampling with replacement and sampling without replacement.

_____ A placebo is identical in appearance to the treatment of interest, but contains no active ingredients.

_____ The least squares line passes through the point $(\bar{x}, \bar{y})$.

_____ The higher the value of the coefficient of determination, the greater the evidence for a causal relationship between x and y.

_____ If two events, A and B, are mutually exclusive, then $P(A \text{ and } B) = P(A)\,P(B)$.

_____ For every random variable, $P(a \leq x \leq b) = P(a < x < b)$.

Problem 8 (3 points each): Circle the single best response for each multiple-choice question.

a) Which of the following statements is false?

   A. The explanatory variables are those variables that have values that are controlled by the experimenter.

   B. The response variable is the variable that the experimenter thinks may be affected by the explanatory variables.

   C. An experimental unit is the smallest entity to which a treatment is applied.

   D. Two variables are confounded if their effects on the response variable can be distinguished.

   E. An experiment in which experimental units are randomly assigned to treatments is called a completely randomized experiment.

b) A researcher wishes to study the relationship between the level of background noise and mental concentration. The treatment (noise level) will have three levels: no noise, low-intensity noise, and high-intensity noise. The subjects are to be divided into three groups, and each group is to receive one of the treatments. He has available to him a set of 60 female volunteers and a set of 90 male volunteers. What experimental design strategy would help him eliminate the introduction of gender as a confounding variable?

   A. stratified sampling

   B. replication

   C. blocking

   D. systematic sampling

   E. double-blind trials

c) Which of the following stem and leaf plots corresponds a data distribution that is positively (right) skewed?

I.

Leaf unit = 1	
2	159
3	
4	389
5	03346
6	004567
7	59
8	7

II.

Leaf unit = 1	
1	0124458
2	11135
3	29
4	239
5	3
6	36

III.

Leaf unit = 1	
1	6
2	
3	
4	
5	3457
6	58
7	0347
8	022346899

   A. I only

   B. II only

   C. III only

   D. I and III only

   E. None of the distributions are positively skewed.

6

d) Suppose that the distribution of weekly water usage for single-family homes in a particular city is mound shaped and approximately symmetric. The mean is 1,300 gallons, and the standard deviation is 200 gallons. What is the approximate value of the 98th percentile?

    A. 1,700 gallons

    B. 1,100 gallons

    C. 1,900 gallons

    D. 1,300 gallons

e) Data on the weight of a pickup truck (pounds) and distance (in feet) required for a truck traveling 40 miles per hour to come to a complete stop for 30 trucks was used to fit the least squares regression line $\hat{y} = 26 + 0.05x$. Which of the following statements is a correct interpretation of the value 0.05 in the equation of the regression line?

    A. On average, the truck weight goes up 0.05 pound for each additional foot required to stop the truck.

    B. On average, the stopping distance is 0.05 foot when the truck weight is 0.

    C. On average, the stopping distance goes up 0.05 foot for each 1-pound increase in truck weight.

    D. Approximately 5% of the variation in the stopping distances can be explained by the linear relationship between stopping distance and truck weight.

f) The Seattle Mariner's 25-man roster has 11 left-handed batters. The team has fallen below 500 so the frustrated manager has decided to randomly select 9 players to start the next game. Let X be the number of left-handed hitters that will start the game. The probability of drawing a starting lineup with 4 lefties is

    A. $\dfrac{\binom{9}{4}\binom{11}{5}}{\binom{25}{9}}$

    B. $\dfrac{\binom{11}{4}\binom{14}{5}}{\binom{25}{9}}$

    C. none of the above

g) A student is collecting data on movie preferences for a class project. The student stands outside a neighborhood theater and asks people exiting the theater if they would be willing to take a short survey. The student is able to get 50 people willing to participate over the course of six hours. The 50 completed surveys is an example of a

A. stratified sample

B. cluster sample

C. simple random sample

D. convenience sample

h) The random variable X has a continuous uniform distribution on the interval from 5 to 21. The height of the probability density function is

A. 0.0385

B. 0.0625

C. 0.0476

D. None of the above

i) The random variable X has a continuous uniform distribution on the interval from 5 to 21. $E(X)$ is

A. 8

B. 13

C. none of the above

j) The random variable X has a continuous uniform distribution on the interval from 5 to 21. The $P(8 \leq X \leq 15)$ is approximately

A. 0.40

B. 0.375

C. 0.50

D. None of the above

k) A car salesman sells on average two cars per week. Assume car sales follow a Poisson distribution. Assuming that the salesman works a five-day workweek, the probability that he will sell one car on a given day is

A. 0.2681

B. 0.1353

C. 0.6703