Name: _____ Date: _____ *Version 3*

Instructions:

a. You have two hours to complete the 42 questions. Questions that will take longer than others are marked with an asterisk (*), to save you time. The questions are roughly in the order that we covered them in class, though not exactly.
b. When you are done with the exam, turn in the exam, your answer sheet, and two evaluations – one for the lecture and one for the practical. Ensure that the evaluation of the practical is clearly marked with your group number (I can help if you need to know your group number)
c. You may not have anything on your desk except a pencil/pen, the exam, a calculator, and your ID. You may not look at or communicate with other students in any way, or leave the room without asking first.

1.	*The scores on a university examination are Normally distributed with a mean of 62 and a standard deviation of 11. If the bottom 5% of students will fail the course, what is the lowest mark that a student can have and still be awarded a passing grade?
A)	40
B)	62
C)	57
D)	44

2.	Researchers are conducting a state-wide survey for the Postal Service. The survey records many different variables of interest. Which of the following variables is categorical?
A)	Age of respondent.
B)	City of residence.
C)	Number of people, both adults and children, living in the household.
D)	Total household income, before taxes, in 2003.

Use the following to answer questions 3-5:

The United States Environmental Protection Agency records data on the fuel economy (miles per gallon) of many different makes of cars.

3. Data on the mileage of 20 randomly selected cars are listed below. The values are ordered for convenience.

12	13	15	16	16	17	18	18	19	19

	20	20	22	23	24	26	26	27	27	29

What is the median mileage for these 20 cars?

A)	19.5 miles per gallon
B)	17.5 miles per gallon
C)	20 miles per gallon
D)	19 miles per gallon

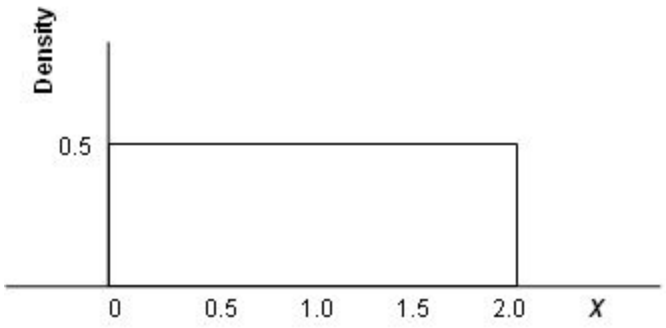
4.	If the value 29 were misrecorded and should really be 21, what would the median mileage be for these 20 cars?
A)	We cannot determine this from the given information.
B)	It would change to 20 miles per gallon.
C)	It would stay the same.
D)	It would change to 21 miles per gallon.

5.	What is the interquartile range for the mileage data?
A)	16.5 miles per gallon
B)	17 miles per gallon
C)	25 miles per gallon
D)	8.5 miles per gallon

6.	A set of midterm exam scores has a median that is much larger than the mean. Which of the following statements is most consistent with this information?
A)	The data set must be so large that it would be better to draw a histogram rather than a stemplot.
B)	A stemplot of the data would be skewed right.
C)	A stemplot of the data would be symmetric.
D)	A stemplot of the data would be skewed left.

7.	There are three children aged 3, 4, and 5 in a room. If another 4-year-old child enters the room, what will happen to the mean and variance of the ages of the children in the room?
A)	The mean and variance will both decrease.
B)	The mean and variance will both stay the same.
C)	The mean will stay the same but the variance will increase.
D)	The mean will stay the same but the variance will decrease.

Use the following to answer question 8:



8.	For this density curve, what percent of the observations lie between 0.5 and 1.2?
A)	70%
B)	25%
C)	35%
D)	50%

9.	Which of the following best describes correlation?
A)	Correlation measures the strength of the linear relationship between two quantitative variables.
B)	Correlation measures the strength of the linear association between two categorical variables.
C)	Correlation measures the strength of the relationship between any two variables.
D)	Correlation measures how much a change in the explanatory variable causes a change in the response variable.

10.	Recall that when we standardize the values of a variable, the standardized value has mean 0 and standard deviation 1. Suppose we measure two variables x and y on each of several subjects. We standardize both variables and compute the least-squares regression line of y on x for these standardized values. Suppose the slope of this least-squares regression line is –0.44. What conclusion can we draw?
A)	The correlation will be –0.44.
B)	The intercept will be 1.0.
C)	The intercept will also be –0.44.
D)	The correlation will be 1.0.

11.	The correlation coefficient between two variables x and y is $r = 0.121$. What conclusion can we draw?

A)	The correlation between x and y is low, but that does not matter. We can still use least-squares regression to calculate an equation of the form $= ax + b$.
B)	Because the correlation is so low, the relationship between x and y is not very strong, thus there is no use in studying this relationship.
C)	Because the correlation is so low, we only know that the linear relationship between x and y is not very strong, but there may a different relationship between the two variables. We need to first look at a scatterplot.
D)	None of the above.

Use the following to answer questions 12-13:

John's parents recorded his height at various ages between 36 and 66 months. Below is a record of the results:

Age (months)	36	48	54	60	66
Height (inches)	34	38	41	43	45

12.	John's parents decide to use the least-squares regression line of John's height on age based on the data in the previous problem to predict his height at age 21 years (252 months). What conclusion can we draw?
A)	Such a prediction could be misleading, because it involves extrapolation.
B)	John's height, in inches, should be about half his age, in months.
C)	The parents will get a fairly accurate estimate of his height at age 21 years, because the data are clearly correlated.
D)	All of the above.

13.	Which of the following could be the equation of the least-squares regression line of John's height on age? (Note: You do not need to directly calculate the least-squares regression line to answer this question.)
A)	Height = 22 + 0.34 × (Age)
B)	Height = 12 × (Age)
C)	Height = Age/12
D)	Height = 60 – 0.22 × (Age)

Use the following to answer question 14 and 18:

Many high school students in the United States take either the SAT or the ACT, which are different college entrance exams. However, some students take both. Data were collected from 60 students who took both college entrance exams. The average SAT score was 912 with a standard deviation of 180. The average ACT score was 21 with a standard deviation of 5. The correlation between the two variables equals 0.817.

14.	To predict the SAT score from a student's ACT score, what is the equation of the least-squares regression line?
A)	$= 156 + 36x$
B)	Cannot be determined from the information given.
C)	$= 294.348 + 29.412x$
D)	$= 0.3027 + 0.0227x$

15.	Fill in the blank. Will a fluoride mouthwash used after brushing reduce cavities? Twenty sets of twins were used to investigate this question. One member of each set of twins used the mouthwash after each brushing, the other did not. After 6 months, the difference in the number of cavities of those using the mouthwash was compared with the number of cavities of those who did not use the mouthwash. This experiment uses _____
A)	a matched pairs design.
B)	double replication.
C)	random placebos.
D)	double-blinding.

Use the following to answer question 16:

Researchers wish to determine if a new experimental medication will reduce the symptoms of allergy sufferers without the side effect of drowsiness. To investigate this question, the researchers gave the new medication to 50 adult volunteers who suffer from allergies. Forty-four of these volunteers reported a significant reduction in their allergy symptoms without any drowsiness.

16.	How could this study be improved?
A)	Repeat the study with only the 44 volunteers who reported a significant reduction in their allergy symptoms without any drowsiness, giving them a higher dosage this time.
B)	Use randomization to divide the volunteers into 2 groups—one to receive the experimental medication and one to receive the placebo.

C)	Include people who do not suffer from allergies in the study in order to represent a more diverse population.
D)	Give the spouses of the volunteers a placebo.

Use the following to answer question 17:

A medicine to remove the redness in eyes was tested in a group of 100 students. Each student took either the medicine or a placebo in both eyes. The specific treatment for each student was decided by flipping a coin. The participants in the study did not know if heads or tails resulted in the medication.

17.	Fill in the blank. The participants were given some free time after receiving their drops. The researcher did not keep track of what they did in that free time. It turns out that many of the people receiving the medicine spent their time outside and many of the people receiving the placebo spent their time inside. The results may now be biased because the variable location (inside vs. outside) is an example of a _____ variable.
A)	response
B)	lurking
C)	control
D)	confounding

Use the information in question 14 to answer the following question:

18.	What fraction of the variation in the values of the SAT scores is accounted for by the linear relationship between SAT and ACT scores?
A)	90.4%
B)	66.7%
C)	Cannot be determined from the information given.
D)	81.7%

19.	A researcher wishes to describe the relationship between the income of a couple and the amount of stress the couple experiences. At low incomes, an increase in income of $10,000 decreases stress ratings by 10 points (out of 100). At high incomes, an increase in income of $10,000 decreases stress ratings by 0.5 points. Of the following, which would be the best way for this researcher to quantify the relationship?
A)	Least-square regression
B)	Spearman's Rho
C)	Pearson's correlation (r)

D)	Cohen's Kappa

20.	In which of the following situations is LOWESS better than a least-square regression line?
A)	When the variables are categorical
B)	When the relationship between two variables is clearly linear
C)	When the relationship between to variables is nonmonotonic
D)	All of the above

21.	A teacher is using two teaching assistants to grade essay exams. The teaching assistants use the teacher's instructions to grade each exam pass or fail. The teacher is interested in how clear her instructions are, so she selects 20 essays and has both assistants grade them. How might she determine how well their pass fail grades coincide with one another?
A)	Pearson's correlation (r)
B)	Spearman's Rho
C)	Kendall's Tau
D)	None of the above

Use the following to answer questions 22-23:

The statistics of a particular basketball player state that he makes 4 out of 5 free-throw attempts.

22.	The basketball player is just about to attempt a free throw. What do you estimate the probability that the player makes this next free throw to be?
A)	0.80
B)	1.2
C)	50-50. Either he makes it or he doesn't.
D)	0.16

23.	During a season, the basketball player makes an average of about 200 free-throw attempts. On how many of these throws would you estimate he will actually score a point?
A)	160
B)	120
C)	80
D)	100

24.	Which of the following is (are) appropriate statements about randomness and/or probability?
A)	Probability describes only what happens in the long run.
B)	In a small or moderate number of repetitions, the observed proportion of an outcome can be far from the probability of the outcome.
C)	A phenomenon is called random if individual outcomes are uncertain but in a large number of repetitions there is a regular distribution of outcomes.
D)	All of the above are appropriate statements.

25.	Event A occurs with probability 0.2. Event B occurs with probability 0.8. If A and B are disjoint (mutually exclusive), then
A)	P(A and B) = 1.0
B)	P(A or B) = 0.16
C)	P(A or B) = 1.0
D)	P(A and B) = 0.16

Use the following to answer questions 26-29:

Ignoring twins and other multiple births, assume babies born at a hospital are independent events with the probability that a baby is a boy and the probability that a baby is a girl both equal to 0.5.

26.	What is the probability that the next three babies are of the same sex?
A)	0.250
B)	0.500
C)	0.125
D)	0.375

27.	Define events A = {the next two babies are boys} and B = {at least one of the next two babies is a boy}. What do we know about events A and B?
A)	They are independent.
B)	They are disjoint.
C)	They are complements.
D)	None of the above.

28.	Define event B = {at least one of the next two babies is a boy}. What is the probability of the complement of event B?
A)	0.250
B)	0.500

C)	0.125
D)	0.375

29.	What is the probability that at least one of the next three babies is a boy?
A)	0.750
B)	0.875
C)	0.333
D)	0.125

Use the following to answer question 30:

A particular city is serviced by three airlines for its passenger traffic. Airline A carries 50% of the passengers, Airline B 30%, and Airline C the remaining 20%. Each of the airlines is responsible for handling its security. The probabilities that a passenger carrying some type of weapon will be detected by Airlines A, B, and C are 0.9, 0.5, and 0.4, respectively.

30.	*If a weapon was detected on a passenger, what is the probability that Airline B detected it?
A)	0.22
B)	0.78
C)	0.15
D)	0.55

Use the following to answer questions 31-32:

It is claimed that 55% of marriages in the state of California end in divorce within the first 15 years. A large study was started 15 years ago and has been tracking hundreds of marriages in the state of California.

31.	*Suppose ten marriages are randomly selected. What is the probability that less than two of them ended in a divorce?
A)	0.0274
B)	0.0130
C)	0.0021
D)	0.0045

32.	*Suppose 100 marriages are randomly selected. What is the probability that less than 20 of them ended in a divorce?

A)	0.0229
B)	Less than 0.0001
C)	0.0055
D)	0.0130

Use the following to answer questions 33-34:

Suppose that a particular candidate for public office is in fact favored by $p = 48\%$ of all registered voters. A polling organization is about to take a simple random sample of voters and will use , the sample proportion, to estimate p.

33.	*How many voters need to be sampled to guarantee that the standard deviation is no more than 0.025?
A)	250
B)	400
C)	249
D)	399

34.	*Suppose that the polling organization takes a simple random sample of 500 voters. What is the probability that the sample proportion will be greater than 0.5, causing the polling organization to predict the result of the upcoming election incorrectly?
A)	0.212
B)	0
C)	0.5
D)	0.185

35.	For which of the following does the random variable X have a binomial distribution?
A)	X is the number of speeding tickets given out at a randomly picked location in a city during a calendar year.
B)	X is the number of defects found in 100 meters of fiber optic cable.
C)	X is the number of people in a random sample of size 50 from a large population that have type-AB blood.
D)	X is the number of pastrami sandwiches sold at a deli in a month.

36.	Which of the following is NOT required for a binomial setting?
A)	A count is being measured
B)	Np is greater than about 10
C)	Observations are independent

D)	All observations have the same probability of success

Use the following to answer questions 37-39:

The distribution of the amount of money undergraduate students spend on books for a term is slightly right-skewed, with a mean of $400 and a standard deviation of $80.

37.	*If a student is selected at random, what is the probability that this student spends more than $425 on books?
A)	0.3773
B)	This cannot be determined from the information given.
C)	0.1125
D)	0.6227

38.	In a simple random sample of 100 undergraduate students, what is the expected value of the sample mean amount of money spent on books?
A)	Anywhere between $320 and $480.
B)	Anywhere between $392 and $408.
C)	This cannot be determined from the information given.
D)	$400

39.	*If a simple random sample of 100 undergraduate students is selected, what is the probability that these students spend more than $425 on books, on average?
A)	0.3773
B)	0.00089
C)	This cannot be determined from the information given.
D)	0.2353

Use the following to answer questions 40-41:

A random variable X is Normally distributed with mean = 75 and = 8. Let Y be a second Normally distributed random variable with mean = 70 and = 12. It is also known that X and Y are independent of one another. Let W be a random variable that is the difference between X and Y, i.e., W = X – Y.

40.	What can be said about the distribution of W?
A)	W is $N(5, 14.4)$
B)	W is $N(0, 20)$

C)	W is $N(0, 14.4)$
D)	W is $N(5, 20)$

41.	*What is the probability that W is greater than –5.5, i.e., P{W > –5.5}?
A)	Not within ± 0.005 of any of the above.
B)	0.767
C)	0.233
D)	0.512

42.	A length of chain is to be constructed by placing 36 component links end-to-end. The length of a link produced by a production process is known to be a random variable with mean $\mu = 2.5$ cm with a standard deviation $\sigma = 0.2$ cm. The 36 links are chosen at random from this process to produce the chain. The mean and standard deviation of the length of chain are, respectively
A)	36 cm, 7.2 cm.
B)	90 cm, 7.2 cm.
C)	36 cm, 1.2 cm.
D)	90 cm, 1.2 cm.