Name _____ Date _____ Per_____

AP REVIEW PACKET #6 INSTRUCTIONS – Due by Friday 4/12 7:50 – 40 points

1) Read Topic 8: Planning and Conducting Experiments, Topic 9: Probability as Relative Frequency, and 10: Combining Independent Random Variables. While you read, answer the questions below.

2) Answer the multiple choice questions in this packet, these will be graded for accuracy. Show all work and then copy your answers into the blanks below.

31. _____	32. _____	33. _____	34. _____	35. _____
36. _____	37. _____	38. _____	39. _____	40. _____

3) Answer the Free-Response Questions.

Optional Bonus Activity: Go to http://learner.org/resources/series65.html. Watch programs 8 & 9 answer the worksheet.

Guided Reading Questions for Packet #6:

1) In experiments, what are explanatory variables called? _____ .

2) A group is treated with some _____ of the explanatory variable, and the _____ of the response variable is measured.

3) Define a) confounded _____.

b) lurking variable_____.

c) placebo effect_____.

4) Why is it important to use randomization in experiments?

5) In double-blinding neither the _____ nor those _____ their responses know who is receiving which treatment.

6) Define relative frequency: And probability:

7) What is the probability that a household chosen at random and found to have an answering machine also has call waiting? (write out the Formula, work, and answer exactly as given in the book)

8) What formula was used to solve #2? (Write out the formula)

9) _____ _____ can be useful in working with conditional probabilities.

10) What is a probability distribution?

11) What application can you use if you forget the order of the variables on the TI-84? _____ _____

Instead of ENTER< click on the _____ button.

12) The geometric probability formula is NOT given on your AP formula sheet. What is it?

13) Read the summary: copy the sentence that describes the content that is most difficult to you.

Name _____ Date _____ Per_____

Packet Multiple Choice Questions: 31. A ski resort rental shop wants to make sure that they have enough of the three different types of skis on hand for all of their customers. A random sample of skiers was taken to determine if males or females tend to prefer one type of ski versus another. The three types of skis they offer for rental are high performance, parabolic, and beginner (short skis). Which is a correct pair of hypotheses for the ski resort shop to test?

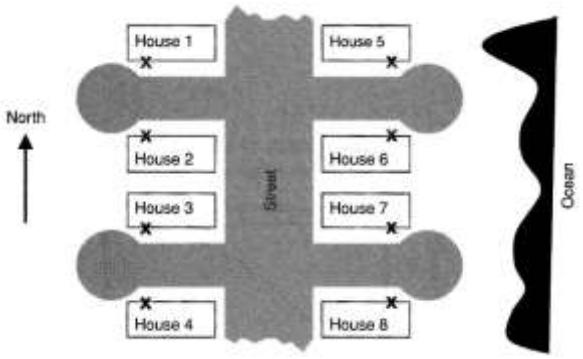
(A) H_0: Gender and type of ski rented are independent.
H_a: Gender and type of ski rented are not independent.

(B) H_0: Gender and type of ski rented are not independent.
H_a: Gender and type of ski rented are independent.

(C) H_0: The proportions of types of skis rented are the same.
H_a: The proportions of types of skis rented are not the same.

(D) H_0: There is an association between gender and type of ski rented.
H_a: There is no association between gender and type of ski rented.

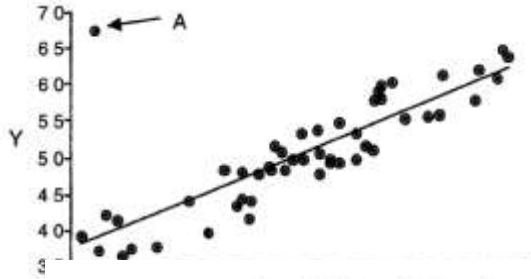
(E) H_0: The proportions of genders renting skis are the same.
H_a: The proportions of genders renting skis are not the same.



32. A garage door manufacturer has developed a new type of door for houses in the Southeast part of the United States. Doors in this area of the country are particularly susceptible to damage from salty ocean spray and the sun's rays, which tend to shine mainly on the north side of the house. An experiment will test the new garage door against the existing type of door on eight houses in a particular residential area. An overhead view of the area is shown below. The location of the garage door on each home is marked with an "X."

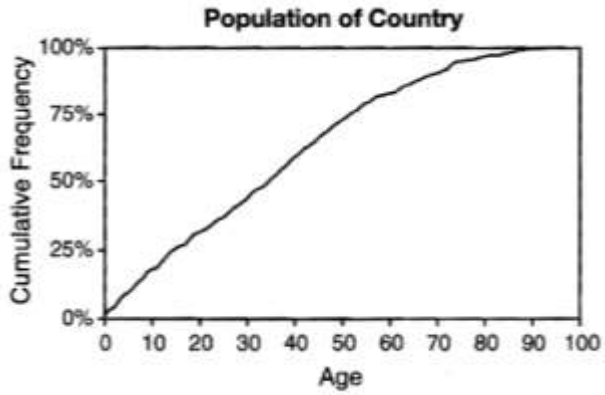
Which of the following blocking schemes is most appropriate to account for variables in this study other than the type of door?
(A) Form the houses into two blocks: {I, 2, 3, 4} and {5, 6, 7, 8}. (B) Form the houses into two blocks: {I, 3, 5, 7} and {2, 4, 6, 8}.
(C) Form the houses into four blocks: {I, 5}, {2, 6}, {3, 7}, and {4,8}. (D) Form the houses into four blocks: {I, 3}, {2, 4}, {5, 7}, and {6,8}.
(E) No blocking is necessary in this experiment.

33. What is the approximate residual of the data point "A" on the scatterplot with the least-squares regression line shown right?



(A) 11 (B) 29 (C) 39 (D) 58 (E) 68

34. The distribution of the population by age in a particular country is represented by the given cumulative relative frequency plot. Which of the following statements about the population is true?



(A) The median age is about 46 years.

(B) The interquartile range of ages is about 27 years.

(C) The mean age is greater than the median age.

(D) There are more people younger than 65 years of age than those older than 30 years of age.

(E) Seventy-five percent of people are older than 50 years of age.

35. A significance test was performed with hypotheses of H_o: $\mu = 15$ and H_a: $\mu < 15$. The P-value for the test was 0.026. Which of the following conclusions is most appropriate for this test?

(A) 2.6% of the time the mean will be exactly 15.

(B) 97.4% of the time the mean will be less than 15.

(C) There is a 2.6% chance of making a Type II error.

(D) There is reason to believe that the mean μ is less than 15.

(E) The true mean is approximately 2.6% of 15.

Name _____ Date _____ Per_____

35. A significance test was performed with hypotheses of H_0: $\mu = 15$ and H_a: $\mu < 15$. The P-value for the test was 0.026. Which of the following conclusions is most appropriate for this test?

(A) 2.6% of the time the mean will be exactly 15.
(B) 97.4% of the time the mean will be less than 15.
(C) There is a 2.6% chance of making a Type II error.
(D) There is reason to believe that the mean, μ, is less than 15.
(E) The true mean is approximately 2.6% of 15.

36. A consumer is considering the purchase of a new appliance and has narrowed the selection down to two brands: Brand M and Brand S. The consumer found the following information about the repair history of the appliances as done by a nonprofit product-testing group.

The consumer can purchase an extended warranty for either brand at a cost of $300. The extended warranty provides free repairs for a 5-year period. For which of the brands under consideration is the cost of the extended warranty less than the expected cost of repairing the appliance over a 5-year period?

(A) Brand M only
(B) Brand S only
(C) Both Brand M and Brand S
(D) Neither Brand M nor Brand S
(E) It cannot be determined without knowing the purchase price of the appliance.

Cost of Brand		Percentage needing repairs within 5 years			
		0 repair	1 repair	2 repairs	3 repairs
M	$350	50%	25%	15%	10%
S	$325	40%	30%	20%	10%

37. A class consists of 10 male and 10 female students. If four students are selected at random, without replacement, what is the probability that all four are male?

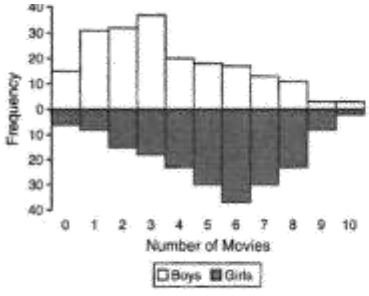
(A) $\frac{20!}{4!16!}(0.5)^4(0.5)^{16}$
(B) $\frac{10!}{4!6!}(0.5)^4(0.5)^6$
(C) $(0.5)^4$
(D) $\frac{\left(\frac{10!}{6!}\right)}{\left(\frac{20!}{16!}\right)}$
(E) $P(z > 2)$

38. A light bulb manufacturer wishes to estimate the mean lifetime (in hours) of its new "long-life" bulb. Thirty bulbs were tested and the lifetime was recorded for each. The mean of the sample was 1,450 hours and the standard deviation of the sample was 150 hours. Assuming all conditions for inference are satisfied, what is the 99% confidence interval for J.1, the mean lifetime of the new "long-life" bulb?

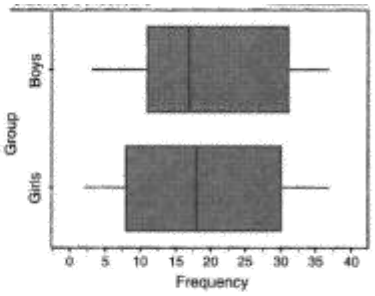
(A) $1450 \pm 2.750\frac{150}{\sqrt{30}}$
(B) $1450 \pm 2.756\frac{150}{\sqrt{30}}$
(C) $1450 \pm 2.750\frac{150}{\sqrt{29}}$
(D) $1450 + 2.756\frac{150}{\sqrt{29}}$
(E) $1450 \pm 0.8389\frac{150}{\sqrt{29}}$

39. A statistics class randomly surveyed 200 boys and 200 girls at their school and asked each respondent how many movies they saw in theaters over summer break. The results are shown in the frequency table below. Which of the following graphs is appropriate to compare the number of movies seen by boys and girls over summer break?

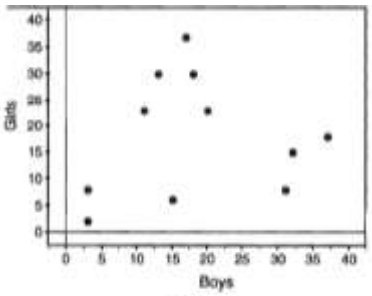
Number of Movies	Frequency of Boys	Frequency of Girls
0	15	6
1	31	8
2	32	15
3	37	18
4	20	23
5	18	30
6	17	37
7	13	30
8	11	23
9	3	8
10	3	2



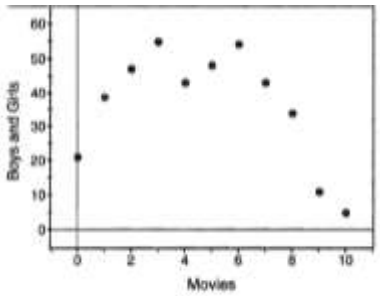
(A)



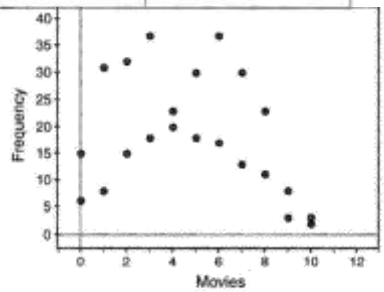
(B)



(C)



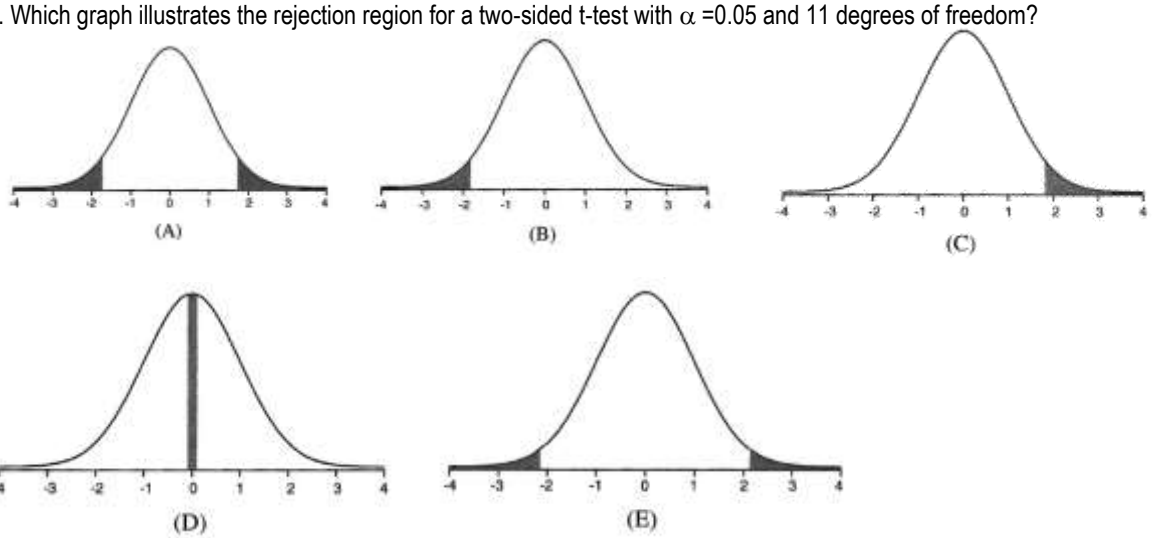
(D)



(E)

Name _____ Date _____ Per_____

40. Which graph illustrates the rejection region for a two-sided t-test with α =0.05 and 11 degrees of freedom?



Packet Free Response Questions:

1) A researcher wants to conduct a study to test whether listening to soothing music for 20 minutes helps to reduce diastolic blood pressure in patients with high blood pressure, compared to simply sitting quietly in a noise-free environment for 20 minutes. One hundred patients with high blood pressure at a large medical clinic are available to participate in this study.

(a) Propose a design for this study to compare these two treatments.

(b) The null hypothesis for this study is that there is no difference in the mean reduction of diastolic blood pressure for the two treatments and the alternative hypothesis is that the mean reduction in diastolic blood pressure is greater for the music treatment. If the null hypothesis is rejected, the clinic will offer this music therapy as a free service to their patients with high blood pressure. Describe Type I and Type II errors and the consequences of each in the context of this study, and discuss which one you think is more serious.

Name _____ Date _____ Per____

2) People with acrophobia (fear of heights) sometimes enroll in therapy sessions to help them overcome this fear. Typically, seven or eight therapy sessions are needed before improvement is noticed. A study was conducted to determine whether the drug D-cycloserine, used in combination with fewer therapy sessions, would help people with acrophobia overcome this fear. Each of 27 people who participated in the study received a pill before each of two therapy sessions. Seventeen of the 27 people were randomly assigned to receive a D-cycloserine pill, and the remaining 10 people received a placebo. After the two therapy sessions, none of the 27 people received additional pills or therapy. Three months after the administration of the pills and the two therapy sessions, each of the 27 people was evaluated to see if he or she had improved.

(a) Was this study an experiment or an observational study? Provide an explanation to support your answer.

(b) When the data were analyzed, the D-cycloserine group showed statistically significantly more improvement than the placebo group did. Based on this result, would the researchers be justified in concluding that the D-cycloserine pill and two therapy sessions are as beneficial as eight therapy sessions without the pill? Justify your answer.

(c) A newspaper article that summarized the results of this study did not explain how it was determined which people received D-cycloserine and which received the placebo. Suppose the researchers allowed the therapists to choose which people received D-cycloserine and which received the placebo, and no randomization was used. Explain why such a method of assignment might lead to an incorrect conclusion.

3) Scientists interested in preserving natural habitats and minimizing the possible extinction of certain bird species conducted a study to determine if it is better for conservation groups to purchase a few large nature preserves or many small preserves in order to meet these goals.

The scientist studied 13 randomly selected islands of different sizes to determine the risk of extinction for bird species. Islands are thought to be a good imitation of what would happen in a nature preserve because of their isolation. If a species lived on only one island, it was considered to be at risk. Scientists have determined that whether or not one species becomes extinct is independent of whether or not another species becomes extinct.

In 1990 scientists counted the number of at-risk species on each of the selected islands. They returned to each of these islands in the year 2000 to see whether the species still existed on the islands. Species that were present in 1990 but absent in 2000 were considered extinct. Data collected by the scientists are given in the table.

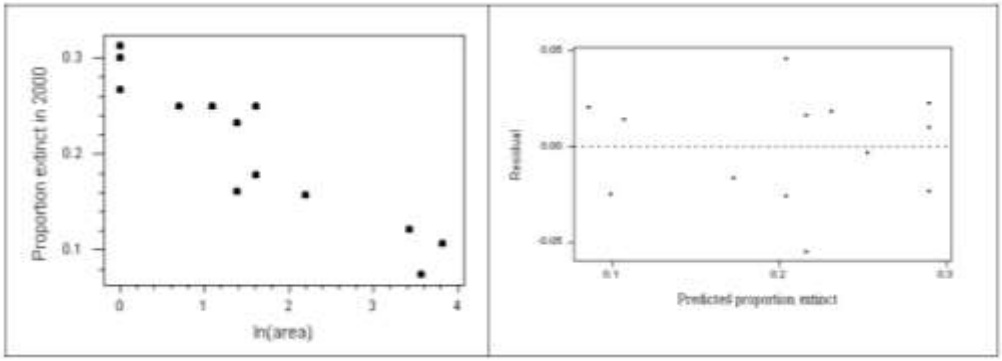
Island	Area (in sq km)	Species at Risk in 1990	Species Extinct by 2000	Proportion Extinct
1	46	75	8	0.11
2	36	67	3	0.04
3	31	66	8	0.12
4	9	51	8	0.16
5	5	28	5	0.18
6	5	20	6	0.30
7	4	43	10	0.23
8	4	31	5	0.16
9	3	28	7	0.25
10	2	32	8	0.25
11	1	30	8	0.27
12	1	20	4	0.20
13	1	16	5	0.31

(a) One scientist involved in the study believes that large islands (those with areas greater than 25 square kilometers) are more effective than small islands (those with areas of no more than 25 square kilometers) for protecting at-risk species. The scientist noted that for this study, a total of 19 of the 208 species on the large islands became extinct, whereas a total of 66 of the 299 species on the small islands became extinct. Assume that the probability of extinction is the same for all at-risk species on large islands and the same for all at-risk species on small islands. Do these data support the scientist's belief? Give appropriate statistical justification for your answer. (Can use the back of the paper or a separate sheet.)

Name _____ Date _____ Per_____

(b) Another scientist who worked on this study thinks that the proportion of species that become extinct is more directly related to the size of the islands than simply to whether the islands are grouped as large or small. This scientist investigated the relationship between the proportion of extinct birds and the area, in square kilometers, of islands. A least squares analysis was conducted on the proportion extinct and ln(area). The regression analysis output, the scatterplot, and the residual plot are shown below.

Predictor	Coef	StDev	T	P
Constant	0.28996	0.01269	22.85	0.000
ln(area)	-0.05323	0.00618	-8.61	0.000
S = 0.02863	R-Sq = 87.1%			



Estimate the slope of the least squares regression line using a 95 percent confidence interval. Interpret your answer in the context of this situation.

(c) In part (a), the scientist assumed that the probability of a species becoming extinct is the same for each of the large islands. Similarly, the scientist assumed that the probability is the same for each of the small islands. Based on your answer in part (b), do you think this is a reasonable assumption? Explain.

(d) A conservation group with a long-term goal of preserving species believes that all at-risk species will disappear whenever land inhabited by those species is developed. It has an opportunity to purchase land in an area about to be developed. The group has a choice of creating one large nature preserve with an area of 45 square kilometers and containing 70 at-risk species, or 5 small nature preserves, each with an area of 3 square kilometers and each containing 16 at-risk species unique to that preserve. Which choice would you recommend and why?

Name _____ Date _____ Per____

Video 8 Questions: DESCRIBING RELATIONSHIPS

1. What is a plot of quantitative variables? _____

2. What is the x-variable called in studies? _____ the y-variable? _____

3. What is a variable that records into which of several categories a case falls? _____

4. How do categorical variables enrich a scatterplot? _____

5. What type of smoothing is found by slicing the scatterplot vertically, calculating the median within each slice, and connecting these medians by a straight line? _____

6. What example in the video illustrates the use of a median trace? _____

7. What is the best fitting line that fits data by minimizing the sum of the squares of the residuals? _____

8. What example is used to illustrate the use of the least squares regression line? _____

9. In the equation $y = a + bx$, what is the formula for b? _____

   What is b in the equation? _____ What is the formula for a? _____

   What does y represent? _____ x? _____

   What is a in the equation? _____

10. Even though you can fit a regression line to any set of data, when is the line valid?

11. What are points with unusually large residuals? _____

12. What are points that deviate strongly in the x-direction? _____

Name _____ Date _____ Per____

Video 9 Worksheet - CORRELATION

1. What is the measure of the strength and direction of the linear relationship between quantitative variables?

   _____

2. What values does ***r*** vary between? _____

3. What indicates a perfect positive correlation? _____ a perfect negative correlation? _____

4. What study in the video illustrates the use of correlation? _____

   Which characteristics showed a strong correlation? _____

   Which characteristics showed a moderately strong correlation? _____

5. In the formula for ***r***, what do $\frac{x - \bar{x}}{s_x}$ and $\frac{y - \bar{y}}{s_y}$ do? _____

   Why does the formula divide by n – 1? _____

   When is ***r*** positive? _____

   When is ***r*** negative? _____

6. What kind of relationships does ***r*** measure? _____

7. What describes the amount of variation in y described by the linear relationship with x? _____

8. What example in the video uses the squared correlation coefficient? _____