

AP Stat Summer Assignment: 80 Flash Cards = 80 HW Points!!

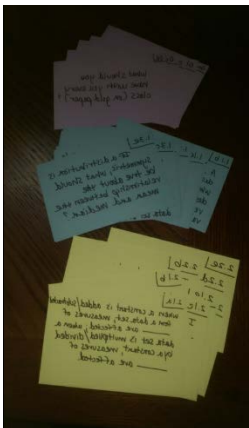
Chapter	Sections	Total Points
0	0	5
1	1.1, 1.2, 1.3	15
2	2.1, 2.2	10
3 and 12	3.1, 3.2, 12.2	15
4	4.1, 4.2, 4.3	8
5	5.1, 5.2, 5.3	7
6	6.1, 6.2, 6.3	10
7-12	H1 - H10	10

You will be creating the fronts only of these cards as the Summer Assignment. You should have them with you every class and should be continually studying them. As we cover the material, we will be adding the answers to the backs of the cards.

Some suggestions:

- Do each “chapter” on a different color card or write with a different color marker- easy to separate the cards for studying purposes.
- Use a zip-lock baggie or large binder clip to keep them all together. You can also buy index cards on a ring.
- You can create an electronic flash card file (there are apps to do this). It may be difficult to study only certain sections however.

There is a positive correlation between studying these flash cards and flash card quiz scores—meaning students who studied these cards scored higher on their flash card quizzes!!!



JOIN EDMODO:

www.edmodo.com

Join Group: 7657pj

Group Title: 17-18 AP STAT NCHS

Have a stat-tastic summer! See you in the fall!! Thank you from Mr. Byerly and Mrs. Fowler

Procedural Information

0a	When is the AP Stat exam? Date and AM or PM?	
0b	What day of the week are your flashcard quizzes?	
0c	What is the minimum number of external studies you must do?	
0d	When are your external studies due?	
0e	What you should have with you every class (on gold paper)?	

1.1 Analyzing Categorical Data

1.1a	A _____ of a categorical variable lists the categories and gives either the count or percent of individuals who fall within that category.	
1.1b	A _____ 1 _____ distribution describe all individuals, whereas a _____ 2 _____ distribution describes individuals with respect to one specific value of the variable.	1 2
1.1c	A _____ is a visual display that helps to compare and determine association amongst categorical variables.	
1.1d	Numbers are used to label choices, 1 = no, 2 = yes, 3 = not sure, are these considered categorical or quantitative?	
1.1e	It is important to include _____ when describing quantitative variables.	

1.2 Displaying Quantitative Data with Graphs

1.2a	When describing quantitative distributions, you should always discuss these four things.	
1.2b	Sketch a distribution that is skewed to the right.	
1.2c	Each quantitative graph should have _____1_____ and _____2_____ on the axes; include a _____3_____ on a stem-and-leaf plot.	1 2 3
1.2d	Which quantitative display is <i>best</i> used when data is given in ranges or intervals?	
1.2e	An area of high frequency in a graph is called a _____.	

1.3 Displaying Quantitative Data with Numbers

1.3a	To visually compare two quantitative distributions, you should use the _____.	
1.3b	What values make up the <i>five number summary</i> ?	
1.3c	To determine if a data set has outliers, multiply the _____ by 1.5 and subtract that value from Q1 and add that value to Q3. Outliers would be located within the data set but outside of these fences.	
1.3d	If a distribution is skewed, use the ___1___ for center and the ___2_____ for spread, since these values are resistant to outliers. If a distribution is symmetric, use the _____3___ for center and the ___4_____ for spread.	1 2 3 4
1.3e	If a distribution is symmetric, what should be true about the relationship between the mean and median?	

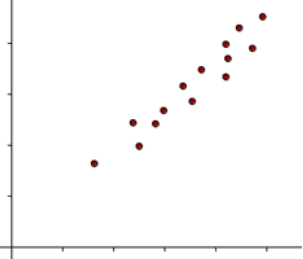
2.1 Describing Location in a Distribution

2.1a	When a constant is added/subtracted from a data set, measures of _____ 1 _____ are affected; when a data set is multiplied/divided by a constant, measures of _____ 2 _____ are affected.	1 2
2.1b	When calculating the _____, subtract the mean from your value and divide by the standard deviation.	
2.1c	In a symmetric distribution, the mean and median have a z-score of _____.	
2.1d	If a data value is at the 75 th percentile, _____% of the distribution is higher than the data value.	
2.1e	A _____ represents the number of standard deviations that the data value is above/below the mean.	

2.2 Density Curves and Normal Distributions

2.2a	This rule approximates the percentage of a normal distribution within one, two and three standard deviations from the mean.	
2.2b	If a normal model is given to you, use _____ 1 _____ to represent the mean and standard deviation. If you calculate these measures from data, use _____ 2 _____ to represent the mean and standard deviation.	1 2
2.2c	Draw a quick sketch of a normal model.	
2.2d	Give the three pieces of information represented by the notation, $N(30,5)$.	N 30 5
2.2e	A normal probability plot would be linear if the distribution is _____.	

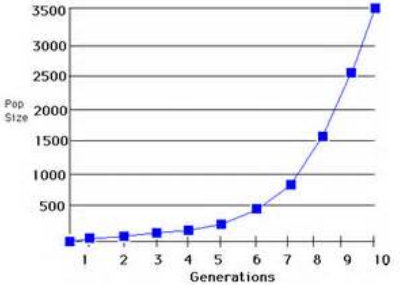
3.1: Scatterplots and Correlation

3.1a	When discussing association, these three things should be discussed.	<p>1)</p> <p>2)</p> <p>3)</p>
3.1b	_____ is a numerical measure of the direction and strength of a linear association.	
3.1c	<p>Describe the association of this scatterplot.</p> 	
3.1d	<p>Write the following r-values in terms of strength, weakest to strongest.</p> <p>$r = +0.8, -0.75, -1, 0$</p>	
3.1e	<p>The linear model helps to _____ about values relating to the data set.</p>	

3.2: Least-Squares Regression

3.2a	<p>In order to calculate a linear regression model from summary statistics, you need to have various components (means, standard deviations, and correlation) to write the model,</p> $\hat{y} = b_0 + b_1x$ <p>b_0 represents _____ 1 _____ and</p> <p>b_1 represents _____ 2 _____.</p>	<p>1</p> <p>2</p>						
3.2b	<p>A _____ is calculated by finding the difference between the actual value and the predicted value.</p>							
3.2c	<p>If a data set has an r-value of -.98 and a linear relationship, what would be the sentence describing the association between the two variables, x and y.</p>							
3.2d	<p>Interpolation is much more reliable than _____.</p>							
3.2e	<p>Write the linear model for the following computer printout (with context):</p> <p>Dependent variable: gestation</p> <p>R-squared = 72.2%</p> <table data-bbox="240 1570 630 1732"> <thead> <tr> <th>Variable</th> <th>Coefficient</th> </tr> </thead> <tbody> <tr> <td>Constant</td> <td>-39.5172</td> </tr> <tr> <td>LifExp</td> <td>15.4980</td> </tr> </tbody> </table>	Variable	Coefficient	Constant	-39.5172	LifExp	15.4980	
Variable	Coefficient							
Constant	-39.5172							
LifExp	15.4980							

12.2 Transforming to Achieve Linearity

12.2a	Transforming data allows us to make _____ using a linear model for nonlinear data.																							
12.2b	If there is a nonlinear pattern in a scatterplot, taking the _____ of the function helps to straighten the data.																							
12.2c	<p>Given the data below, creating a(n) _____ regression would not be appropriate.</p>  <table border="1" data-bbox="267 695 667 978"> <caption>Population Size over Generations</caption> <thead> <tr> <th>Generations</th> <th>Pop Size</th> </tr> </thead> <tbody> <tr><td>1</td><td>100</td></tr> <tr><td>2</td><td>150</td></tr> <tr><td>3</td><td>200</td></tr> <tr><td>4</td><td>250</td></tr> <tr><td>5</td><td>300</td></tr> <tr><td>6</td><td>400</td></tr> <tr><td>7</td><td>600</td></tr> <tr><td>8</td><td>1000</td></tr> <tr><td>9</td><td>1500</td></tr> <tr><td>10</td><td>3500</td></tr> </tbody> </table>	Generations	Pop Size	1	100	2	150	3	200	4	250	5	300	6	400	7	600	8	1000	9	1500	10	3500	
Generations	Pop Size																							
1	100																							
2	150																							
3	200																							
4	250																							
5	300																							
6	400																							
7	600																							
8	1000																							
9	1500																							
10	3500																							
12.2d	Which two plots should you look at to determine if you have a good model for your data?	<p>1)</p> <p>2)</p>																						
12.2e	<p>Given the following transformation model, determine the predicted value for age, if $x = 2$.</p> $\sqrt[3]{age} = 2x + 1$																							

4.1: Sampling and Surveys

4.2: Experiments

4.3: Using Studies Wisely

4.1a	A _____ is chosen in such a way that every group of n individuals in the population has an equal chance to be selected as the sample.	
4.1b	Name one type of bias that can affect an experiment or study.	
4.1c	If 29 assignments need to be made using a random digit table, you will need to use _____ digit(s).	
4.2a	The four principles of experimental design are:	1) 2) 3) 4)
4.2b	A _____ effect occurs when a subject responds favorably to a treatment that is a control (the response to a dummy treatment).	
4.2c	_____ is similar to stratifying in data sampling because individuals are divided into groups based on similar characteristics to compare responses to treatments in an experiment.	
4.3a	Inferences about the population and drawing conclusions using cause and effect are only valid if subjects are _____ assigned to groups.	
4.3b	True or False: In reference to studies or experiments: All individual data must be kept confidential. Only statistical summaries for groups of subjects may be made public.	

5.1: Randomness, Probability and Simulation 5.2: Probability Rules

5.3: Conditional Probability and Independence

5.1a	The probability of an event occurring is always between ____1____ and ____2____.	1 2
5.1b	The law of large numbers says that if we observe _____ repetitions of any chance process, the proportion of times that specific outcome occurs approaches a single value.	
5.2a	A ____1____ model is a description of some chance process that lists a sample space S and a ____2____ for each outcome.	1 2
5.2b	The probability of an event and it's complement always equal _____.	
5.2c	<p>What word fills in the blank: <u>AND</u> or <u>OR</u>?</p> <p>The general addition rule can be used to find $P(A \text{ ___ } B) = P(A \cup B) =$</p> $P(A) + P(B) - P(A \cap B)$ <p>by using a two-way table or a venn diagram.</p>	
5.3a	<p>What word fills in the blank: <u>AND</u> or <u>OR</u>?</p> <p>The general multiplication can be used to find $P(A \text{ ___ } B) =$</p> $P(A \cap B) = P(A) \cdot P(B A)$ <p>by using a tree diagram.</p>	
5.3b	True or false: Two mutually exclusive events can be independent.	

6.1: Discrete and Continuous Random Variables

6.2: Transforming and Combining Random Variables

6.3: Binomial and Geometric Random Variables

6.1a	<p>Find the expected number of times at bat for the given model.</p> <table border="1" data-bbox="203 535 844 651"> <tr> <td>Times at bat</td> <td>0</td> <td>1</td> <td>2</td> </tr> <tr> <td>Probability</td> <td>.25</td> <td>.25</td> <td>.5</td> </tr> </table>	Times at bat	0	1	2	Probability	.25	.25	.5		
Times at bat	0	1	2								
Probability	.25	.25	.5								
6.1b	<p>A continuous random variable is usually represented by a _____ model.</p>										
6.2a	<table border="1" data-bbox="203 903 844 1060"> <tr> <td></td> <td>mean</td> <td>Standard deviation</td> </tr> <tr> <td>PD1</td> <td>4</td> <td>2</td> </tr> <tr> <td>PD2</td> <td>6</td> <td>1</td> </tr> </table> <p>Determine the expected number of absences for the difference in PD1 and PD 2.</p>		mean	Standard deviation	PD1	4	2	PD2	6	1	
	mean	Standard deviation									
PD1	4	2									
PD2	6	1									
6.2b	<table border="1" data-bbox="203 1396 844 1554"> <tr> <td></td> <td>mean</td> <td>Standard deviation</td> </tr> <tr> <td>PD1</td> <td>4</td> <td>2</td> </tr> <tr> <td>PD2</td> <td>6</td> <td>1</td> </tr> </table> <p>Determine the standard deviation of absences for the difference in PD1 and PD 2.</p>		mean	Standard deviation	PD1	4	2	PD2	6	1	
	mean	Standard deviation									
PD1	4	2									
PD2	6	1									
6.3a	<p>To determine the probability for geometric and binomial distributions, you have to know the probability/proportion of _____.</p>										

Name _____

Date _____

Pd _____

6.3b	Geometric or binomial: You want to know the probability of getting 3 red marbles out of 30 marbles.	
6.3c	Geometric or binomial: You want to know the probability of picking a red marble on the 5 th one you pick.	
6.3d	pdf or cdf: You want to know the probability of picking 3 red marbles out of 30 marbles.	
6.3e	pdf or cdf: You want to know the probability of picking more than 3 red marbles out of a bag 30 marbles.	
6.3f	Geometric or binomial: $P(x=4) = (.25)^3(.75)$	

Hypothesis Tests

<p>H1</p>	<p>Ho: mean = 8</p> <p>Ha: mean < 8</p> <p>The 95% CI for a mean is (4 , 10).</p> <p>Does this interval support the claim?</p>	
<p>H2</p>	<p>Ho: mean = 4</p> <p>Ha: mean > 4</p> <p>$\bar{x} = 5$</p> <p>Draw a quick sketch of the normal model used in the mechanics for this hypothesis test.</p>	
<p>H3</p>	<p><i>Among randomly selected pets, 27% of the 188 dogs have fleas.</i></p> <p>What type of test would be appropriate?</p>	
<p>H4</p>	<p><i>Tags placed on garbage cans allow the disposal of up to 30 pounds of garbage. Is this strong evidence that residents overload their garbage cans?</i></p> <p>Write the alternate hypothesis for this situation.</p>	
<p>H5</p>	<p><i>Tags placed on garbage cans allow the disposal of up to 30 pounds of garbage. Is this strong evidence that residents overload their garbage cans?</i></p> <p>Write the null hypothesis for this situation.</p>	

Name _____

Date _____

Pd _____

H6	<p>$p = .125$</p> <p>$\alpha = .025$</p> <p>Would you reject or fail to reject the null hypothesis?</p>	
H7	<p>If you know</p> <p>$\alpha = .05$</p> <p>$p = .042$</p> <p>and you shaded to the left,</p> <p>which of the following is your hypothesis:</p> <p>$\text{mean} < 8$</p> <p>$\text{mean} > 8$</p> <p>$\text{mean} \neq 8$</p>	
H8	<p>If you know</p> <p>$\alpha = .05$</p> <p>$p = .042$</p> <p>Would you reject or fail to reject the null hypothesis?</p>	
H9	<p>True or false?</p> <p>$H_0: \text{mean} = 9$</p> <p>$H_a: \text{mean} > 9$</p> <p>Shade to the left</p> <p>$P(t > 1.75)$</p>	
H10	<p>True or false?</p> <p>$H_0: \text{mean} = 9$</p> <p>$H_a: \text{mean} > 9$</p> <p>Shade to the right</p> <p>$P(t > 1.75)$</p>	

Name _____ Date _____ Pd _____