Good luck to: ____ Period ________

Statistics Honors Final Exam Review Packet 2019

Final Exam:	Thursday, June 13th	Periods 4 & 5/6
	Friday, June 14th	Periods 3 & 7/8/9
	Monday, June 17th	Periods 2 & 9/10/11
	Tuesday, June 18th	Periods 1 & 12

Please bring the following to class on the day of your exam:

1.) Graphing calculator – I DO NOT HAVE ANY TO LEND OUT!!!
Make sure your batteries work. If you borrowed a school calculator, you will be turning it in once you are done with your exam. You must turn in the same calculator that was issued to you. The serial number must match the on you your signed paperwork. If you lost it, you must either pay $125 or buy a brand new one (unopened and with the store receipt).

2.) #2 pencils for the scantron section of your exam

3.) Statistics textbook– if you lost it, the cost is $160.

4.) This completed packet…it will be collected and graded on the day of the exam. I reserve the right to check work each day.

FINAL EXAM:

15 multiple choice (2 pts ea. = 30 pts) & choice of 2 open ended (10 pts ea. = 20 pts)

UTILIZE THE FORMULA SHEET PROVIDED ON THE NEXT PAGE WHILE COMPLETING THIS REVIEW. THIS IS THE SAME FORMULA SHEET THAT WILL BE PROVIDED THE DAY OF THE EXAM. IF IT'S NOT ON THE FORMULA SHEET, YOU MUST MEMORIZE IT!

Below is a list of which review assignments are DUE on the given days. Each assignment will be checked at the end of the period. An answer key will be provided to you to check and make corrections. If you are absent, all work must still be completed and shown on the day you return to class to get credit

Due at the END of the period on:

Monday, June 10th	Pages 4 – 7
Tuesday, June 11th	Pages 8 – 10
Wednesday, June 12th	Pages 11 – 14

Formula Sheet Statistics Honors Semester II

Descriptive Statistics

$$\bar{x} = \frac{\sum x_i}{n} \qquad s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1} \qquad s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$$

$IQR = Q_3 - Q_1$
$< Q_1 - 1.5 \cdot IQR$
$> Q_3 + 1.5 \cdot IQR$

$$z = \frac{x - \mu}{\sigma} \qquad normalcdf(\text{min}, \text{max}, \mu, \sigma) \qquad invnorm(percentile, \mu, \sigma)$$

Linear Regression

LSRL: $\hat{y} = a + bx$ **slope:** $b = r \cdot \frac{s_y}{s_x}$ **residual:** $y - \hat{y}$

y - intercept: $a = \bar{y} - b\bar{x}$

correlation coefficient: $r = \frac{\sum\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)}{n-1}$ or $r = \frac{1}{n-1}\sum\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)$

Probability

$${}_nC_r = \frac{n!}{(n-r)!r!} \qquad {}_nP_r = \frac{n!}{(n-r)!}$$

$P(A \text{ and } B) = P(A) \cdot P(B)$

$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

$P(B \setminus A) = \frac{P(A \text{ and } B)}{P(A)}$

Probability Density Function
$\mu = \sum x_i \cdot p_i$

$\sigma^2 = \sum(x_i - \mu)^2 \cdot p_i$

Rules for Mean and Variance

$\mu_{ax} = a \bullet \mu_x$

$\mu_{x+b} = \mu_x + b$

$\mu_{ax+b} = a \bullet \mu_x + b$

$\sigma_{ax} = a \bullet \sigma_x$

$\sigma_{x+b} = \sigma_x$

$\sigma_{ax+b} = a \bullet \sigma_x$

$\mu_{x+y} = \mu_x + \mu_y$

$\mu_{x-y} = \mu_x - \mu_y$

$\sigma^2_{x+y} = \sigma^2_x + \sigma^2_y$

$\sigma^2_{x-y} = \sigma^2_x + \sigma^2_y$

Binomial Distribution $B(n,P)$

$P(X=k) = {}_nC_r \bullet p^k (1-p)^{n-k}$

$\mu = np$

$\sigma = \sqrt{np(1-p)}$

calculator commands:

binompdf(n, p, k)

binomcdf$(n, p,$ maximum value$)$

Sampling Distribution of Proportions: $\mu_{\hat{p}} = p$ $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$

Sampling Distribution of Means: $\mu_{\bar{x}} = \mu$ $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

I. MULTIPLE CHOICE: **SHOW ALL WORK FOR FULL CREDIT in the space provided.**

1.) The Bureau of Labor Statistics announced last month it interviewed a sample of 50,000 households; **4.5%** of these households were unemployed. 4.5% is a …

A) population B) parameter C) sample D) statistic E) mean

2.) Choose the correct word from the work bank:

population	sample	experiment	treatment	simple random sample (SRS)

A) A specific experimental condition imposed on the subjects is called a(n): ________________________

B) The entire group of individuals that we're trying to infer information about: _____________________

C) This deliberately imposes some treatment on individuals in order to observe their response: _________

D) T allows all individuals from the population an equal chance of being chosen: ___________________

E) A part of the population that is actually examined to gather information about: ___________________

3.) The Census Bureau interviewed a sample of 40 households in the Cleveland metropolitan area to learn what percent of their spending goes toward housing. Due to the Central Limit Theorem, the sampling distribution of the sample mean $(\overline{x})$ has a distribution that is ….

A) Skewed to the right

B) Approximately normal

C) Skewed to the left

D) Not enough information given

4.) Your Statistics Honor's teacher notes that there is a 40% chance of a student earning an A average as a final grade. What is the expected number of A's that will be earned in a class of 20 students?

A) 12

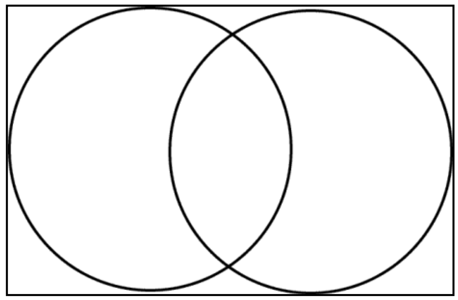
B) 4

C) 8

D) 20

5.) A poll was taken of students on the language(s) studied in school. 65% of the students studied Spanish. 8% of the students studied Spanish and French and 10% studied French only. What is the probability that a person at random studied Spanish given they studied French?

A) 0.08 B) 0.4444 C) 2.25 D) 0.5556



6.) Drew is trying to hit as many homeruns as possible during the Baseball Championship game. The probability of him hitting a homerun on any one try is 1/3. What is the probability that Reid hits exactly 2 homeruns in 6 tries?

A) 0

B) 0.3333

C) 0.6708

D) 0.3292

7.) If two 4-sided dice are rolled. What is the probability of rolling a sum of a 2 **or** a 5?

A) 0.3125

B) 0.25

C) 0.6875

D) 0.1389

8.) Travis just picked a card 2 times (with replacement). What is the probability that he picked a king and queen.

A) 0.0044

B) 0.0059

C) 0.0060

D) 0.0045

9.) In a population of families, the number of cell phones per household owned is a random variable X with the given probability distribution:

X	0	1	2	3	4
$P(X)$	0.1	0.2	0.1	0.5	0.1

What is the probability that the random variable X is greater than or equal to 3?

A) 0.6

B) 0.5

C) 0.4

D) 0.3

10.) All the values in a data set are increased by 4, what happens to the "new" mean?

A) multiplied by 4

B) stays the same

C) decreases by 4

D) increases by 4

11.) All the values in a data set are multiplied by 2, what happens to the "new" standard deviation?

A) multiplied by 2

B) increases by 2

C) decreases by 2

D) stays the same

12.) All the values in a data set are multiplied by 4, what happens to the "new" mean?

A) multiplied by 4

B) stays the same

C) decreases by 4

D) increases by 4

13.) All the values in a data set are increased by 2, what happens to the "new" standard deviation?

A) multiplied by 2

B) increases by 2

C) decreases by 2

D) stays the same

14.) A scatterplot of **<u>log y versus log x</u>** looks approximately like a positively sloping straight line. We may conclude that the model that best represents the data is:

A) exponential

B) linear

C) power

D) the answer cannot be determined from the information given.

15.) Student council advisors asked a sample of 150 high school students if they enjoyed their prom banquet hall. Suppose that 75% all EBHS high school students enjoyed the prom banquet hall. If $\hat{p}$ is the proportion of the sample who enjoyed their prom banquet hall, what is the mean o the sampling distribution of $\hat{p}$?

A) 112.5

B) 0.75

C) 0.25

D) 0.65

16.) You must choose a simple random sample (SRS) of 6 students from a class of 20 students.

a.) How would you label the students? ______________________________

b.) Starting at line 115, select 6 students from your SRS. Write down the numbers that represent the 6 students in your sample: ______________________________

17.) When choosing a simple random sample, name 3 different ways to choose the individuals.

1) ______________________________

2) ______________________________

3) ______________________________

II. OPEN-ENDED: SHOW ALL WORK FOR FULL CREDIT in the space provided.

1.) Television executives and companies who advertise on TV are interested in how many viewers watch particular television shows. According the 2005 Nielsen ratings, *Survivor: Guatemala* was one of the most-watched television shows in the United States during every week that it aired. Suppose that the true proportion of U.S. adults who watched *Survivor: Guatemala* is $p = 0.37$. We took a simple random sample of 100 adults.

(a) If $\hat{p}$ is the proportion of the sample who watch Survivor: Guatemala, what is the mean of the sampling distribution of $\hat{p}$?

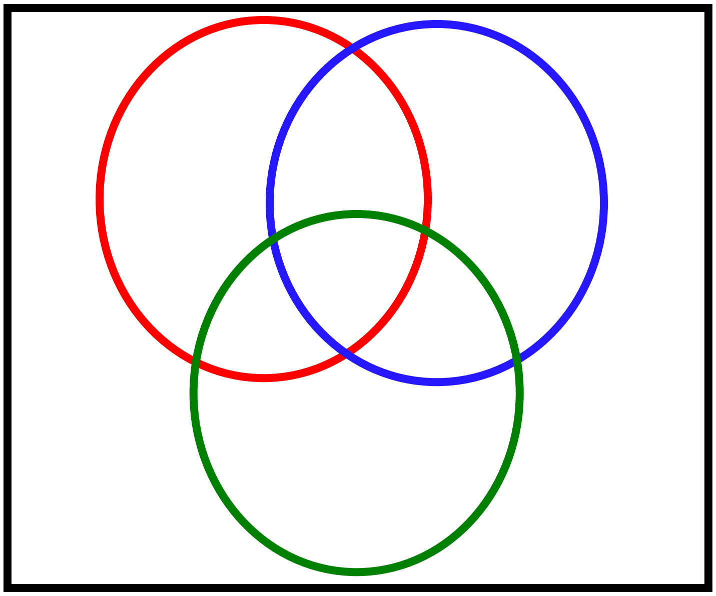
(b) What is the standard deviation o the sampling distribution of $\hat{p}$?

Include justification for why you can use the formula for standard deviation of the sampling distribution of $\hat{p}$ *in this setting.*

(c) **JUSTIFY** that you can use a normal approximation for the sampling distribution of $\hat{p}$.

(d) Find the probability that a sample of size n = 100 yields a $\hat{p}$ ***within plus or minus 1 percentage point*** of the true value p . *Draw, label, and shade, a normal curve.*

2.) Below is a Venn Diagram showing the probability that a randomly chosen student travels to the following destinations during summer vacation: Beach (B), Mountains (M), and Disney (D)



Given:

60% travel to the beach
10% travel to the mountains
15% travel to Disney

In addition…

1% of students travel to all three places
50% travel to beach only
2% travel to mountains & beach
2% travel to Disney only

(a) Find probability that a student travels to none of these destinations.

(b) Find the probability that a student travels to Disney and the beach.

(c) Find the probability that a student does not travel to the mountains.

(d) Find the probability that a student travels to Disney given that he/she travels to the beach.

(e) Find the probability that a student travels to the beach given that he/she travels to Disney.

3.) *How does the speed you are traveling in a car influence the distance you need to come to a complete stop?*
A statistics class gathered data to answer this question. The table below shows the speed *(in miles per hour = mph)* and the distance *(in feet)* needed to come to a complete stop when the brake was applied.

Speed (mph)	Distance (feet)
6	1.42
9	4.92
19	18.00
30	44.75
32	52.08
40	84.00
48	110.33

What is the explanatory variable? ______________________________

What is the response variable? ________________________________

(a) Determine if the above data is a ***linear,*** an ***exponential,*** or a ***power function***. Explain.
SHOW ALL WORK FOR FULL CREDIT.

(b) Write the equation for the (***linear transformed) LSRL***.

(c) Write the equation for the ***Non-Linear Regression function*** *(if the data is not linear)*.

(d) What does the ***regression equation*** predict a speed of 55 mph?

4.) A statistics' student, Tyler, opened a bag of M&Ms, dumped them out, and ate all the ones with the M&M on top. When he finished, he put the remaining 30 M&Ms back in the bag and repeated the same process over and over until all the M&Ms were gone. Below is a table showing the number of M&Ms remaining at the end of each "course."

Course	M&M'S remaining
1	30
2	13
3	10
4	3
5	2
6	1

What is the explanatory variable? ________________________________

What is the response variable? __________________________________

(a) Determine if the above data is a ***linear,*** an ***exponential,*** or a ***power function***. Explain. SHOW ALL WORK FOR FULL CREDIT.

(b) Write the equation for the ***(linear transformed) LSRL***.

(c) Write the equation for the ***Non-Linear Regression function*** *(if the data is not linear)*.

(d) What does the ***regression equation*** predict for the 5^{th} course?

5.) Given: number of sides for each of the two dice.

Dice 1: 6-sided ***Dice 2: 6-sided***

(a) List all of the possible ***outcomes*** if you roll the first die.

(b) List all of the possible ***outcomes*** for the second die.

(c) Draw a ***tree diagram*** or write the ***sample space*** for rolling the two dice at the same time.

Find the ***probability*** *for each of the following:*

(d) The sum of the dice being greater than 9: ______________________________

(e) Rolling two sixes or rolling two threes: __________________________________

(f) The sum of the dice being between 5 and 11: ____________________________

(g) The sum of the two dice being less than 7:_______________________________

(h) Rolling a 4 on the 1st roll and an 6 on the 2nd roll: _________________________

(i) Both are even numbers:____

(j) Rolling a 3 on the 1st roll and a 2 on the 2nd roll: ____________________________

(k) The sum of the two dice being 10: ______________________________________

*Now, roll both dice together 20 times. Record the **outcomes** in the table below.*

Event	Die 1	Die 2	Sum of both Dice
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			

i.) What was the most common ***outcome***? Why do you think this is?

ii.) What ***outcomes*** were certain? Why?

iii.) What ***outcomes*** were impossible? Why?

6.) Complete the linear transformation on the random variable X.

Find the **mean** and **standard deviation** for each.

						μ	σ
a) *X*	*0*	*3*	*5*	*7*	*9*		
b) $X + 2$							
c) $X - 8$							
d) $-3X$							
e) $2X-4$							

X and Y are independent random variables:

GIVEN: $\mu_X = 3$ $\mu_Y = 7$ $\sigma_X = 4$ $\sigma_Y = 9$

(A) Suppose all the data values for **random variable X were multiplied by 4,** find the following of the new data:

i) Mean

ii) Standard deviation

iii) Variance

(B) Suppose all the data values for **random variable X were increased by 16,** find the following of the new data:

i) Mean

ii) Standard deviation

iii) Variance

(C) Suppose all the data values for **random variable Y were divided by 2 and subtracted by 6,** find the following of the new data:

i) Mean

ii) Standard deviation

iii) Variance

(D) Find the mean of:

i) X – Y

(E) Find the standard deviation of:

i) X – Y