Statistics 300 : Fall 2016 (Mon/Wed class)

Instructor : L. C. Larsen

Name (print): ______________________________

Los Rios ID#: ______________________________

Signature: ______________________________

Exam : Unit 1

Time allowed : 2 hours and 5 minutes

Resources allowed:

== > Textbook (Author: Nividi and Monk)
== > Notes/helps written by the student
== > Quiz and exam solutions written by instructor
== > Quiz and exam solutions written by the student
== > Calculator/laptop/tablet of choice (no outside contacts)
== > Instructor at 916-346-6324

Resources not allowed:

== > Consultants other than the instructor
== > No phones, unless used as a calculator only

1. Use the data for the sample represented in the following table to answer parts (a) through (e).

(1 point, 1 minute)
(a) What is the probability that a randomly selected person will have "High School" as their highest level of schooling?

Highest Level of Schooling	Income Quartile First	Second	Third	Fourth	Total
Graduate	100	200	400	500	1200
College	100	200	500	400	1200
High School	200	400	300	300	1200
Elementary	500	300	200	200	1200
None	700	300	100	100	1200
Total	1600	1400	1500	1500	6000

(3 points, 3 minutes)
(b) What is the probability that a person randomly selected from this sample will have income in the "Second Quartile" given that the person has "Graduate" level schooling?

(3 points, 3 minutes)
(c) What is the probability that a person randomly selected from this sample will have "high school" education and income <u>greater than</u> the "Second Quartile"?

(3 points, 3 minutes)
(d) What is the probability that a person randomly selected from this sample will have income in the "First Quartile" or "High School" as their highest education level?

(3 points, 3 minutes)
(e) What is the probability that a person randomly selected from this sample will not be in the "First Quartile" for income and also not have "Elementary" level schooling?

2. (5 points; 5 minutes)
(a) Daily maximum temperatures in the month of May have a bell-shaped distribution with a mean of 86 ^{o}F and a standard deviation of 4 ^{o}F, approximately <u>what percent</u> of the temperatures in May next year should be between 78 ^{o}F and 94 ^{o}F? To get any points for this problem, you must show how you got your answer!

Answer:

(3 points; 3 minutes)
(b) For the bell-shaped distribution in part (a), are any of the values between 86 ^{o}F and 90 ^{o}F unusual?

YES **Why?** ______________________
or ______________________
NO ______________________

(9 points; 8 minutes)
3. Complete the columns in the "Frequency Distribution" table using the data values given below, and answer the two questions below the table.

Frequency Distribution

Class Limits Lower	Upper	Tally	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
0.0	5.0					
>5.0	10.0					
>10.0	15.0					

Class Midpoint

Data:					
	14.86	1.08	9.25	2.25	10.83
	9.16	14.41	12.27		

Class Width

What is the frequency for Class #2 ? ____________

What is the upper limit for Class #1 ? ____________

(4 points; 5 minutes)

4. **For each of the underlined segments in the situations below, select the appropriate term from the list provided and write it in the blank next to the description or situation. Choose the term that is best connected to the underlined text in the description or situation.**

Terms:			
	1. randomization	**5. placebo**	**9. parameter**
	2. replication	**6. block**	**10. statistic**
	3. confounding	**7. experimental unit**	**11. population**
	4. blinding	**8. treatment**	

(a.) An experiment is done to estimate **the average of the responses of all autistic children** to large doses of vitamins. The study involved 600 autistic children in each of 5 age groups. In each age group, 200 children were given a pill with no vitamins, 200 a pill with the standard dose, and 200 a pill with a large dose. The children stayed in their family homes, and each family believed their child was receiving the "large dose". Cameras in each home recorded the behavior of each child which was scored on a "20 point scale" for "severity of autism". Conclusions were based on the difference between the average score of the large dose group and the average score of the standard dose group. The study could not control for the possible effects of unique factors in each household that may also affect autism. ____________________

(b.) An experiment is done to estimate the average of the responses of all autistic children to large doses of vitamins. The study involved 600 autistic children in each of 5 age groups. In each age group, 200 children were given a pill with no vitamins, 200 a pill with the standard dose, and 200 a pill with a large dose. Children stayed in their family homes, and each family believed their child was receiving the "large dose". Cameras in each home recorded the behavior of each child which was scored on a "20 point scale" for "severity of autism". Conclusions were based on the difference between the average score of the large dose group and the average score of the standard dose group. **The study could not control for the possible effects of unique factors in each household that may also affect autism.** ____________________

(c.) An experiment is done to estimate the average of the responses of all autistic children to large doses of vitamins. The study involved 600 autistic children in each of 5 age groups. In each age group, 200 children were given a pill with no vitamins, 200 a pill with the standard dose, and 200 a pill with a large dose. Children stayed in their family homes, and **each family believed their child was receiving the "large dose".** Cameras in each home recorded the behavior of each child which was scored on a "20 point scale" for "severity of autism". Conclusions were based on the difference between the average score of the large dose group and the average score of the standard dose group. The study could not control for the possible effects of unique factors in each household that may also affect autism. ____________________

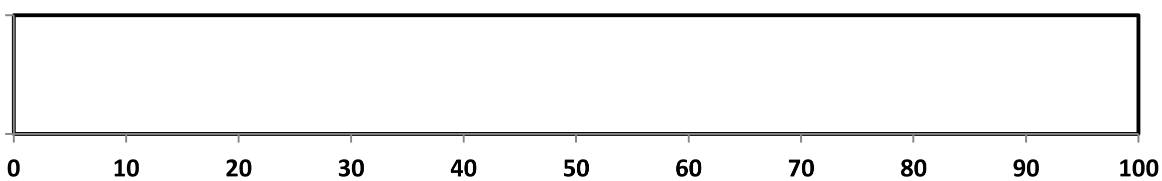
(d.) An experiment is done to estimate the average of the responses of all autistic children to large doses of vitamins. The study involved 600 autistic children in each of 5 age groups. In each age group, **200 children were given a pill with no vitamins, 200 a pill with the standard dose, and 200 a pill with a large dose**. Children stayed in their family homes, and each family believed their child was receiving the "large dose". Cameras in each home recorded the behavior of each child which was scored on a "20 point scale" for "severity of autism". Conclusions were based on the difference between the average score of the large dose group and the average score of the standard dose group. The study could not control for the possible effects of unique factors in each household that may also affect autism. ____________________

(4 points; 4 minutes)

5. **The US Department of Education released a report that said 73% of high-shool seniors in the 2014-2015 school year were able to graduate. In random samples of 20 seniors from the class of 2014-2015, the number that graduated follows a Binomial distribution. If you select a random of 20 high-school seniors from the many thousands in the class of 2014-2015, what is the probability that exactly 12 will be students who graduated?**

(5 points and 4 points; 10 minutes)

6. **(a) For the set of 47 values shown below in sorted order, prepare a Boxplot <u>inside</u> the rectangle that is above the number line.**



5	10	13	14	19	22	25	28	28	32
36	38	40	45	45	48	53	53	58	63
67	70	71	73	76	79	80	84	84	89
89	94	97	97	101	102	103	103	103	107
107	111	116	117	120	120	120			

(b) For the set of 47 values shown above in sorted order, what percentile is represented by the value 103?

(8 points; 10 minutes)

7. Answer parts (a), (b), and (c).
Use the columns in the table in any way you wish to use them.

(a) Is this distribution "valid" ("proper") (circle "YES" or "NO")? **YES** **NO**

Explain why?: ______________________________

x	P(x)				
14	0.4				
33	0.3				
51	0.2				
62	0.1				

(b) Write the formulas for the mean, the variance, and the standard deviation of this discrete probability distribution.

$\mu =$ ______________________________

$\sigma^2 =$ ______________________________

$\sigma =$ ______________________________

(c) Write the values for the mean, the variance, and the standard deviation of this discrete probability distribution. The work above must show how you determined these values.

$\mu =$ ____________

$\sigma^2 =$ ____________

$\sigma =$ ____________

(4 points and 2 points; 5 minutes)

8. You have four M&Ms, one green, one blue, one red, and one yellow. Two of the M&Ms will be picked, one at a time, without replacement. List the sample space for this procedure. Example: {RY} for first "red" and second "yellow".

If the three M&Ms are picked in a random order without replacement, what is the probability that the red M&M will be picked but the blue M&M will not be picked?

(3 points; 3 minutes)

9. A lucky person is picked at random as the Grand Prize winner in a lottery. The winner will pick one of 4 cars, one of 5 houses, and one of 6 TVs. How many different ways can the winner pick a combination of one car, one house, and one TV?

(3 points; 3 minutes)

10. A person has eight siblings (brothers and sisters). The person also has a picture of each sibling. If the pictures are to be hung at equal height along one wall of the living room, how many different ways can the pictures be arranged?

(5 points; 6 minutes)

11. A creator of carnival rides wants to use the 10th and 90th percentiles of the heights all 10-year-old children to help in designing a new "roller coaster" ride. A random sample of 450 10-year-old children is selected and their heights are measured. The 10th and 90th percentiles of these heights were 48 inches and 56 inches.

Use the information in the "story" to answer the following:

(a) What is the population of interest to the creator of carnival rides?

(b) The creator of carnival rides wanted to know the value of what population parameter(s)?

(c) The creator of carnival rides determined the value(s) of what statistic(s)?

(d) What were the value(s) that the creator of carnival rides determined for the statistic(s) in part (c)?

(e) Was a sample or a census used, and how did you decide your answer?

(12 points; 8 minutes)

12. Use the data below to determine the value of each statistic. Write an expression for each statistic or describe how it is calculated in principle (do NOT describe how to use the calculator to determine the result)**.**

Data
66
51
72
80
51
59
48
66

	Expression or Description	Value of statistic
median		
variance		
mode		
range		
mean		
standard deviation		

(3 points; 3 minutes)

13. Last year, a small High School had 60 students, of which 39 were girls (G) and 21 were boys (B). If 5 of the 60 students are selected at random (without replacement), what is the probability that the exact order will be {G,G, B, B, G} ?
(Show your work.)

(6 points; 8 minutes)

14. Last year, the dropout rate for all High School students was 18.5%. Use the "relative frequency" approach to probability and assume that decisions to drop out of school are independent. Then, decide if it would be unusual for a random sample of 1077 students this year to have 250 that will drop out and 827 that will not drop out.
(Hint: This problem will require the "binomial distribution" and something else to assess "unusual".)

(6 points; 6 minutes)

15. **Circle the best answer for each situation.**

Official surveyors call 4,000 people who are randomly selected from those that used Medical health services in the last 10 years. Of the 4000 people, 22% agree to take the the survey, and these are asked if they know their estate must repay the costs when they die.

Simple Random	Systemmatic
Stratified Random	Cluster
Convenience	Census

A government program auditor reviews Medical case files. Files for every Medical provider (hospital, doctor, pharmacy, etc.) willl be reviewed. For each provider, the auditor will examine every 200^{th} file.

Simple Random	Systemmatic
Stratified Random	Cluster
Convenience	Census

The DMV provides a list of all registered vehices to a research group for study. The researchers attach a random value between zero and one to each vehicle's record. The 200 records with the smallest random values are selected as the sample to study.

Simple Random	Systemmatic
Stratified Random	Cluster
Convenience	Census

16. **Circle the best answer for each situation.**

The Department of Corrections (Prisons) selects a group of 800 prisoners released in 2010 and follows their lives for the next 30 years to find out what types of decisions decrease the percent that return to prison at a later time.

observational retrospective	observational cross-sectional
observational prospective	experiment

The Department of Corrections randomly selects 5000 prisoners in 2013 and conducts a detailed examination of their lives before they went to prison, to learn about factors that may lead to criminal behavior and imprisonment.

observational retrospective	observational cross-sectional
observational prospective	experiment

The Department of Corrections releases 10 groups of 40 prisoners. In each group, the prisoners are as much alike as possible (age, gender, criminal record, ethnicity, religion, etc.). In each group, half go into the army and the other half does not, so the effect of army service on post-release life can be evaluated.

observational retrospective	observational cross-sectional
observational prospective	experiment