

Statistics 300 : Fall 2017 (Saturday class)

Instructor : L. C. Larsen

Name (print): _____

Los Rios ID#: _____

Signature: _____

Exam : Unit 1

Time allowed : 2 hours and 10 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				Total
	First	Second	Third	Fourth	
Graduate	87	121	437	591	1236
College	191	285	446	580	1502
High School	274	203	328	267	1072
Elementary	239	254	151	139	783
None	772	389	131	62	1354
Total	1563	1252	1493	1639	5947

(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 "None" education or "Graduate" education?

(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 a person in the "First Quartile" for income who also has "Graduate" level schooling?

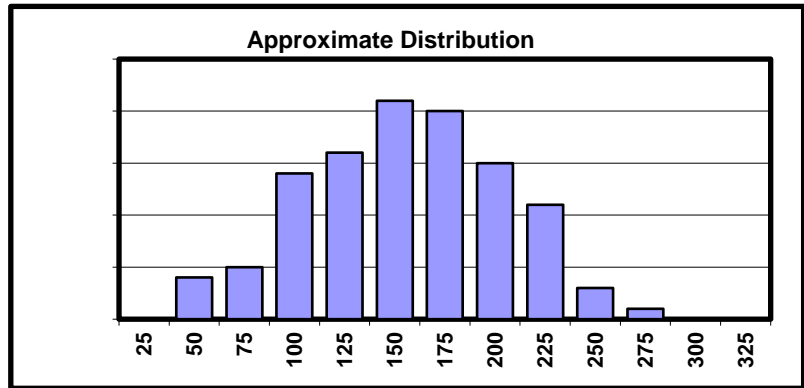
(4 points and 3 points; 8 minutes)

2. (a) Use the information provided to determine the lower number and the higher number between which you expect to find almost all (~99.7%) of the data.

Mean = 144.6
Standard Deviation = 49.8

Lower Number = _____

Higher Number = _____



(b) Would a value of 250 or more from this distribution be considered "unusual" ? [must show why]

Circle: YES of NO

(10 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
6.5 7.0					
>7.0 7.5					
>7.5 8.0					

Class Midpoint	Class Boundary

Data:	7.7	7.3	7.4	8.0	7.7
	7.7	6.5	7.6	7.8	7.5

Class Width

What is the frequency for Class #3 ? _____

What is the lower limit for Class #2 ? _____

(4 points; 6 minutes)

4. For the study described below, select the appropriate statistical terms from the list provided and write them in the blanks, choose the term that is best connected to the underlined text.

Terms:	randomization	placebo
	replication	block
	confounding	experimental unit
	blinding	treatment

a. _____

Engineers test the strength of concrete for use by homeowners. Forty homeowners take part in a the study. Twenty homeowners are randomly assigned to use "concrete mix #1" and the other twenty homeowners use "concrete mix #2", which has extra ingredients. Neither the homeowners nor the engineers know who is using which mix. Each homeowner mixes the concrete and pours a square slab 10 feet by 10 feet and 4 inches thick. No effort is made to control factors like daily temperature that could affect the strength of the concrete. The experiment shows that the extra ingredients improve the concrete strength by a large amount.

b. _____

Engineers test the strength of concrete for use by homeowners. Forty homeowners take part in a the study. Twenty homeowners are randomly assigned to use "concrete mix #1" and the other twenty homeowners use "concrete mix #2", which has extra ingredients. Neither the homeowners nor the engineers know who is using which mix. Each homeowner mixes the concrete and pours a square slab 10 feet by 10 feet and 4 inches thick. No effort is made to control factors like daily temperature that could affect the strength of the concrete. The experiment shows that the extra ingredients improve the concrete strength by a large amount.

c. _____

Engineers test the strength of concrete for use by homeowners. Forty homeowners take part in a the study. Twenty homeowners are randomly assigned to use "concrete mix #1" and the other twenty homeowners use "concrete mix #2", which has extra ingredients. Neither the homeowners nor the engineers know who is using which mix. Each homeowner mixes the concrete and pours a square slab 10 feet by 10 feet and 4 inches thick. No effort is made to control factors like daily temperature that could affect the strength of the concrete. The experiment shows that the extra ingredients improve the concrete strength by a large amount.

d. _____

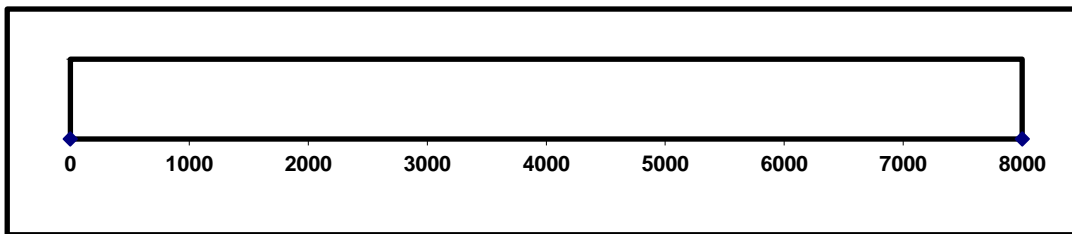
Engineers test the strength of concrete for use by homeowners. Forty homeowners take part in a the study. Twenty homeowners are randomly assigned to use "concrete mix #1" and the other twenty homeowners use "concrete mix #2", which has extra ingredients. Neither the homeowners nor the engineers know who is using which mix. Each homeowner mixes the concrete and pours a square slab 10 feet by 10 feet and 4 inches thick. No effort is made to control factors like daily temperature that could affect the strength of the concrete. The experiment shows that the extra ingredients improve the concrete strength by a large amount.

(4 points and 5 points; 10 minutes)

5. Use the dataset at the bottom of this page to answer parts (a), (b) and (c).
There are 130 values in the dataset, in rows of 10, sorted from the smallest at the top to the largest at the bottom.

(a) What percentile is represented by the value 2136 ?

- (c) Using the number line below, make a Boxplot to represent the distribution of the dataset.
Use a simple "5-number summary" to prepare this boxplot.



100	107	144	149	170	193	200	226	263	294
322	340	344	363	372	385	402	440	475	514
523	545	584	599	627	657	669	697	715	740
752	770	778	830	863	963	988	1015	1042	1070
1073	1087	1169	1223	1278	1335	1365	1430	1484	1521
1551	1564	1603	1613	1657	1727	1791	1798	1877	1904
1934	1948	1948	1954	1962	1966	1970	1980	1985	1989
1994	2004	2012	2020	2029	2032	2041	2047	2057	2063
2063	2068	2070	2080	2090	2098	2102	2112	2121	2125
2128	2131	2136	2137	2140	2142	2146	2155	2161	2161
2165	2174	2183	2193	2195	2204	2210	2216	2219	2222
2228	2229	2229	2231	2237	2243	2251	2255	2264	2266
2272	2279	2286	2289	2293	2301	2310	2315	2318	2324

(4 points and 2 points; 5 minutes)

- 6.** 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, meaning the first is "red" and the second is "yellow".

If the two M&Ms are picked in a random order and 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)

- 7.** A lucky person is picked at random as the Grand Prize winner in a lottery. The winner will pick one of four cars, one of five houses, and one of six TVs. How many different ways can the winner pick a combination of one car, one house, and one TV?

(3 points; 3 minutes)

- 8.** 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)

9. 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(s) 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?

(14 points; 8 minutes)

10. 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	Expression or Description	Value of statistic
42.0 40.1 41.4 40.5 41.3 40.0 41.3 40.2 41.0	median	
	variance	
	mode	
	range	
	mean	
	standard deviation	
	mid-range	

(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.)

(3 points; 3 minutes)

- 14.** A final exam in statistics must have 10 out of 26 possible problems. If the professor decides to choose the 10 problems at random and then arrange them in a random order, how many different ways could the test turn out?

(3 points; 3 minutes)

- 15.** A statistics exam will have 2 different versions so students will be discouraged from trying to cheat. The class has 36 students and will be divided into two groups of 18 each. Group 1 will get test #1 and group 2 will get test #2. How many ways could the professor divide the class into two groups of 18?

(3 points; 3 minutes)

- 16.** From a very large population (millions of values), what is the probability that two random values (treat them as independent) will both be between the 30th and the 62nd percentiles (P_{30} and P_{60}).

(3 points; 3 minutes)

17. Circle the best answer for each situation.

Official surveyers 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 200th file.

Simple Random	Systemmatic
Stratified Random	Cluster
Convenience	Census

The DMV provides a list of all registered vehicles 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

18. Circle the correct choice in each box in relation to the underlined text.

	Are the data ... ?	Are the data ... ?
a. A grocery store sells <u>eggs that are small, medium, large , and extra large</u> . On September 11, 2001, the store sold 127 pounds of eggs.	Qualitative Quantitative and Discrete Quantitative and continuous	Nominal Interval Ordinal Ratio
a. A grocery store sells eggs that are small, medium, large , and extra large. On <u>September 11, 2001</u> , the store sold 127 pounds of eggs.	Qualitative Quantitative and Discrete Quantitative and continuous	Nominal Interval Ordinal Ratio
a. A grocery store sells eggs that are small, medium, large , and extra large. On September 11, 2001, the store sold <u>127 pounds of eggs</u> .	Qualitative Quantitative and Discrete Quantitative and continuous	Nominal Interval Ordinal Ratio