Statistics 300 : Fall 2008

Instructor : L. C. Larsen

Student name & ID#: _____

Student signature: _____

Exam : Unit 1

Time allowed : 2 hours and 5 minutes

Exam window: 9/26, 9/29, 9/30, 2008.

Resources allowed:

== > Open textbook (Author: Triola)
== > Open notes/helps written by the student
== > Quiz and exam solutions written by instructor
== > Quiz and exam solutions written by the student
== > Calculator/laptop of choice
== > Instructor at 916-346-6324

Resources not allowed:

== > Consultants

Statistics 300 **Name:** _____ **Spring 2009**

Instructor: Lawrence C. Larsen **Tue/Thu 7:00-9:05 p.m.**

Exam #1

(7 points; 8 minutes)

1. **A small ferry boat carries people and cars across a river. The boat can carry 10 people and 2 cars. Five cars are waiting to cross the river -- 2 are Red, 2 are Green, and one is Blue. The car owners all claim to have arrived at the same time, so the ferry boat operator decides to pick one at random to get on the boat first and another at random to get on the boat second.**

   (a) List the sample space for the boat operators' procedure (e.g. $\{G_2,G_1\}$)

   (b) List the possible color sequences for the first two cars (e.g. {G,G}) and their probabilities

(7 points; 7 minutes)

2. **Given: X ~ Binomial(n = 2000, p = 0.72) and Y ~ Binomial(n = 800, p = 0.44)**
   Which would be more unusual, X = 1392 or Y = 378 ?

Statistics 300 **Name:** _____ **Spring 2009**
Instructor: Lawrence C. Larsen **Tue/Thu 7:00-9:05 p.m.**

Exam #1

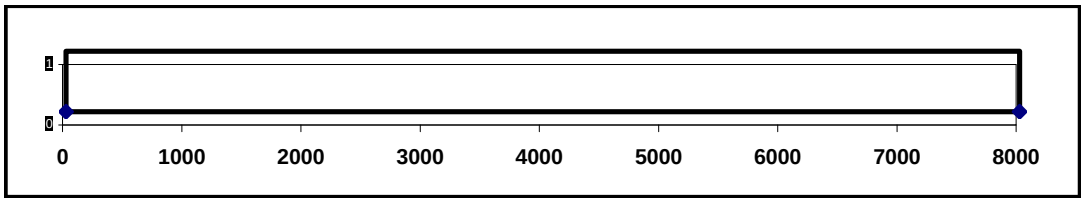
(12 points; 10 minutes)

3. Use the dataset at the bottom of this page to answer parts (a), (b) and (c). There are 199 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 1087 ?

(b) What is the value of the 75^{th} percentile, P_{75} ?

(c) Using the number line below, make a Boxplot to represent the distribution of the dataset.



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
2329	2337	2342	2345	2348	2354	2357	2360	2366	2368
2377	2382	2385	2391	2399	2404	2407	2471	2599	2769
2785	2895	2945	2969	3156	3296	3351	3531	3655	3819
3850	3852	3930	4026	4051	4231	4371	4509	4665	4795
4814	4836	4961	5023	5071	5257	5305	5416	5531	5593
5656	5670	5821	5941	6126	6308	6396	6492	6593	6716
6734	6773	6816	6902	7024	7183	7323	7431	7456	

Statistics 300 **Name:** _____ **Spring 2009**
Instructor: Lawrence C. Larsen **Tue/Thu 7:00-9:05 p.m.**

Exam #1

(3 points; 5 minutes)

4. For each of the following "sampling" situations circle RANDOM, STRATIFIED, SYSTEMATIC, CLUSTER, CONVENIENCE, or CENSUS as the type of sampling conducted.

a. An MTV program shows music videos and asks all viewers to call a free 800 number to rate each video on a scale of 1 to 10.

Simple Random	Systematic
Stratified Random	Cluster
Convenience	Census

b. A cable TV company rates the popularity of TV shows among its customers by constantly tracking the channel to which each of its cable connect boxes is tuned.

Simple Random	Systematic
Stratified Random	Cluster
Convenience	Census

c. A research company rates the popularity of TV shows by taking random samples of all Americans in each of 5 income groups withing each of 5 age groups.

Simple Random	Systematic
Stratified Random	Cluster
Convenience	Census

(3 points; 5 minutes)

5. For each of the following studies circle all of the characteristics that are appropriate.

a. The Department of Corrections (Prisons) selects a group of 5000 prisoners released in 2001 and studies key characteristics of their lives to find out what types of decisions decrease the percent that return to prison at a later time.

retrospective	observational study
cross-sectional	
	experiment
prospective	

b. The Department of Corrections releases a group of 400 prisoners who share alike in key characteristics (race, education, family history, type of crime, etc.). Half of the 400 go into the army and the others do not, so the effect of military service can be studied.

retrospective	observational study
cross-sectional	
	experiment
prospective	

c. The Department of Corrections randomly selects 5000 prisoners 2006 and conducts a detailed health exam on each one in order to study the present state of health in the prison population at that time.

retrospective	observational study
cross-sectional	
	experiment
prospective	

Statistics 300 **Name:** _____ **Spring 2009**
Instructor: Lawrence C. Larsen **Tue/Thu 7:00-9:05 p.m.**

Exam #1

(8 points; 5 minutes)

6. For each of the discrete probability distributions below, calculate the mean, variance, and standard deviation.

(a)

X	P(X)			
7	0.23			
19	0.66			
24	0.11			

(b)

X	P(X)			
0	0.136			
1	0.279			
2	0.423			

(8 points; 5 minutes)

7. Use the small set of data below to complete the identified parts the frequency distribution.

Class Limits		Tally	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
Lower	Upper					
10	20					
30	40					
50	60					
70	80					

Data:	67	24	33	12	68	38
	42	57	46	56	24	70
	20	59				

_____ = the class width.

_____ = the lower class limit for class #2.

_____ = the frequency of class #3.

(4 points; 6 minutes)

8. The California Association of Realtors wants to estimate the percentage of all single-family homes in the state that have more than two bathrooms. For this purpose, the Association takes a random sample of 800 addresses for single-family homes and determines the number of bathrooms in each of these homes. There were 96 homes with more than two bathrooms, so the estimate for all single family homes in CA is 12%.

(a) What is the population of interest in this situation?

_____

_____

(b) What is the parameter of interest in this situation?

_____

_____

(c) What statistic was used in this situation?

_____

_____

(d) Was a sample or a census used for this study, and why did you choose your answer?

_____

_____

Statistics 300 **Name:** _____ **Spring 2009**
Instructor: Lawrence C. Larsen **Tue/Thu 7:00-9:05 p.m.**

Exam #1

(14 points; 10 minutes)

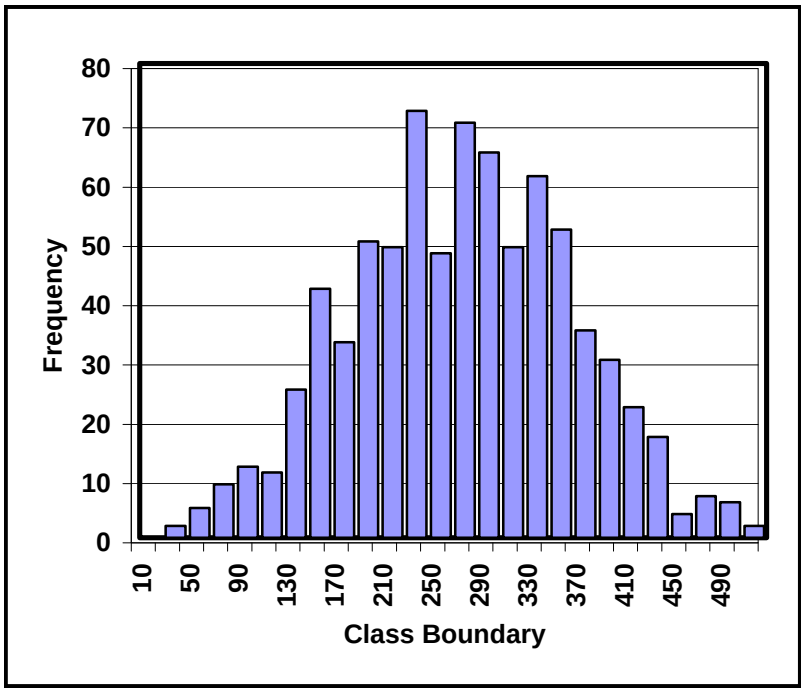
9. For the sample of data given below, provide the formula (expression) or description of calculation (not how to use the calculator) **for each statistic listed and also provide the value of each statistic. The mean and standard deviation must be calculated using your calculator's "statistics mode".**

Data:
116
120
116
113
117
119
110
116
118
117
110
117

Statistic	Formula / Description	Value
Mid-range		
Mode		
Mean		
Median		
Range		
Standard deviation		
Variance		

(6 points; 6 minutes)

9. A sample of 1000 data values were collected in a random sample. A graph of the distribution is shown. The 16th percentile (P_{16}) was 160 and the 84th percentile (P_{84}) was 340. Use all this information to estimate (not a wild guess) the standard deviation of the data.



(3 points; 3 minutes)

10. A final exam in statistics must have 10 out of 16 possible problems. If the professor decides to choose the 10 problems at random arrange them in a random order, how many different ways could the test turn out?

(3 points; 3 minutes)

11. A statistics exam will have 2 different versions so students will be discouraged from trying to cheat. If the class has 40 students and half will be assigned to each of the two versions, how many different ways could the professor divide the class into two groups?

(5 points; 4 minutes)

12. A different statistics professor likes multiple choice problems. That professor gives an exam that has 10 problems with 4 possible answers in each one. If a student decides to use the calculator's random number function to guess on each problem, what is the probability that the student will guess the correct answer on exactly 4 of the 10 problems?

(5 points; 4 minutes)

13. Another statistics professor also likes multiple choice problems. That professor gives an exam that has 6 problems with 5 possible answers in each one. If a student decides to use the calculator's random number function to guess on each problem, what is the probability that the student will guess the correct answer on at least one of the 6 problems?

(3 points; 3 minutes)

14. 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:	**1. randomization**	**5. placebo**
	2. replication	**6. block**
	3. confounding	**7. experimental unit**
	4. blinding	**8. treatment**

Best term

a. _____ **A total of 60 children were included in a study of a new medication. The study used 30 similar girls and 30 similar boys that already used the standard medication every day. In the study, 10 boys and 10 girls were given a "medication" that had no effect at all, 10 boys and 10 girls were given the standard medication, and 10 boys and 10 girls were given the new medication. So, each child received one of the three types of medication. Before the study began, each child was equally likely to be assigned to each one of the medications. To prevent "bias", neither the children nor the experimenters knew which medication each child was being given.**

b. _____ **A total of 60 children were included in a study of a new medication. There were 30 girls and 30 boys in the study, who were already using the standard medication every day. In the study, 10 boys and 10 girls were given a "medication" that had no effect at all, 10 boys and 10 girls were given the standard medication, and 10 boys and 10 girls were given the new medication. So, each child received one of the three types of medication. Before the study began, each child was equally likely to be assigned to each one of the medications. To prevent "bias", neither the children nor the experimenters knew which medication each child was being given.**

c. _____ **A total of 60 children were included in a study of a new medication. There were 30 girls and 30 boys in the study, who were already using the standard medication every day. In the study, 10 boys and 10 girls were given a "medication" that had no effect at all, 10 boys and 10 girls were given the standard medication, and 10 boys and 10 girls were given the new medication. So, each child received one of the three types of medication. Before the study began, each child was equally likely to be assigned to each one of the medications. To prevent "bias", neither the children nor the experimenters knew which medication each child was being given.**

Statistics 300 **Name:** _____ **Spring 2009**
Instructor: Lawrence C. Larsen **Tue/Thu 7:00-9:05 p.m.**

Exam #1

(3 points; 3 minutes)

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

	Are the data ... ?	Are the data ... ?
a. The total gallons of all the gasoline used by Americans to drive to work today.	**Qualitative** **Quantitative and Discrete** **Quantitative and continuous**	**Nominal** **Interval** **Ordinal** **Ratio**
b. The number of "subcompact", "compact", "mid-size", and "standard" cars used by Americans to drive to work today.	**Qualitative** **Quantitative and Discrete** **Quantitative and continuous**	**Nominal** **Interval** **Ordinal** **Ratio**
c. The total profit of all the gasoline companies that sell gasoline to Americans who drive to work today.	**Qualitative** **Quantitative and Discrete** **Quantitative and continuous**	**Nominal** **Interval** **Ordinal** **Ratio**

(3 points; 3 minutes)

16. A standard California license plate for a car has 4 numbers (digits) and 3 letters in the format "DLLLDDD". Each D can be a digit from 0 through 9 and each L can be any one of the 26 letters in our alphabet. How many standard license plates are possible?

(3 points; 3 minutes)

17. A bowl contains 20 jelly beans. Five are "Cherry", 8 are "Orange", 2 are "Lemon" and 5 are "Grape". What is the probability of getting the sequence "O,O,G" if 3 jelly beans are taken out of the bowl (and not put back in between picks)?

(3 points; 3 minutes)

18. Use the information in the table to answer the probability questions (a) - (c).

Number of Years at Current Job	Type of Position in Current Job: Entry Level	Regular Staff	Middle Manager	Executive Manager	Total
0 to 5 years	160	40	10	0	210
6 to 10 years	80	90	20	4	194
> 10 years	40	100	40	8	188
Total	280	230	70	12	592

(a) What is the probability that someone picked at random from the 592 individuals in this table will be someone who has been at their current job for 6 to10 years?

(b) What is the probability that someone picked at random from the 592 individuals in this table will be someone who has been at their current job for 6 to10 years given that they are a Middle Manager?

(c) What is the probability that someone picked at random from the 592 individuals in this table will be someone who has been at their current job for 6 to10 years or be a Regular Staff person?