Solution Statistics 300 Name: Fall 2011 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. Effect cars are waiting to cross the river $\frac{1}{2}$ are Red $\frac{1}{2}$ are Green, and $\frac{1}{2}$ are 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. $\{G_2,G_1\}$) R.R. BaR, GIR, G2B, BR. R261 61R2 G2R2 BR2 R.G. R_2G_2 G_1G_2 G_2G_1 BG_1 R. Gz R2B G, B G, B BG2 R, B $\begin{array}{cccc} & 7/2 & 7/2 \\ 7/2 & 7/2 & 7/2 \\ 7/2 & 7/2 & 7/2 \\ 7/2 &$ (b) List the possible color sequences for the first two cars (e.g. {G,G}) and their probabilities Probabilities BR RG BG RR RB (7 points; 7 minutes) 2. Given: X ~ Binomial(n = 2000, p = 0.27) and Y ~ Binomial(n = 800, p = 0.34) Which would be more unusual, X = 576 or Y = 246 ? $\mu_{y} = 800(0.34) = 272$ $\mu_{x} = 2000 (0.27) = np = 540$ Jy = npg = 800(.34)(.66) $\sigma_{x} = Mpg = [2000(.27)(.73)]$ = 13.40 = 19.85 $\overline{z}_{g} = \frac{y - \mu_{y}}{\sigma_{y}} = \frac{246 - 272}{13.40}$ $(\overline{z}_{x}) = -1.99$ is most $(\overline{z}_{y}) = -1.99$ $Z_{\chi} = \frac{\chi - \mu_{\chi}}{\sigma_{\chi}} = \frac{576 - 540}{19.85}$ Zy > = 1.81

Fall 2011 Tue/Thu 7:00-9:05 p.m.

Exam #1

is at location L, where $L = \left(\frac{k}{100}\right) N = \left(\frac{57}{100}\right) 170 = 98.9$ = 2146 Vound up to 97 location

(12 points; 10 minutes)

3. Use the dataset at the bottom of this page to answer parts (a), (b) and (c). There are 170 values in the dataset, in rows of 10, sorted from the smallest at the top to the largest at the bottom.

k = 90

Name:

(a) What percentile is represented by the value 2969 ? $k = \left(\frac{\# q}{100} \sqrt{100} \le \frac{2969}{100}\right) 100 = \left(\frac{153}{170}\right) 100$

(b) What is the value of the 57th percentile, P_{57} ?

= 2146

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

	с на с								_		
				1					۱ ۲		
					,				- 1		
		0	1000 20	00 3000	4000	5000	6000	7000	8000		
	L					6					
			P -	Q1 =	(locati	m 43)	= 116	9			
Dim	alt		125 -	00 -	Nacator	e 85 \$ 86	allage	d) - 2	ngy		
Por	por		P.50 =	QL =	Cicculion	DONTOO	UV	5-1			
	- 1		D =	Q2 -1	locatio	n (28)	= 23	15			
m	IN) · · · · ·	175	~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~					·		
	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 5	V 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	1
	2063	2068	2070	2080	2090	2098 477	¹⁴ 2102	2112	2121	2125	,th
100	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 P	2264	2266	
	2272	2279	2286	2289	2293	2301	2310	2315 0	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	2969	2969	3351	3531	3655	3819	
	3850	3852	3930	4026	4051	4231	4371	4509	4665	4795	
170	-									$\langle \ $	

Statistics 300

Name:

Instructor: Lawrence C. Larsen

Fall 2011 Tue/Thu 7:00-9:05 p.m.

Exam #1

Solutio

(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.



(3 points; 5 minutes)

- 5. For each of the following studies circle <u>all</u> of the characteristics that are appropriate.
 - a. The Department of Corrections (Prisons) selected a group of 5000 prisoners released in 2001. Over the next 10 years, the 5000 prisoners were studied to find out what types of decisions affect the percent that end up returning to prison during that time.
 - b. The Department of Corrections released a group of 400 prisoners who share many key characteristics (race, education, family history, type of crime, etc.). Half of the 400 are randomly selected to serve in the army while the others do not serve in the army. The The purpose was to study the effect of army service.
 - c. The Department of Corrections randomly selected 5000 prisoners in 2006 and conducted a detailed health exam on each one in order to study the present state of health in the prison population at that time.

I might drop this problem.



retrospective	observational
	study
cross-sectional	
	experiment
prospective	

Statistics 3	00	Name:
Instructor:	Lawrence C. Larsen	

Exam #1

(8 points; 5 minutes)

6. For each discrete probability distribution, calculate the mean, variance, and standard deviation.

			T	2 2	T 2
(a)	х	P(X)	$\mu = \Sigma (\chi \cdot P(x))$	$\sigma = \Sigma(x - \mu) \cdot P(x)$	0=0
	6	0.23	1.38	36.17	
	18	0.11	1.98	0,03	
	23	0.66	15.18	13.13	
	ZPX)	1.00	µ= 18,54	$\sigma^2 = 49.33$	$(\sigma = 7.02)$



(9 points; 6 minutes)

20

= the class width.

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

0.

Class	s Limits			Relative	Cumulative	Cumulative Relative	Class	Class
Lower	Upper	Tally	Frequency	Frequency	Frequency	Frequency	Boundaries	Midpoint
10	20	(1)	2	2/14	2	2/14	25	15
30	40	(11	3	3/14	5	5/14	45	35
50	60	(1)	3	3/14	8	8/14	45	55
70	80	HHT	6	6/14	14	$ 4 _{14} = $	03	75
			N = 14	T		· ~~~~		-
	Data:	68	57 19	66 69	79 70	27 39		
		AT	64	22	40			

0.143

0.357 .57

.0

Ø

Statistics 3	00		
Instructor:	Lawrence	C. Larsen	

Exam #1

(4 points; 6 minutes)

8. The California Healthcare Association wants to estimate the percentage of California households that have health isurance. For this purpose, the Association takes a random sample of 1330 addresses for California households and determines whether each household has health insurance. The Association found that 1104 of the 1330 households had health isurance, so the estimated percentage for all CA households was 83%.

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

SUVANE

Name:

1í IMSUN every household NAVHNA Total NO yes and =1 COUNT -< (b) What is the parameter of interest in this situation? have health house Ø

(c) What statistic was used in this situation and what was the value of that statistic?

noucen insurance а value =

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

P ause () () (I was a random



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".

Name:

	Statistic	Formula / Description	 Value	
Data:	Mid-range	$(\min + \max) = 2$	(49+89)=2 = 69	
52 89 83 83	Mode	The most frequent (or common) value	(83) occurs 2 times	
49 69	Mean	Zx n	69.37	calculator = TX
49 52	Median	The value in the middle when the data are in sorted order	(69+72)+2 = (70,5)	
58	Range Media N	Marx - Min	89 - 49 = (40)	
72	Standard deviation	$\Sigma(\chi - \overline{\chi})^2$	(15.17)	cale. Cotor
83		1 m-1		=S
83	Variance	$\sum (x - \overline{x})^2$	(229.98)	= 52
89		n-1		

Statistics 300 Instructor: Lawrence C. Larsen Name:

was

Solution

190

Exam #1

(4 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 (P₁₆) was 160 and the 84th percentile (P₈₄) was 340. Use all this information to estimate (not a wild guess) the standard deviation of the data. (Hint: What percent of the data are between P₁₆ and P₈₄?)



For a bell-shaped distribution 68% of the data should be between $(\mu - \sigma)$ and $(\mu + \sigma)$, which powers 200

So, <u>(340-160)</u> ~ 0

Name:

Instructor: Lawrence C. Larsen

Exam #1

(3 points; 3 minutes)

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

26 10 13 1.93 × 10

amanged is a different order, the tests would be different. So, order makes a difference

(3 points; 3 minutes)

11. 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 diveded into two groups. 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?

Order does not make a difference. Each group will 18 students (Yz of the dass), have 9075,135,300 or (9.08 × 109 10

(5 points; 4 minutes)

12. A different statistics professor likes multiple choice problems. That professor gives an exam that has 12 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 5 of the 12 problems?

 $P(\chi = 5) = 12^{\circ} (0.25)(0.75)^{7}$ Binomial M = 12p=0.25 q = 0.750-1032 X=5 125

(5 points; 4 minutes)

13. A very long multiple choice exam has 800 problems and each problem as 5 answers from which to choose. If someone guessed at random on every question, would it be unusual for that person to get 100 problems correct?

Z-Score! Binomial: N=500 and g=0.2 g=0.8 $7 = \frac{100 - 160}{11.21}$ N=100 M=np = 800 (0,2) = 160 $=\frac{-60}{11.31}=(-5.31)$ Mpg = 1 800 (.2) (.8) = 11.31 Z172 So YPS; UNUSUA

Name:	

Instructor: Lawrence C. Larsen

Exam #1

(3 points; 3 minutes)

Statistics 300

15. Circle the correct choice in each box in relation to the <u>underlined text</u>.



(3 points; 3 minutes)

16. A country's standard 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 one of the 26 letters in the English alphabet, but no digit or letter can occur more than once in the license plate. How many different standard license plates are possible?

 $(26P_3)$ = $(10P_4)(26P_3)$ (10)(26)(25)(24)(7)(8)(7) = (4)fundamental counting rule

(3 points; 3 minutes)

17. A bowl contains 18 jelly beans. Five are "Cherry", 6 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)?

 $P(0,0,G) = \left(\frac{6}{18}\right) \left(\frac{5}{14}\right) \left(\frac{5}{16}\right) = \frac{150}{4896} \neq 0.0306$

Name:

Instructor: Lawrence C. Larsen

Exam #1

(8 points; 8 minutes)

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

	Туре				
Number of Years with the Company	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

20

(a) What is the probability that someone picked at random from the 592 individuals in this table will be someone who has been with the Company for less than 6 years?

that is	someone	in the	(0to5) rou
Prob	= 210		= 0.355
	592		

(b) What is the probability that someone picked at random from the 592 individuals in this table will be a Middle Manager given that they have been with the company for 6 to 10 years?

P(middle Mgr | 6 to 10 years) = (-194 P(Middle Mgr and 6to 10yrs) 20/592 P(6 to 10 yrs)

(c) What is the probability that someone picked at random from the 592 individuals in this table will be someone who has been with the Company for 6 to10 years or be at and Entry Level job?

P(6 to 10 years or Entry Level) - P (both at the same time, = P(6 to 10 gars) + P(Entry Level) 666