Statistics 300 Name: Solution Fall 2010
Instructor: Lawrence C. Larsen Tu/Th 7:30 – 9:35 pm

Exam #1

(3 points : 3 minutes)

1. A store sells five different types of home appliances – refrigerators (R), washing machines (W), dryers (D), stoves (S), and automatic dishwashers (A). How any different ways could the next 10 sales happen? Example: W,S,R,R,D,W,D,W,R,A.

The types can repeat, so ...

$\underbrace{5 \cdot 5 \cdot 5 \cdots 5}_{\text{ways ways } \cdots} = 5^{10} = 9{,}765{,}625$

(3 points : 3 minutes)

2. A store sells five different types of home appliances – refrigerators (R), washing machines (W), dryers (D), stoves (S), and automatic dishwashers (A). How any different ways could the sales manager arrange one of each type of appliance side-by-side so customers will see them as they enter the store?

$5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5! = 120$

OR ${}_5P_5 = 120$

(3 points : 3 minutes)

3. A store sells five different types of home appliances – refrigerators (R), washing machines (W), dryers (D), stoves (S), and automatic dishwashers (A). But, the rate of defects is 18% . If the defects occur in a random fashion, what is the probability that the next five sales include at least one defective appliance?

$P(\text{defect}) = 0.18$
$P(\text{good}) = 0.82$

$P(\text{at least one defective}) = 1 - P(\text{all good})$
$= 1 - (0.82)^5 = 0.629$

(3 points : 3 minutes)

4. A store that sells five different types of home appliances – refrigerators (R), washing machines (W), dryers (D), stoves (S), and automatic dishwashers (A) – employs 20 people on the sales floor. Management has decided to select 5 of the 20 to take part in special training. How many different groups of 5 could management choose?

group counts, not order within the group.

${}_{20}C_5 = 15{,}504$

(4 points : 4 minutes)

5. A store that sells FOUR different types of home appliances – refrigerators (R), washing machines (W), dryers (D), stoves (S) – will only sell 2 items tomorrow. On the back of this page, list the sample space for the possible sales tomorrow. The following are examples of the possible sales: {A,D} and {W,W}.

If all sales are equally likely, what is the probability of at least one "W" tomorrow?

7/16

seven of the simple events have a W

Statistics 300 Name: Solution Fall 2010
Instructor: Lawrence C. Larsen Tu/Th 7:30 – 9:35 pm

Exam #1

(11 points : 10 minutes)

6. Use the table to answer questions (a) through (d).

Age Group	Number of Traffic & Parking Tickets Last Year							Total
	0	1	2	3	4	5	6 +	
1	131	13	18	17	18	13	17	227
2	154	15	20	15	19	15	11	249
3	150	19	12	14	20	14	19	248
4	167	11	17	17	19	18	12	261
5	160	14	20	16	17	19	10	256
6	152	20	19	15	12	12	17	247
7	144	13	10	18	13	11	15	224
8	145	20	11	14	20	12	18	240
Total	1203	125	127	126	138	114	119	1952

(a) If a person from this table is selected randomly, what is the probability that person will have received exactly ~~received~~ 3 tickets last year?

126 received 3 tickets
1952 ways to pick anyone

$\frac{126}{1952} = 0.0645$

(b) If a person from this table is selected randomly, what is the probability that person will not be a person from age group 4 who received 5 tickets last year?

$P(\text{Not group 4 with 5 tickets}) = 1 - P(\text{group 4 w/ 5 tick})$

$= 1 - \frac{18}{1952} = \frac{1934}{1952} = 0.9908$

(c) If a person from this table is selected randomly, what is the probability that person will have received zero tickets last year, given that the person is from age group 2?

$P(\emptyset \text{ tickets} \mid \text{group 2}) = 154/249 = 0.6185$

(OR) $= P(\emptyset \text{ Tick} \cap \text{Grp 2}) / P(\text{Grp 2}) = \frac{(154/1952)}{(249/1952)} = \frac{154}{249} = 0.6185$

(d) If a person from this table is selected randomly, what is the probability that person will have received less than 2 tickets or be from age group 5?

$P(< 2 \text{ tickets OR Group 5}) = P(< 2 \text{ Tickets}) + P(\text{Grp 5}) - P(\text{overlap})$

$= \frac{1328}{1952} + \frac{256}{1952} - \frac{174}{1952} = \frac{1410}{1952} = 0.722$

Statistics 300 Name: Solution Fall 2010
Instructor: Lawrence C. Larsen Tu/Th 7:30 – 9:35 pm

Exam #1

(4 points : 5 minutes)

7. Determine the type of sampling for each situation.

(a) Robert wants to test gum balls that come in bags of 10 in the stores. His friends help him get one bag from each of 50 different stores. The sample includes 50 blue, 180 red, 70 white, 120 green, and 80 orange gum balls. Then he tests them for poisons.

Simple random	Systematic
Stratified Random	Cluster
Convenience	Census

(b) Robert wants to test the gum balls now on the shelves of Drug stores. They are sold in bags of 10. He works with the stores to get random samples of 100 gum balls for each of the colors – blue, red, white, green, and orange. Then he tests them for poisons.

Strata

Simple random	Systematic
Stratified Random	Cluster
Convenience	Census

(c) Robert wants to test the gum balls now on the shelves of Drug stores. They are sold in bags of 10. He works with the stores to get 50 randomly selected bags from all the bags at the stores. Then he tests all 500 of the gum balls for toxins. Red ones have the most poisons.

clusters of 10 gum balls.

Simple random	Systematic
Stratified Random	Cluster
Convenience	Census

(d) Robert wants to test the gum balls that are currently being sold in Drug stores in his area. He works with the manufacturers to get a sample of 500 gum balls so that each gum ball they make next week is equally likely to be picked, independent of the rest.

Simple random	Systematic
Stratified Random	Cluster
Convenience	Census

(6 points : 7 minutes)

8. Based on the data on the next page (not a bell-shaped set), answer questions (a) and (b). There are 467 values arranged in sorted order, 46 rows of 10 plus 7 more.

(a) Estimate the "center" using a reasonable approach from this class (do not calculate the mean).

mid-range $\frac{\min + \max}{2} = \frac{128 + 1490}{2} = 809$

OR Median $L = 467\left(\frac{50}{100}\right) = 233.5 \uparrow 234$

Value @ $L = 234$ is 810

(b) Estimate the standard deviation using a reasonable approach from this class (do not calculate the standard deviation itself).

Range Rule

st. dev. $\approx \frac{\text{Max} - \text{Min}}{4} = \frac{1490 - 128}{4} = 340.5$

Data for problem of the previous page.

128	132	138	143	148	154	157	158	163	164
166	169	175	176	180	180	182	182	183	189
190	194	195	201	206	206	207	210	211	216
222	225	228	232	236	238	243	245	248	253
254	254	257	257	257	263	267	268	269	269
275	279	279	281	284	285	289	291	292	293
299	302	307	308	310	316	318	318	320	323
324	329	331	332	334	336	341	343	349	349
349	349	353	359	364	367	369	372	377	379
381	383	387	388	389	395	397	398	401	405
408	410	412	414	420	424	425	427	431	432
432	435	441	444	448	449	449	451	453	455
461	462	467	469	470	472	476	478	483	486
491	497	503	509	511	517	517	520	520	523
526	531	531	536	536	542	547	553	555	555
561	561	565	565	568	571	571	571	574	580
586	586	591	594	599	604	610	612	615	615
616	622	625	625	628	634	639	644	644	647
652	654	657	659	662	668	668	669	673	674
675	681	682	682	686	690	695	697	703	706
709	710	714	716	717	720	725	731	733	735
735	740	741	741	746	748	754	760	765	768
768	772	774	780	781	784	784	786	787	793
799	804	810	810	816	819	824	825	830	835
840	846	849	851	851	851	857	858	863	868
868	869	871	875	879	884	889	892	892	895
899	902	903	904	907	912	912	912	913	919
925	928	930	934	937	943	946	952	956	957
957	957	962	967	967	969	971	977	982	982
985	988	993	997	998	1001	1005	1010	1010	1011
1011	1013	1013	1015	1019	1022	1027	1029	1032	1032
1033	1037	1039	1042	1046	1052	1058	1064	1066	1069
1070	1075	1075	1076	1081	1082	1085	1090	1093	1097
1098	1101	1103	1105	1108	1109	1114	1114	1119	1121
1121	1121	1122	1123	1124	1125	1127	1133	1135	1137
1141	1146	1146	1151	1151	1156	1156	1160	1165	1167
1170	1176	1182	1185	1191	1193	1195	1196	1200	1202
1202	1205	1209	1209	1214	1218	1220	1222	1225	1225
1228	1233	1234	1234	1238	1242	1244	1245	1249	1251
1257	1259	1263	1264	1264	1264	1269	1273	1279	1282
1287	1292	1296	1296	1298	1304	1306	1306	1309	1311
1315	1319	1320	1325	1326	1328	1332	1334	1337	1341
1345	1350	1352	1356	1362	1365	1369	1375	1377	1379
1381	1383	1384	1385	1388	1391	1393	1398	1404	1408
1410	1416	1418	1418	1423	1429	1434	1437	1438	1443
1444	1450	1451	1457	1461	1463	1467	1471	1476	1479
1479	1482	1483	1484	1484	1484	1490			

Statistics 300 Name: Solution Fall 2010
Instructor: Lawrence C. Larsen Tu/Th 7:30 – 9:35 pm

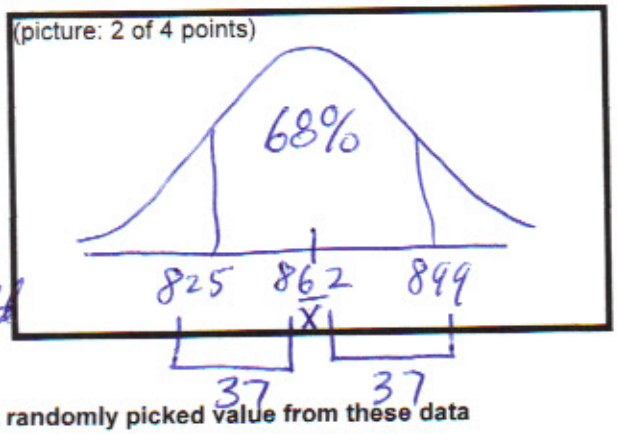
Exam #1

9. The values in a large set of data have a bell-shaped distribution. The mean of the distribution is 862 and the standard deviation is 37.

(4 points : 4 minutes)
(a) Approximately what percent of the data do you expect to find between 825 and 899?

825 is one st. dev. below the mean and 899 is one st. dev. above the mean. The empirical rule says that about 68% of the data should be between these limits.

(picture: 2 of 4 points)



(3 points : 3 minutes)
(b) What is the probability (approximately) that a randomly picked value from these data will be greater than 825 given that the value is less than the mean? (use the picture)

Bell-shaped distributions are symmetric, so 50% of values should be below the mean. Also, 34% or $(\frac{1}{2})68\%$, are greater than 825 and below the mean.

$$P(>825 \mid <\text{mean}) = \frac{P(>825 \text{ and } <\text{mean})}{P(<\text{mean})} = \frac{0.34}{0.5} = 0.68 \text{ or } 68\%$$

(6 points : 7 minutes)
10. **More than a million babies are born in the U.S. each year. According to Wikipedia, 30.2% of the births in 2005 were done by Caesarian section. If we select a random sample of 7 births from 2005 in the U.S., what is the probability that exactly three of them will be births done by Caesarian section?**

Binomial:
- N = Number of trials = 7 (known)
- Two possible outcomes: Caesarian and Not
- $P(\text{Caesarian}) = 0.302$ for each birth
- Independent (random)
- x = Number of Caesarian births in 7 trials

$p = 0.302$
$q = 0.698$

$$P(x=3) = {}_7C_3 (0.302)^3 (0.698)^4$$
$$= (35)(0.302)^3 (0.698)^4 = 0.2288$$

Statistics 300 Name: Solution Fall 2010
Instructor: Lawrence C. Larsen Tu/Th 7:30 – 9:35 pm

Exam #1

(8 points : 8 minutes)

11. Using the data shown below the table, fill in the values for all of the boxes.

Class Limits Lower	Class Limits Upper	Tally	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
80	90	I	1	1/12 = 0.083	1	1/12
100	110	IIII I	6	6/12 = 0.50	7	7/12
120	130	IIII	5	5/12 = 0.417	12	12/12 = 1

N = 12

Class Midpoints	Class Boundaries
85	
	95
105	
	115
125	

Class Width
20

100 − 80
120 − 100
105 − 85
125 − 105

Data: ~~130~~ ~~128~~ ~~129~~ ~~83~~ ~~99~~ ~~96~~ ~~129~~ ~~115~~ ~~101~~ ~~114~~ ~~96~~ ~~127~~

N = 12

(8 points : 8 minutes)

12. Calculate the mean, variance, and standard deviation for this discrete probability distribution and provide the formulas for each in their respective boxes.

x	P(x)	x · P(x)	(x − μ)² · P(x)	
14	0.52	7.28	79.44	
31	0.36	11.16	7.75	
66	0.12	7.92	188.56	

Σ = 1.00 Σ = 26.36 = μ 275.75 = σ²

$\sqrt{275.75} = 16.61 = \sigma$

Is the distribution valid? YES NO WHY? $\Sigma P(x) = 1$

Formulas for:

$\mu = \Sigma x \cdot P(x)$

$\sigma^2 = \Sigma (x-\mu)^2 \cdot P(x)$

$\sigma = \sqrt{\Sigma (x-\mu)^2 \cdot P(x)}$

Statistics 300 | Name: Solution | Fall 2010
Instructor: Lawrence C. Larsen | Tu/Th 7:30 – 9:35 pm

Exam #1

(14 points : 10 minutes)

13. For the sample of data in this problem, provide a math definition or formula for the given statistics. Then calculate the value of each statistic for the sample of data. You must use the special statistical functions of your calculator to get the mean and the standard deviation.

Data: ~~63~~ ~~72~~ ~~75~~ ~~67~~ ~~73~~ ~~69~~ ~~72~~ ~~63~~

63 } 2x
63 }
67
69
72 } 2x
72
73
75

$n = 8$

	Definition / Formula	Value for these data
Mode	the most common (frequently occuring) value	63 and 72
Mean	$\frac{\Sigma x}{n}$	69.25
Median	Value in the middle when data are in sorted order	70.5 = (69 + 72)/2
Mid-range	$\frac{\text{Min} + \text{Max}}{2}$	69 = $\frac{63+75}{2}$
Variance	$\frac{\Sigma(x-\bar{x})^2}{n-1} = s^2$	20.79 = $(4.559)^2$
Standard Deviation	$\sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = s$	4.559
Range	Max − Min	12 = 75 − 63

Statistics 300 Name: Solution Fall 2010
Instructor: Lawrence C. Larsen Tu/Th 7:30 – 9:35 pm

Exam #1

$Z = \frac{x - \mu}{\sigma}$

(5 points : 4 minutes)

14. Population "A" has a mean of 239 and a standard deviation of 43. Population "B" has a mean of 49 and a standard deviation of 8. A random value from population "A" was 285, and a random value from population "B" was 37. Which ~~or~~ of the two random values was most unusual? Circle your answer and explain your choice.

$Z_A = \frac{285 - 239}{43} = 1.07$ A's random value

B's random value

$Z_B = \frac{37 - 49}{8} = -1.5$

Why? Because $|-1.5| > |1.07|$

~~The~~ so $|Z_B| > |Z_A|$

(8 points : 5 minutes)

15. Circle the correct choice in each box based on the underlined test in each part.

	Are the data ... ?	Are the data ... ?
(a) <u>A dog, a cat, and a bird</u> compete in three races – a short one, a medium one, and a long one. The amount of energy used by each animal in each race is measured. The bird won 2 races, the dog won 1 race, and the cat won 0 races.	Qualitative Quantitative and discrete Quantitative and continuous	Nominal Interval Ordinal Ratio
(b) A dog, a cat, and a bird compete in three races – <u>a short one, a medium one, and a long one</u>. The amount of energy used by each animal in each race is measured. The bird won 2 races, the dog won 1 race, and the cat won 0 races.	Qualitative Quantitative and discrete Quantitative and continuous	Nominal Interval Ordinal Ratio
(c) A dog, a cat, and a bird compete in three races – a short one, a medium one, and a long one. The amount of energy used by each animal in each race is measured. <u>The bird won 2 races, the dog won 1 race, and the cat won 0 races.</u>	Qualitative Quantitative and discrete Quantitative and continuous	Nominal Interval Ordinal Ratio
(c) A dog, a cat, and a bird compete in three races – a short one, a medium one, and a long one. <u>The amount of energy used by each animal in each race is measured</u>. The bird won 2 races, the dog won 1 race, and the cat won 0 races.	Qualitative Quantitative and discrete Quantitative and continuous	Nominal Interval Ordinal Ratio

Statistics 300 Name: Solution Fall 2010
Instructor: Lawrence C. Larsen Tu/Th 7:30 – 9:35 pm

Exam #1

(8 points : 8 minutes)

17. For the data shown below (in sorted order, 240 values with 10 values per row), answer parts (a) and (b).

(a) What is the value of the 88th percentile (P_{88}) ? KNOW k and you want x

$L = \left(\frac{k}{100}\right)N = \left(\frac{88}{100}\right)240 = 211.2 \uparrow$ 212 location

where $x = 701$

$701 = P_{88}$

(b) What percentile is represented by the value 665 ? KNOW x and you want k

$k = \left(\frac{\#\text{ of values} < x}{\text{Total \# of values}}\right)100 = \left(\frac{181}{240}\right)100 = 75.42$

$665 = P_{75}, P_{75.4}, P_{75.42}$

386	387	390	392	395	398	399	401	401	401
401	404	405	405	408	410	412	415	416	416
417	419	422	424	425	427	428	430	430	430
432	433	434	436	439	442	445	447	449	450
450	452	453	454	457	459	462	462	463	466
467	468	470	471	474	476	479	481	482	484
485	485	488	490	493	495	495	497	499	502
504	507	509	509	512	512	512	513	513	514
514	514	516	518	519	519	520	520	523	526
527	530	532	534	537	538	538	539	542	544
546	549	552	555	556	557	558	558	560	561
563	563	566	567	570	571	571	572	575	577
579	580	582	583	586	588	589	589	590	592
592	594	597	600	603	605	608	610	612	612
612	615	617	620	623	624	627	630	630	631
632	632	633	634	634	635	636	638	640	641
643	643	645	645	647	648	648	649	649	651
651	653	656	656	656	659	661	663	663	663
664	665	665	665	667	669	671	671	672	673
673	673	676	676	676	676	678	681	681	684
685	685	688	691	693	695	695	695	697	700
700	701	701	702	703	705	706	707	709	712
712	712	715	716	716	717	717	718	718	720
720	720	722	723	725	726	727	730	730	733

$L = 212$

Statistics 300 Name: Solution Fall 2010
Instructor: Lawrence C. Larsen Tu/Th 7:30 – 9:35 pm

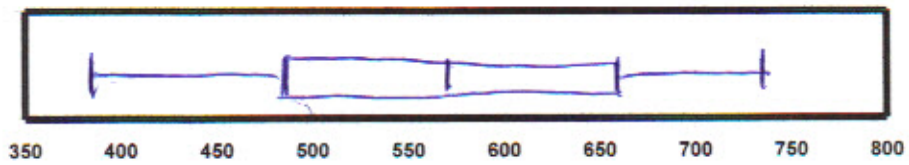
Exam #1

(4 points; 4 minutes)

18. According to the "percentile" function built into this program, the listed percentiles have the given values. Use the values given to construct a boxplot representing the distribution of the data.

P_0 =	386.0
P_5 =	405.0
P_{10} =	424.9
P_{15} =	444.6
P_{20} =	462.8
P_{25} =	484.8
P_{30} =	508.4
P_{35} =	518.7
P_{40} =	538.0
P_{45} =	559.1
P_{50} =	578.0
P_{55} =	595.4
P_{60} =	621.2
P_{65} =	635.4
P_{70} =	649.0
P_{75} =	663.3
P_{80} =	673.6
P_{85} =	691.3
P_{90} =	705.1
P_{95} =	718.0
P_{100} =	733.0

Draw Boxplot here



Boxplot requires 5 numbers:
Min, Q_1, Q_2, Q_3, and max
(P_0) (P_{25}) (P_{50}) (P_{75}) (P_{100})
= 386 / 484.8 / 578 / 663.3 / 733