97 points possible.

Statistics 300 **Name:** Solution **Fall 2012**
Instructor: Lawrence C. Larsen **Tue./Thu. 7:00 - 8:50 p.m.**

Exam #1

(4 points; 4 minutes)

1. **Identify each of the following "sampling" situations as RANDOM, STRATIFIED, SYSTEMMATIC, CLUSTER, or CONVENIENCE sampling.**

(a) The manager of a shoe store learns that some of the shoes they sell may be defective. The store sells 30 different styles, so the manager arranges to take random samples of 10 pair of shoes from each of the 30 different styles in the store. Each pair of shoes in the samples is inspected to find out if it is defective.

Type is sampling is: Stratified

(b) Health officials are worried that houses built before 1970 may have been painted with paint containing lead, which can be a danger to children. To study older homes, an experiment is funded in which 380 homes built before 1970 will be inspected. The 380 houses are chosen so that all possible samples of 380 of these homes are equally likely.

Type is sampling is: Random -0

(c) The government wants to know how many people in small California towns have health insurance. A survey is planned in which 40 small towns are selected at random. In each town, all the people will be studied to find out whether or not they have health insurance.

Type is sampling is: Cluster

(d) At a coal mine, the coal is loaded into railroad cars. The workers that load the cars are told to take a sample of the coal from every 200th car that they load. The sample of coal will be tested for the presence and the amount of toxic metals.

Type is sampling is: Systemmatic

(6 points; 7 minutes)

2. **A collection of 30 dice includes 4 bad dice (not balanced correctly) and 26 dice that are good (balanced correctly).**

(a) What is the probability that a random sample of four dice (without replacement) will have the exact sequence : "Good", "Bad", "Bad", "Good" ?

26 Good
4 Bad

$\left(\frac{26}{30}\right)\left(\frac{4}{29}\right)\left(\frac{3}{28}\right)\left(\frac{25}{27}\right) = 0.0119$

P(G and Bad and Bad and Good)

-0

(b) What is the probability that a random sample of two dice (without replacement) will have at least one "Good" die?

Use Complement Rule

P(at least one Good) = 1 − P(all Bad)

$= 1 - \left(\frac{4}{30}\right)\left(\frac{3}{29}\right)$

$= 1 - 0.0138$

$= 0.9862$

-0

-0

Statistics 300 Name: Solution Fall 2012
Instructor: Lawrence C. Larsen Tue./Thu. 7:00 - 8:50 p.m.

Exam #1

(14 points; 8 minutes)

3. **For the data in this problem, provide the requested information for the specified sample statistics.**
If you provide a description, it must give in words what the formula or procedure would do.
You must use your calculator's statistics functions for the mean and standard deviation.

Data:
~~50~~
~~9~~
~~14~~
~~10~~
86
~~9~~
~~63~~
~~78~~

9
9 mode
10
14
50
63
78
86

	Formula or description	Value
Range =	Max − Min	86 − 9 = 77
s =	$\sqrt{\frac{\sum (x-\bar{x})^2}{n-1}}$	33.138 (calculator)
s^2 =	$\frac{\sum (x-\bar{x})^2}{n-1}$	$(33.138)^2$ = 1098.125
$\bar{x}$ =	$\frac{\sum x}{n}$	39.875 (calculator)
Median =	Value in the middle when data are in sorted order	(14 + 50)/2 = 32
Mode =	The most common value.	9 occurs twice.
Midrange =	$\frac{\text{Min} + \text{Max}}{2}$	$\frac{9+86}{2} = \frac{95}{2}$ = 47.5

Statistics 300 Name: Solution Fall 2012
Instructor: Lawrence C. Larsen Tue./Thu. 7:00 - 8:50 p.m.

Exam #1

4. Use the data in the following table to answer parts (a) through (d).

	Employment Type for People in Sample				
State	Farm	Factory	Gov'ment	Service	To
Arizona	9	38	12	41	# 100
California	23	37	11	29	# 100
Nevada	22	40	9	29	# 100
Oregon	24	39	15	22	# 100
Washington	17	42	13	28	# 100
Total	95	196	60	149	# 500

(1 point, 1 minute)

(a) What is the probability that a randomly selected person from this sample will be a factory worker?

$\frac{196}{500}$ factory workers / total in table

(3 points, 3 minutes)

(b) What is the probability that a person randomly selected from this sample will be from Washinton given that the person is a factory worker?

$P(WA \mid factory) = \frac{42}{196}$ WA and factory / factory

OR

$= \frac{P(WA \text{ and } factory)}{P(factory)} = \frac{42/500}{196/500} = \left(\frac{42}{500}\right)\left(\frac{500}{196}\right) = \frac{42}{196}$

(3 points, 3 minutes)

(c) What is the probability that a person randomly selected form this sample will be from Oregon or work in government?

$P(OR \text{ OR } Gov't) = P(OR) + P(Gov't) - P(OR \text{ and } Gov't)$

$= \frac{100}{500} + \frac{60}{500} - \frac{15}{500} = \frac{145}{500}$

(2 points, 2 minutes)

(d) What is the probability that a person randomly selected form this sample will be from Arizona or Washington?

$P(AZ \text{ OR } WA) = P(AZ) + P(WA) - \emptyset$ (No overlap to subtract)

$= \frac{100}{500} + \frac{100}{500} = \frac{200}{500}$

Statistics 300 **Name:** Solution **Fall 2012**
Instructor: Lawrence C. Larsen **Tue./Thu. 7:00 - 8:50 p.m.**

Exam #1

(4 points; 3 minutes)

5. **For the data at the bottom of the page (given in sorted order), what percentile is represented by the value 160 ? [note: There are 8 rows of 10 plus 6 more values for a total of 86 values]**

$160 = P_k$ ($k = ?$)

$$k = \left(\frac{\#\text{ of values} < 160}{\text{total \# of values}}\right) 100$$

$$= \left(\frac{72}{86}\right) 100 = 83.72$$

$160 = P_{83.72}$

$= P_{83.7}$

$= P_{84}$

(4 points; 3 minutes)

6. **For the data below (given in sorted order), what is the 17th percentile? [note: There are 8 rows of 10 plus 6 more values for a total of 86 values]**

$P_{17} = ?$

$$\text{location of value} = L = \left(\frac{k}{100}\right) N = (0.17)86$$

$= 14.62 \uparrow$

$= 15^{\text{th}}$ position

$P_{17} = 34$

15^{th} position

4	5	8	12	14	17	21	25	26	29
30	30	31	32	34	36	38	41	42	43
45	48	50	54	57	60	60	62	64	65
66	70	73	75	76	77	80	82	84	84
84	86	89	91	91	93	97	97	98	100
101	104	104	106	110	111	111	113	116	120
123	126	128	132	135	136	138	140	144	147
150	153	160	160	160	164	165	167	169	171
171	174	177	180	186	186				

86 values.

Statistics 300 Name: Solution Fall 2012
Instructor: Lawrence C. Larsen Tue./Thu. 7:00 - 8:50 p.m.

Exam #1

(8 points; 7 minutes)

7. **In the context of discrete probability distributions, what are the expressions (formulas) for these parameters ?**

parameter	expression (formula)
μ	$\sum x \cdot P(x)$
σ	$\sqrt{\sum (x-\mu)^2 \cdot P(x)}$
σ^2	$\sum (x-\mu)^2 \cdot P(x)$

For each valid probability distribution below, calculate the mean, variance, and standard deviation. [Be sure to provide the formula for each parameter.]
(Use the extra empty columns however you wish.)

(a)

x	P(x)
0	0.2
1	0.5
2	0.2

$\sum = 0.9$

Not valid

(b)

x	P(x)	$x \cdot P(x)$	$(x-\mu)^2 \cdot P(x)$
10	0.7	7	11.2
20	0.2	4	7.2
30	0.1	3	25.6

$\sum = 1$ valid

$\sum = 14 = \mu$

$\sum = 44 = \sigma^2$

$\sqrt{44} = \sigma = 6.63$

Statistics 300 **Name:** Solution **Fall 2012**
Instructor: Lawrence C. Larsen **Tue./Thu. 7:00 - 8:50 p.m.**

Exam #1

(5 points; 5 minutes)

8. **A bank for businesses determines that a new business will earn $70,000.00 in profits if it is successful, but it will lose $40,000.00 if it fails. If the probability of success is 0.65 and the probability of failure is ~~0.45~~ 0.35, what is the expected value of a decision to start the new business?**

$E(x) = \sum x \cdot P(x)$ = the mean (μ)

x	$P(x)$	$x \cdot P(x)$
70,000	0.65	45500
−40,000	~~0.45~~ 0.35	−14000
		$\sum = \$31{,}500 = E(x)$

(10 points; 8 minutes)

9. **Complete the columns in the "Frequency Distribution" table using the data values given below and answer the two additional questions at the bottom of the page.**

Frequency Distribution

Class Boundary	Class Limits Lower	Class Limits Upper	Tally	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency	Class Midpoint
	10	20	I	1	1/8	1	1/8	15
25								
	30	40	II	2	2/8	3	3/8	35
45								
	50	60	III	3	3/8	6	6/8	55
65								
	70	80	II	2	2/8	(8)	8/8 = 1	75

N = (8)

Data: ~~35~~ ~~69~~ ~~71~~ 22 ~~65~~ ~~52~~ ~~46~~ ~~26~~

What is the lower class limit of the second class? 30

What is the class width for this frequency table? 20 = 30 − 10 two consecutive lower class limits

(OR) 45 − 25 = 20 two consecutive class boundaries

Statistics 300 Name: Solution Fall 2012
Instructor: Lawrence C. Larsen Tue./Thu. 7:00 - 8:50 p.m.

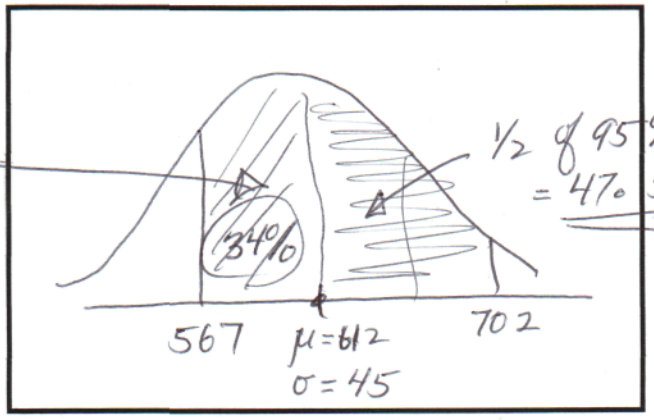
Exam #1

10. A very large dataset has a bell-shaped distribution with mean = 612 and standard deviation = 45. With respect to this dataset, answer parts (a) and (b) below.

(5 points; 6 minutes)

(a) What is the approximate percentage of this dataset that can be found between 567 and 702 ?

picture is required and it is worth 2 points



½ of 68% = 34%

½ of 95% = 47.5%

34% + 47.5% = 81.5%

answer

distance | 45 | | distance = 90 |
σ 2σ

$\frac{90}{45} = 2$

(4 points; 4 minutes)

(b) Which of the two values below would be more unusual if selected at random ?
(circle the most unusual value and provide a appropriate basis for your decision.)

550 $z = \frac{x-\mu}{\sigma} = \frac{550-612}{45} = |-1.38| = 1.38$

680 $z = \frac{x-\mu}{\sigma} = \frac{680-612}{45} = |1.51| = 1.51$

most unusual

Statistics 300 Name: Solution Fall 2012
Instructor: Lawrence C. Larsen Tue./Thu. 7:00 - 8:50 p.m.

Exam #1

(5 points; 6 minutes)

11. An experiment was done in which living sperm cells were exposed to intense radiation. As a result, 13% of the cells had damaged (mutated) DNA. In a random sample of 600 of the irradiated sperm cells, would it be unusual to find more than 93 damaged cells?

Binomial

Known # of trials = 600

Two outcomes each trial: damaged and not damaged

Independent

P(damaged) = 0.13 for all trials

X = total # of damaged cells in 600 trials

93 is unusual if $\left|\frac{93-\mu}{\sigma}\right| > 2$.

$\mu = n \cdot p = (600)(0.13) = 78$

$\sigma = \sqrt{npq} = \sqrt{(600)(0.13)(0.87)} = 8.24$

$Z = \frac{93-78}{8.24} = 1.82$

Not unusual

(5 points; 6 minutes)

12. An experiment was done in which living sperm cells were exposed to intense radiation. As a result, 13% of the cells had damaged (mutated) DNA. In a random sample of 10 of the irradiated sperm cells, what is the probability the sample will have exactly 2 damaged cells?

$= N$

$= X$

Binomal

$$P(X) = {}_nC_X (p)^X (1-p)^{(n-X)}$$

$$= {}_{10}C_2 (0.13)^2 (0.87)^8$$

$$= (45)(0.13)^2(0.87)^8 = 0.2496$$

Statistics 300 Name: Solution Fall 2012
Instructor: Lawrence C. Larsen Tue./Thu. 7:00 - 8:50 p.m.

Exam #1

(6 points; 6 minutes)

13. For each example of data, circle the correct answer in each column, A, B, and C. (Be sure to focus on the underlined portion of each statement.)

	A	B	C
a. The <u>years (such as 1964)</u> when Halley's Comet will pass through our solar system.	(Qualitative) Quantitative	If quantitative: (Discrete) Continuous	Nominal Ordinal (Interval) Ratio
b. The <u>weights</u> of asteroids that will pass within 10,000 miles of earth.	(Qualitative) Quantitative	If quantitative: Discrete (Continuous)	Nominal Ordinal Interval (Ratio)
c. The <u>names of comets</u> that will be visible from earth during the next 100 years.	(Qualitative) Quantitative	If quantitative: ~~Discrete~~ ~~Continuous~~	(Nominal) Ordinal Interval Ratio

(3 points; 3 minutes)

14. A high school has 1200 students. A random sample of 30 students will have their hearing tested. How many samples of 30 students are possible?

If the same 30 are picked, the order does not change the sample.

$${}_{1200}C_{30} = 6.2 \times 10^{59}$$

(3 points; 3 minutes)

15. A high school debate team had 7 members when it won the state championship. Individual pictures of the 7 members will be printed at the top of the front page in the town newspaper. How many ways can the 7 pictures be arranged from left to right?

7 pictures. All arrangements are different. $7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5040$

$N! = 7! = 5040$ $\quad {}_7P_7 = 5040$

(4 points; 4 minutes)

16. A high school has 1400 students. There are 500 in the freshmen class, 400 in the sophomore class, 300 in the Junior class, and 200 in the senior class. Two students from each class will be picked randomly to go on a trip the Washington, D.C. How many groups 8 students could be chosen to make the trip?

$$\underset{2\ \text{Frosh}}{{}_{500}C_2} \cdot \underset{2\ \text{Soph}}{{}_{400}C_2} \cdot \underset{2\ \text{Junior}}{{}_{300}C_2} \cdot \underset{2\ \text{Senior}}{{}_{200}C_2} = 8.9 \times 10^{18}$$

(2 points; 2 minutes)

17. You are a senior at the high school in problem #16. What is the probability you will get to go on the trip?

$$\frac{2}{200} = 0.01$$

$$\frac{{}_{199}C_1}{{}_{200}C_2} = \frac{\text{groups with you}}{\text{all possible groups}} = \frac{199}{19900} = 0.01$$

Statistics 300 **Name:** Solution **Fall 2012**
Instructor: Lawrence C. Larsen **Tue./Thu. 7:00 - 8:50 p.m.**

Exam #1

(5 points : 5 minutes)

18. For the situation below, select from the list of statistical term the one best connected to the bold and underlined part of the description.

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

a. Experimental Units — An experiment was done using **18 piglets**, 3 from each of six mother pigs. All of the piglets were exposed to swine flu virus, but three different vaccines were randomly assigned within each of the 6 sets of 3 piglets. None of the people working in the study knew which piglets received which vaccine. One of the vaccines had been deactivated so it would have no effect.

b. Blocks — An experiment was done using 18 piglets, **3 from each of six mother pigs**. All of the piglets were exposed to swine flu virus, but three different vaccines were randomly assigned within each of **the 6 sets** of 3 piglets. None of the people working in the study knew which piglets received which vaccine. One of the vaccines had been deactivated so it would have no effect.

c. Treatments — An experiment was done using 18 piglets, 3 from each of six mother pigs. All of the piglets were exposed to swine flu virus, but **three different vaccines** were randomly assigned within each of the 6 sets of 3 piglets. None of the people working in the study knew which piglets received which vaccine. One of the vaccines had been deactivated so it would have no effect.

d. Blinding — An experiment was done using 18 piglets, 3 from each of six mother pigs. All of the piglets were exposed to swine flu virus, but three different vaccines were randomly assigned within each of the 6 sets of 3 piglets. **None of the people working in the study knew which piglets received which vaccine**. One of the vaccines had been deactivated so it would have no effect.

e. Placebo — An experiment was done using 18 piglets, 3 from each of six mother pigs. All of the piglets were exposed to swine flu virus, but three different vaccines were randomly assigned within each of the 6 sets of 3 piglets. None of the people working in the study knew which piglets received which vaccine. **One of the vaccines had been deactivated so it would have no effect.**