92 points possible 66

Stat300: L.C.Larsen — Exam #1 — Name: Solution — Fall 2012 — Saturdays 9:00 a.m. to 1:05 p.m.

(6 points; 5 minutes)

1. An experiment consists of picking one of the four aces in a deck of cards, and rolling a die one time. List the possible outcomes in the sample space for this experiment.

A1 1, A1 2, A1 3, A1 4, A1 5, A1 6

A2 1, A2 2, A2 3, A2 4, A2 5, A2 6

A3 1, A3 2, A3 3, A3 4, A3 5, A3 6

A4 1, A4 2, A4 3, A4 4, A4 5, A4 6

24 outcomes in the sample space

A1 = A ♠
A2 = A ♡
A3 = A ◇
A4 = A ♣

If an outcome is selected at random from your sample space, what is the probability of getting the ace of spades for the card and an odd number (1, 3, or 5) for the die?

3/24

(4 points; 4 minutes)

2. The rules for making a password say that it must be made of 8 items. The 8 items must start with 2 of the 4 symbols, followed by 3 of the 8 letters, and end with 3 of the 6 digits shown below. Each item can only be used once, and each arrangement is a different password. How many different passwords could be formed?

The symbols are: & * % $
The letters are: A B C D E F G and H
The digits are: 1 2 3 4 5 and 6

2 symbols, 3 letters, 3 #s

Example 1: & * D E F 2 3 6

Example 2: * & E D F 6 2 3

$$(4 \cdot 3) \cdot (8 \cdot 7 \cdot 6) \cdot (6 \cdot 5 \cdot 4) = 483{,}840$$

$$({}_4P_2) \cdot ({}_8P_3) \cdot ({}_6P_3) = 483{,}840$$

(3 points; 3 minutes)

3. If the 10 digits (0, 1, 2, 3, 4, 5, 6, 7, 8, and 9) are written on one ping pong ball each and four of the 10 balls are selected, how many different sets of four digits could be formed? Example: the sets {0 1 7 2} and {7 2 0 1} are the same.

order does Not make a difference

$${}_{10}C_4 = 210$$

13

Stat300: L.C.Larsen
Exam #1 Name: Solution

Fall 2012
Saturdays 9:00 a.m. to 1:05 p.m.

(5 points; 6 minutes)

4. The Army wants to know the 95th percentile (P_{95}) of the scores of ALL its soldiers in a test of physical endurance that simulates combat. Because ti would be too expensive to put all the soldiers through the test, the Army selected 800 soldiers at random to do the test and then they use the 95th percentile of the sample scores as their estimate of the population parameter they want to know.

it

Use the information in the "story" to answer the following:

(a) What is the population of interest?

All the scores of all the soldiers in the Army when they take the endurance test

(c) What parameter was important for the Army to know?

P_{95} (95th percentile) of the sores of all the soldies in the test of endurance

(d) What statistic did the Army use instead?

the P_{95} for the ~~sample~~ 800 scores from the sample of soldies that took the endurance test.

(d) If all the 800 test scores were sorted from smallest to largest, at what location in the list would you find the statistic that the Army used?

the 95th percentile, P_{95}, would be found at location:

$$L = \left(\frac{95}{100}\right)800 = 760^{th}$$

761^{st}

average of the two values at these locations

+5

Stat300: L.C.Larsen
Exam #1 Name: Solution

Fall 2012
Saturdays 9:00 a.m. to 1:05 p.m.

(4 points; 4 minutes)

5. **A subway train has three safety systems. In an emergency, the first safety system will work 90% of the time, the second safety system will work 80% of the time, and the third will work 70% of the time. What is the probability that the train will handle the next emergency safely?** (Assume that the three systems operate independently of each other.)

P(1st OR 2nd or 3rd works) is NOT the answer.

$$1 - P(\text{all fail}) = 1 - (0.10)(0.20)(0.30) = 1 - 0.006 = 0.994$$

P(1st fails) and P(2nd fails) and P(3rd fails)

(4 points; 3 minutes)

6. **For the set of 39 values shown at the bottom of the page in sorted order, to what percentile does the value 99 correspond?**

33 values are less than 99. $99 = P_k$ ($k = ?$)

$$k = \left(\frac{(\#\text{ of values} < 99)}{(\text{total \# of values})}\right)100 = \left(\frac{33}{39}\right)100 = 84.6$$

$99 = P_{84.6}$ OR P_{85}

23	23	25	29	33	34	34	35	37	40
40	45	50	54	57	61	62	63	63	66
67	67	68	73	78	79	81	83	87	88
92	93	98	99	99	99	99	101	106	

+8

(5 points; 5 minutes)

7. All scores of two different tests have bell-shaped distributions. The scores for Test A have a mean of 110 points and a standard deviation of 42 points. The scores for Test B have a mean of 330 and a standard deviation of 68 points. A student takes Test A and earns a score of 91 points. Another student takes Test B and earns a score of 350 points. Which of the two test scores is the most unusual? Circle the answer and explain your choice.

The score on Test A. Why? Because $|Z_A| > |Z_B|$

$0.452 > 0.294$

The score on Test B.

Test A:

$x = 91$, $\mu = 110$, $\sigma = 42$

$Z_A = \left(\frac{91-110}{42}\right) = \boxed{-0.452}$

Test B:

$x = 350$, $\mu = 330$, $\sigma = 68$

$Z_B = \frac{350-330}{68} = \frac{20}{68} = \boxed{0.294}$

(9 points; 7 minutes)

8. Complete the columns in the "Frequency Distribution" table using the data values given below.

Frequency Distribution

Class Limits Lower	Upper	Tally	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
10	50	~~IIII~~ I	6	6/12	6	6/12
60	100	I	1	1/12	7	7/12
110	150	II	2	2/12	9	9/12
160	200	III	3	3/12	12	12/12 = 1
			N = 12			

Class Midpoint	Class Boundary
30	55
80	105
130	155
180	

Class Width
50

Data: ~~28~~ ~~157~~ ~~52~~ ~~76~~ ~~138~~ ~~47~~ ~~25~~ ~~171~~ ~~29~~ ~~43~~ ~~187~~ ~~133~~

(60-10) lower class limits

(105-55) class boundaries

+14

Stat300: L.C.Larsen
Exam #1 Name: Solution

Fall 2012
Saturdays 9:00 a.m. to 1:05 p.m.

9. Give a short definition of statistics (1 point; 1 minute):

The ART and Science of making sense out of data.

(4 points; 4 minutes)

10. For each situation below, select the appropriate statistical term from the list provided and write it in the blank next to the description or situation. Choose the term that is best connected to the underlined text in the description or situation.

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

Treatments Educators studied four different ways of teaching children to read. All of the second-grade classrooms in the state were grouped according to 5 different levels of economic/social status. Each way of teaching children to read was assigned at random to 20 different classrooms in each status group. The "status" groups did not take into account whether the classrooms were in urban, suburban, or rural locations. Teachers and students were not told which way of teaching was assigned to their classrooms, but everyone figured it out after a few weeks.

Blinding failed, which led to confounding Educators studied four different ways of teaching children to read. All of the second-grade classrooms in the state were grouped according to 5 different levels of economic/social status. Each way of teaching children to read was assigned at random to 20 different classrooms in each status group. The "status" groups did not take into account whether the classrooms were in urban, suburban, or rural locations. Teachers and students were not told which way of teaching was assigned to their classrooms, but everyone figured it out after a few weeks.

Blocking Educators studied four different ways of teaching children to read. All of the second-grade classrooms in the state were grouped according to 5 different levels of economic/social status. Each way of teaching children to read was assigned at random to 20 different classrooms in each status group. The "status" groups did not take into account whether the classrooms were in urban, suburban, or rural locations. Teachers and students were not told which way of teaching was assigned to their classrooms, but everyone figured it out after a few weeks.

Experimental UNITS Educators studied four different ways of teaching children to read. All of the second-grade classrooms in the state were grouped according to 5 different levels of economic/social status. Each way of teaching children to read was assigned at random to 20 different classrooms in each status group. The "status" groups did not take into account whether the classrooms were in urban, suburban, or rural locations. Teachers and students were not told which way of teaching was assigned to their classrooms, but everyone figured it out after a few weeks.

/5

Stat300: L.C.Larsen
Exam #1 Name: Solution

Fall 2012
Saturdays 9:00 a.m. to 1:05 p.m.

(8 points; 10 minutes)

11. For each discrete probability distribution, calculate the mean, variance, and standard deviation. Use the columns in the tables as you wish to use them.

x	P(x)	$x \cdot P(x)$		$(x-\mu)^2 \cdot P(x)$	
16	0.263	4.208		~~188.11~~	58.8
24	0.183	4.392		~~129.4~~	8.8
35	0.285	9.975		~~125.4~~	4.7
46	0.269	12.374		~~92.8~~	60.9
$\Sigma P(x) = 1$ (valid)		$\Sigma = \mu = 30.95$			$\Sigma = \sigma^2 = 133.2$

$\sigma = \sqrt{\sigma^2} = \sqrt{133.2} = 11.54$

x	P(x)				
0	0.25				
1	0.05				
2	0.15				
3	0.20				
4	0.13				
5	0.08				
6	0.06				
$\Sigma P(x) = 0.92$ (Not valid)					

Write the formulas for the mean, the variance, and the standard deviation of a discrete probability distribution.

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

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

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

+8

Stat300: L.C.Larsen
Exam #1 Name: Solution

Fall 2012
Saturdays 9:00 a.m. to 1:05 p.m.

(14 points; 8 minutes)

12. 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
~~38~~
~~48~~
57
~~34~~
~~33~~
~~40~~
~~34~~
~~45~~

sorted order
33
$\boxed{\begin{matrix}34\\34\end{matrix}}$ mode
38
median → 40
45
48
57

	Expression or Description	Value of statistic	
median	Value in the middle when data are put in sorted order.	$(38+40)\div 2$ 39	
range	Maximum − minimum	$57-33=$ 24	
variance	$\frac{\Sigma(x-\bar{x})^2}{n-1} = s^2$	$70.41 = s^2 = (8.391)^2$	
mode	the most frequently occurring value.	34	
midrange	$\frac{\text{Minimum} + \text{Maximum}}{2}$	$(33+57)/2=$ 45	
standard deviation	$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$	8.391	from calculator
mean	$\frac{\Sigma x}{n}$	41.125	from calculator

+14

Stat300: L.C.Larsen
Exam #1 Name: Solution

Fall 2012
Saturdays 9:00 a.m. to 1:05 p.m.

(6 points; 5 minutes)

13. An investigative reporter believes that 23% of the high school students in a major city are high on illegal drugs on any given day. If the reporter is correct and a random sample of 300 high school students from this town are tested for illegal drug use, would it be unusual to find 75 positive tests in the sample of 300?

Binomial $n = 300$ $x = 75$ $p(+) = 0.23$ $q = 0.77$

use "z"

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

$$z = \frac{75-69}{7.29} = \frac{6}{7.29} = 0.823$$

$$\boxed{\begin{aligned} \mu &= n \cdot p = 300(0.23) = 69 \\ \sigma &= \sqrt{npq} = 7.29 = \sqrt{(300)(.23)(.77)} \end{aligned}}$$

$|z| < 2$, so "not unusual"

(6 points; 6 minutes)

14. In another major city, the rate of illicit drug use by high school students is only 17%. If a random sample of 12 high school students is tested, what is the probability that exactly 3 will test positive for illegal drug use?

$n = 12$ $x = 3$
$p = 0.17$ $q = 0.83$

$$P(x=3) = {}_nC_x (p)^x (1-p)^{n-x}$$

$$= {}_{12}C_3 (0.17)^3 (0.83)^9$$

$$= 0.202$$

(5 points; 5 minutes)

15. Based on many years of experience, an investment advisor tells you that your investments will earn \$20,000 with 70% probability, earn \$0 with 20% probability, and lose \$12,000 with 10% probability. If your advisor is correct, what is the expected value of your income from investments this year?

Expected value $= E(x) = \sum x \cdot P(x)$

x	$P(x)$	$x \cdot P(x)$
20,000	0.70	14,000
0	0.20	0
-12,000	0.10	-1,200

$\Sigma =$ \$12,800 = Expected value of investment income

+17

(3 points; 3 minutes)

16. Circle the correct sampling plan for each situation,

A store manager wants to survey customers, so a bell is installed that will ring at a random time in each hour. When the bell sounds, each cashier records the next customer's gender and the total cost of the items purchased.

Simple Random	Systemmatic
Stratified Random	Cluster
Convenience	Census

A store manager wants to survey customers in a representative way. The survey is designed to make it easy for all customers to take part, so a questionaire is given to every customer along with their receipt. Questionaires can be returned on the customer's next visit.

Simple Random	Systemmatic
Stratified Random	Cluster
Convenience	Census

A store manager wants to survey customers in a representative way. The survey is designed to have 25 days selected randomly next year. On each of the 25 days, all customers will answer three questions and receive $5 in store credit.

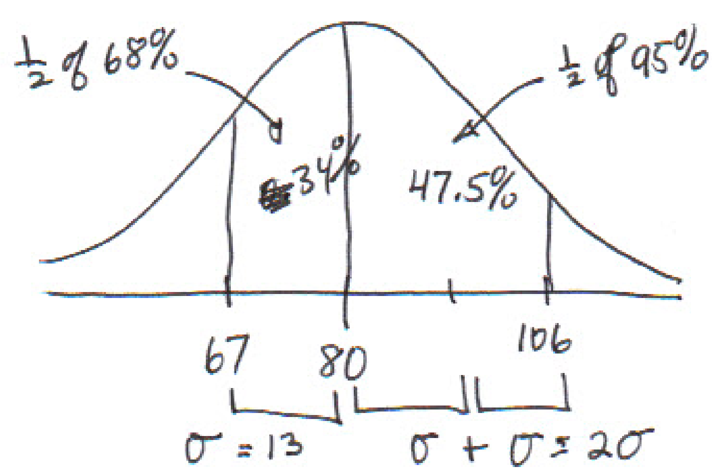
Simple Random	Systemmatic
Stratified Random	Cluster
Convenience	Census

No choice: the security guard is armed.

(5 points; 5 minutes)

17. A random variable has a bell-shaped distribution. The mean of the distribution is 80, and the standard deviation of the distribution is 13. What percent of the data do you expect to be found between 67 and 106?

Empirical rule for "bell-shaped" distributions



34.0
47.5
81.5

81.5%

+8

Statistics 300 : Fall 2012

Instructor : L. C. Larsen

Student name & ID#:

Student signature:

Exam : Unit 1

Time allowed : 2 hours and 5 minutes

Resources allowed:

== > Textbook (Author: Triola)
== > 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 (no outside messages)
== > Instructor at 916-346-6324

Resources not allowed:

== > Consultants other than the instructor
== > No phones, unless used as a calculator only