REVIEW PROBLEMS.

(1)

Elon 2201C

3.3 Another orthodontist's office employees receive the following hourly wages, in dollars.

15.67	23.45	18.95	20.79	25.49	
25.49	20.79	25.49	18.95	23.45	15.67

a. Using measures of central tendency, determine the shape of this data.
b. Determine the first and third quartiles for this data.

3.15 One of the leading business periodicals recently conducted a study of its subscribers to determine the total credit card debt for each customer. A sample of 50 subscribers responded to the survey, with the following results, in dollars.

$1,366	$0	$1,692	$2,973	$2,426	$2,090	$2,429	$3,306	$3,050	$2,085
3,269	2,261	3,011	3,617	2,273	2,960	3,203	347	0	2,441
2,516	3,727	2,085	2,010	700	2,301	2,096	2,008	2,653	3,088
2,257	8,345	2,523	1,948	2,685	3,393	2,591	1,209	3,621	300
3,612	2,380	0	2,681	2,506	3,076	4,065	2,218	3,287	3,712

a. Develop a box and whisker plot for these sample data.
b. Based on the box and whisker plot, does it appear that the distribution of credit card debt is skewed? If so, in which direction is it skewed? Discuss.

3.25 The Price Corporation has built six homes during the past year. The number of square feet in each home (treated as the population of interest) is listed as follows:

square feet = {1,560; 2,340; 1,990; 1,750; 4,000; 2,200}

a. Compute the range.
b. Compute the variance.
c. Compute the standard deviation.
d. Write a short paragraph that describes these data. Please feel free to also compute measures of the center and include these values in your discussion.

3.28 Grover's Pay n' Pak sells hardware supplies to "do-it-yourselfers." One of the things the company prides itself on is fast service. It uses a number system and takes customers in the order in which they arrive at the store. Recently, the assistant manager tracked the time customers spent in the store from the time they took a number until they left. A sample of 16 customers was selected, and the following data (measured in minutes) were recorded.

15	14	16	14	14	14	13	8
12	9	7	17	10	15	16	16

a. Compute the mean, median, mode, range, interquartile range, and standard deviation.
b. Develop a box and whisker plot for these data.

2

1.11 4.11 The Skateworld Company operates ice rinks in several major cities throughout the United States. During each session of open skating, one customer is selected at random to receive a free pass for a future open skating session. At a recent session there were 150 males and 130 females skating.

a. What is the probability that the person selected for the free pass will be a female?
b. Referring to part a, what method of probability assessment is used to determine the probability?
c. Suppose the company decides to give free passes to two customers. Are the events that a female received the first pass and a male received the second pass independent? Why or why not?

4.21 A paint store carries three brands of paint. A customer arrives and wants to buy another gallon of paint to match paint that she purchased at the store previously. She can't recall the brand name and does not wish to return home to find the old can of paint. She selects two of the three brands of paint at random and buys them.

a. What is the probability that she matched the paint?
b. Her husband also goes to the paint store and fails to remember what brand to buy. He also purchases two of the three brands of paint at random. Determine the probability that both the woman and her husband fail to get the correct brand of paint. (Hint: Are the two events independent?)

4.35 The Ace Construction Company has submitted a bid on a state government project in Delaware. The price of the bid was predetermined in the bid specifications. The contract is to be awarded on the basis of a blind drawing from those who have bid. Five other companies have also submitted bids.

a. What is the probability of Ace Construction winning the bid?
b. Suppose that there are two contracts to be awarded by a blind draw. What is the probability of Ace winning both contracts?
c. Referring to part b, what is the probability of Ace not winning either contract?
d. Referring to part b, what is the probability of Ace winning exactly one contract?
e. Referring to part b, what is the probability of Ace winning at least one contract?

4.47 The following probability distributions are given for two discrete random variables, x and y.

x	y	$P(x)$	$P(y)$	$P(xy)$
100	500	0.25	0.25	0.10
200	300	0.40	0.40	0.50
300	400	0.20	0.20	0.30
400	600	0.15	0.15	0.10

Compute the correlation coefficient and indicate what it means with respect to these two variables.

(3)

4.25 A local ski area offers private ski lessons with professionally qualified ski instructors. There are three ski instructors available. One is Austrian, one is German, and the third is from the United States. According to company policy, the instructors are assigned randomly. Thus, when a customer calls, a random selection is made and the selected instructor is scheduled with that customer.

a. On a given day, five customers call for lessons. Of these, four are assigned to the German instructor and one to the American. What is the probability of this happening if the assignments are random?

b. On a different day, three customers call for lessons and all three are assigned to the German instructor. What is the probability of this happening?

c. Referring to parts a and b, compute the probability that both the outcomes for day one and day two happen. Based on this probability, is there any cause for concern that the ski-lesson assignment may not be random? Explain.

4

4.49 The Seremonte Emergency Medical Department has recorded the number of emergency calls received each day for the past 200 days. These data are shown in this frequency distribution.

Calls	Number of Days
0	22
1	20
2	40
3	55
4	28
5	20
6	5
7	10
	200

a. Determine the probability distribution based on the frequency distribution.
b. What is the mean of the probability distribution?
c. What is the standard deviation of the probability distribution?
d. Compute the coeffiecient of variation.
e. Each emergency call requires a team of three individuals to respond. How many employees must Seremonte have so they can respond to at least 75% of the emergency calls?

4.9 A major airline has tracked its on-time status during the past year for flights originating in San Francisco and Los Angeles. The following data reflect the data for 400 flights.

Origination	On-Time Status		
	Early	*On Time*	*Late*
San Francisco	25	50	100
Los Angeles	50	100	75

a. Based on these data, what is the probability that flight from one of the two cities will arrive early?
b. What is the probability that a flight will have originated in Los Angeles?
c. Given that the flight originated in Los Angeles, determine the probability that it will arrive early. What would this probability have to be if the event arriving early were independent from the event in Los Angeles?
d. If three flights are selected at random, list the sample space indicating the possible "on-time" status for all three.

(5)

5.3 Use the binomial distribution table to calculate the following probabilities for a binomial random variable where $n = 20$ and $p = 0.40$.

a. $P(x = 10)$
b. $P(7 < x < 12)$
c. $P(x \geq 12)$

5.5 Use the counting rule for combinations to determine:

a. the number of ways 4 items can be selected from 8 items
b. the number of ways 6 items can be selected from 10 items
c. the number of ways 3 items can be selected from 10 items
d. the number of ways 7 items can be selected from 10 items

5.17 A survey by KRC Research for *U.S. News* reported that 37% of people plan to spend more on eating out after they retire. If eight people are randomly selected, determine the following probabilities.

a. Exactly five people plan to spend more on eating out after they retire.
b. Fewer than four people plan to spend more on eating out after they retire.
c. More than two plan to spend more on eating out after they retire.

5.28 If $\lambda t = 3.5$ for a Poisson-distributed variable, find the following:

a. $P(2 \leq x \leq 5) =$
b. $P(x = 3) =$
c. $P(x \geq 1) =$

5.55 A commuter airline has studied the passenger counts on a flight between Boston and Atlanta and has found that the number of passengers who purchase tickets for this flight is approximately normally distributed, with a mean of 72 and a standard deviation of 4. The data were determined for all days, regardless of the number of tickets sold on the flight. Keep in mind, some people do not show up for their flight.

a. If the capacity on the plane is 85, what percentage of the time should the flight be full?
b. The catering manager who is responsible for snack and beverage provisions on the flight plans to stock 90 snack packs. What is the probability that there will be 8 or fewer left over, assuming that each passenger gets one snack pack?
c. Comment on the potential problems in assuming that the number of fliers on a flight is normally distributed? What type of variable is the number of fliers? Discuss.

(6)

6.67 A continuous random variable is uniformly distributed between 20 and 60.

a. What is the probability a randomly selected value will be above 50?

b. Calculate the probability a randomly selected value will be exactly 45.

c. Determine the probability that a randomly selected value will be between 25 and 35.

d. Find the probability that a randomly selected value will be less than 34.

6.3 Consider the following values to represent a population:

129	330	100	200	150	105	100	130	190	400	120

a. If a random sample of $n = 3$ items includes the following, compute the sampling error.

150	100	400

b. Determine the range for the possible sampling error when a sample of size $n = 3$ is used. (Hint: find the sampling error for the three smallest values and the sampling error for the three largest values.)

c. Refer to part b and determine the range of potential sampling error if the sample size is increased to 5. Discuss the impact of sample size on the potential for extreme sampling error.

6.7 The owner of Miller's Union 76 in Rochester, New York, has tracked gasoline sales for several years and is confident that the true mean sale is 16.9 gallons. Assuming that this is the population mean, if a random sample of 10 fill-ups is collected, what is the sampling error if the gallons per fill are

13.3	19.8	22.6	15.0	19.3	9.7	17.5	22.4	18.0	13.0

6.11 A population is normally distributed, with a mean of 1,000 and a standard deviation equal to 200.

a. Determine the probability that a random sample of size 5 selected from this population will have a sample mean less than 970.

b. Referring to part a, suppose a second sample of size 10 is selected. What is the probability that this sample will have a mean that is less than 970?

c. Why are the answers to parts a and b different? Discuss.

6.31 Given a population in which the proportion of items with a desired attribute is $p = 0.65$, if a sample of 100 is taken:

a. Determine the probability the proportion of successes in the sample will be less than 0.63.

b. Referring to part a, suppose the sample size is increased to $n = 200$, what is the probability that the sample proportion will be less than 0.63? Discuss why the answers in parts a and b differ.

6.35 Given a population in which the probability of a success is $p = 0.40$, if a sample of 1,000 is taken:

a. Calculate the probability the proportion of successes in the sample will be less than 0.42.

b. What is the probability the proportion of successes in the sample will be greater than 0.44?

7

7.9 Given the following data from a simple random sample for the population of interest, compute the 95% confidence interval estimate. (What assumption must be made about the population?)

7.15 Agri-Beef, Inc., is a large midwestern farming operation. The company has been a leader in employing statistical analysis techniques in its business. Recently, John Goldberg, operations manager, requested that a random sample of cattle be selected and fed a special diet. The cattle were weighed before the start of the new feeding program and at the end. John wished to estimate the average daily weight gain for cattle on the new program. Two hundred cattle were tested, and the sample results were

$$\bar{x} = 1.2 \text{ lb. per day gain}$$
$$s = 0.50 \text{ lb}$$

a. Obtain a 95% confidence interval estimate for the true average daily weight gain.
b. Provide a 90% confidence interval estimate for the true average daily weight gain.
c. Discuss the difference between the estimates found in parts a and b and indicate the advantages and disadvantages of each.

7.27 Suppose, as part of your job, you are asked to estimate a population mean using a 90% confidence interval and a margin of error of 60. What size sample is required if the following pilot sample is used to determine a value to use for the population standard deviation?

3,239	3,144	2,960	2,507	2,842
3,134	3,249	2,908	2,754	2,715

7.36 The Northwest Pacific Phone Company wishes to estimate the average number of minutes its customers spend on long-distance calls per month. The company wants the estimate made with 99% confidence and a margin of error of no more than 5 minutes.

a. A previous study indicated that the standard deviation for long-distance calls is 21 minutes per month. What should the sample size be?
b. Determine the required sample size if the confidence level were changed from 99% to 90%.
c. What would the required sample size be if the confidence level was 95% and the margin of error was 8 minutes?

7.37 The quality manager for a major automobile manufacturer is interested in estimating the mean number of paint defects in cars produced by the company. She wishes to have her estimate be within ± 0.10 of the true mean and wants 98% confidence in the estimate. The data in the CD-ROM file called *CarPaint* contains a pilot sample that was conducted for the purpose of determining a value to use for the population standard deviation. How many additional cars need to be sampled to provide the estimate required by the quality manager?

(8)

7.43 A sample of $n = 300$ items has been randomly selected. Of these, 55 contain the attribute of interest. Based on this information, compute a 90% confidence interval estimate for the proportion of items in the population that have this attribute.

7.55 Watson, Harris & Tonkin is a CPA firm in Columbus, Ohio. As part of an audit of a large retail company, the firm wishes to estimate the proportion of credit card accounts that are past due. They want this estimate to have a margin of error no greater than 0.03, and they wish to use a 95% confidence interval estimate.

a. What size sample should they select if they have no idea what the percentage might be and they want to make sure the sample size is large enough to meet their requirements?
b. Suppose that the sample size determined in part a is used and the sample proportion is 0.18. Construct the confidence interval estimate for the population.
c. Referring to parts a and b, discuss why the margin of error obtained in the interval estimate is actually smaller than what the firm wanted.

8.1 For each of the following claims, list the appropriate null and alternative hypotheses:

a. The mean is larger than 20.
b. The mean equals 50.
c. The mean is at least 35.
d. The mean is more than 87.
e. The mean is at most 6.

8.5 Given the following null and alternative hypotheses,

$$H_0: \mu \geq 4{,}000$$
$$H_A: \mu < 4{,}000$$
$$\alpha = 0.05$$

and

$\bar{x} = 3{,}980$ $s = 205$ $n = 100$

a. Establish the appropriate decision rule.
b. Indicate the appropriate decision based on the sample information, using both the p-value and $\bar{x}$.
c. Provide the two research hypotheses that could have produced the null and alternative hypotheses in this problem.

8.9 Determine the p-value for each of the following hypothesis scenarios:

a. $H_0: \mu = 1{,}346$ versus $H_A: \mu \neq 1{,}346$ and $z = 2.36$.
b. $H_0: \mu \geq 4{,}000$ versus $H_A: \mu < 4{,}000$ and $z = -1.85$.
c. $H_0: \mu \leq 24.78$ versus $H_A: \mu > 24.78$ and $z = 0.84$
d. $H_0: \mu \leq 200$ versus $H_A: \mu > 200$ and $z = -2.06$ (be careful here).

8.23 Calculate the critical values for the following situations:

a. $H_A: p > 0.4$, $n = 150$, $\alpha = 0.05$
b. $H_A: p < 0.7$, $n = 200$, $\alpha = 0.10$
c. $H_A: p \neq 0.85$, $n = 100$, $\alpha = 0.10$

8.27 Given the following null and alternative hypotheses,

$$H_0: p \leq 0.24$$
$$H_A: p > 0.24$$

test the null hypothesis based on a random sample of $n = 100$, where $\bar{p} = 0.27$. Assume an $\alpha = 0.05$ level.

a. Use $\bar{p}_\alpha$ as the test statistic to test the hypothesis. Be sure to show clearly the decision rule.
b. Use z as the test statistic to test the h...

9

8.31 A large number of complaints have been received in the past six months regarding airlines losing fliers' baggage. The airlines claim the problem is much smaller than newspaper articles have indicated. In fact, one airline spokesman claimed that fewer than 1% of all bags fail to arrive at their destinations with the passengers. To test this claim, 800 bags were randomly selected at various airports in the United States when they were checked with this airline. Of these, 6 failed to reach their destinations when their owners arrived.

a. Is this sufficient evidence to support the airline spokesman's claim? Test using a significance level of 0.05. Discuss.

b. Estimate the proportion of bags that fail to arrive at the proper destinations using a technique for which 95% confidence applies.

①

Solutions

3-3.

a.

Review Questions – Econ 2201

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} = 234.19/11 = \$21.29$$

To compute the median, rank the observations and find the middle value.

15.67 15.67 18.95 18.95 20.79 20.79 23.45 23.45 25.49 25.49 25.49

Median = \$20.79
Mode = 25.49
Because the mean (21.29) is greater than the median (20.79) this data is right-skewed

b. The 1st quartile is equal to the 25th percentile

$i = \frac{p}{100}(n+1)$ = (25/100)(12) = 3 or 3rd observation = 18.95

The 3rd quartile is equal to the 75th percentile

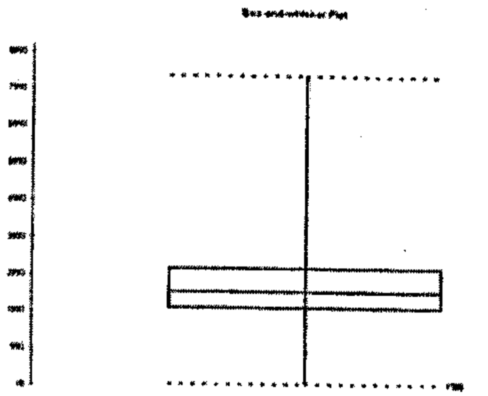
$i = \frac{p}{100}(n+1)$ = (75/100)(12) = 9 or 9th observation = 25.49

3.15

a.

Box-and-whisker Plot

Five-number Summary	
Minimum	0
First Quartile	2085
Median	2506
Third Quartile	3145.5
Maximum	8345



b. Based on the box and whiskers plot, it appears that the data are skewed. The box is not evenly centered within the range of the whiskers. The data appear to be right-skewed. Also, the mean is 2,847.8 which is higher than the median

3-25.

The data are assumed to be a population since the six homes is every home built by the Price Corporation.

a. Range = 4,000 – 1,560 = 2,440

b.

X	$X - \mu$	$(X - \mu)^2$
1,560	(747)	557,511.11
2,340	33	1,111.11
1,990	(317)	100,277.78
1,750	(557)	309,877.78
4,000	1,693	2,867,377.78
2,200	(107)	11,377.78
13,840		3,847,533

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x-\mu)^2}{N} = 3,847,533/6 = 641,255.7$$

c.

$$\sigma = \sqrt{\sigma^2} = \sqrt{641,255.5} = 800.7843$$

d. Student answers will vary.

3-28.

a.

(3)

X	X - $\bar{x}$	(X - $\bar{x}$)^2
15	1.875	3.515625
14	0.875	0.765625
16	2.875	8.265625
14	0.875	0.765625
14	0.875	0.765625
14	0.875	0.765625
13	-0.125	0.015625
8	-5.125	26.26563
12	-1.125	1.265625
9	-4.125	17.01563
7	-6.125	37.51563
17	3.875	15.01563
10	-3.125	9.765625
15	1.875	3.515625
16	2.875	8.265625
16	2.875	8.265625
210		141.75

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 210/16 = 13.125$$

To compute the median, rank the observations and compute the average of the middle two.

7 8 9 10 12 13 14 14 14 14 15 15 16 16 16 17

Median = (14 + 14)/2 = 14

Mode = 14

Range = 17 – 7 = 10

The 1st quartile is equal to the 25th percentile

$i = \frac{p}{100}(n+1)$ = (25/100)(16+1) = 4.25 so 25 percent of the distance between the 4th and 5th values = 10 +.25(10-12) = 10.5

The 3rd quartile is equal to the 75th percentile

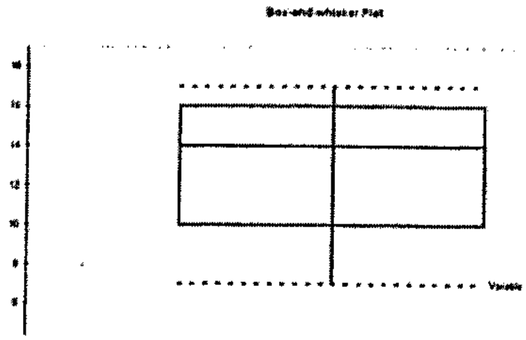
$i = \frac{p}{100}(n+1)$ = (75/100)(16+1) = 12.75 so 75 percent of the distance between 12th and 13th values = 15 +.75(16 - 15) = 15.75

Interquartile Range = 15.75 – 10.5 = 5.25

$$S^2 = \frac{\sum_{i=1}^{n}(x-\bar{x})^2}{n-1} = 141.75/(16\text{-}1) = 9.45$$

$$S = \sqrt{S^2} = \sqrt{9.45} = 3.0741$$

b.



4.9.

Origination	Early	On-Time	Late	Total
San Francisco	25	50	100	175
Los Angeles	50	100	75	225
Total	**75**	**150**	175	400

a. P(Early) = 75/400 = 0.1875.

b. P(Los Angeles) = 225/400 = 0.5625.

c. P(Early Given Los Angeles) = 50/225 = 0.2222. P(Early) = 50/400 = 0.125

d. Let E = Early, O= On-Time, and L = Late. The events are as follows:

Elementary Event	Flight 1	Flight 2	Flight 3
Event 1	E	E	E
Event 2	E	E	O
Event 3	E	E	L
Event 4	E	O	E
Event 5	E	O	O
Event 6	E	O	L
Event 7	E	L	E
Event 8	E	L	O
Event 9	E	L	L
Event 10	O	E	E
Event 11	O	E	O
Event 12	O	E	L
Event 13	O	O	E
Event 14	O	O	O
Event 15	O	O	L
Event 16	O	L	E
Event 17	O	L	O
Event 18	O	L	L
Event 19	L	E	E
Event 20	L	E	O
Event 21	L	E	L
Event 22	L	O	E
Event 23	L	O	O
Event 24	L	O	L
Event 25	L	L	E
Event 26	L	L	O
Event 27	L	L	L

The sample space consists of the 27 elementary events listed in the table. Sample Space = {Event 1, Event 2, Event 3, …, Event 27}

4.11.

5

Male	150
Female	130
Total	280

a. P(Female) = 130/280 = 0.4643

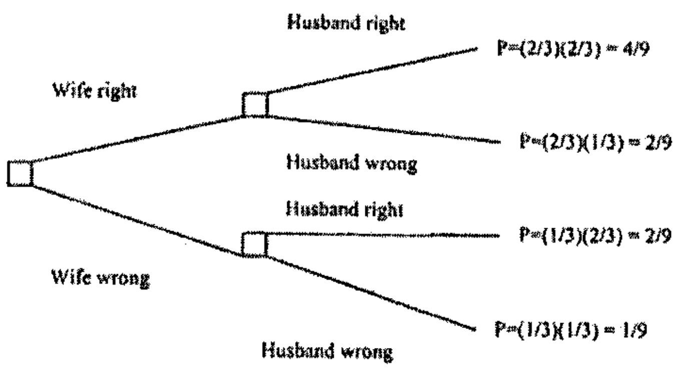
b. Relative frequency approach since it is based on actual observations.

c. No. Once the first customer is selected, the probabilities for both a male and a female customer being selected as the second winner change. Therefore the events are not independent.

4.21.

a. P(matched) = 2/3

b. P(both wrong) = (1/3)*(1/3) = 1/9



4.25.

a. List the number of ways four German instructors and one American instructor could be assigned and then determine the probabilities of each of these ways.

A-G-G-G-G = (1/3)(1/3)(1/3)(1/3)(1/3) = 0.00412
G-A-G-G-G = (1/3)(1/3)(1/3)(1/3)(1/3) = 0.00412
G-G-A-G-G = (1/3)(1/3)(1/3)(1/3)(1/3) = 0.00412
G-G-G-A-G = (1/3)(1/3)(1/3)(1/3)(1/3) = 0.00412
G-G-G-G-A = (1/3)(1/3)(1/3)(1/3)(1/3) = 0.00412

The sum of these probabilities is 0.02060

b. P(G-G-G) = 1/3(1/3)(1/3) = 0.03704

c. (0.0206)(0.03704) = 0.00076; because this probability is so low that if it really did happen then the assignments are probably not made randomly.

6

4.35.

There are a total of 6 companies.

a. P(Ace) = 1/6 = 0.1667

b. P(Win1 and Win2) = (1/6)(1/6) = 0.0278

c. P(Lose1 and Lose2) = (5/6)(5/6) = 0.6944

d. P(Win1 and Lose2) + P(Lose1 and Win2) = (1/6)(5/6) + (5/6)(1/6) = 0.2778

e. P(Win1 and Lose2) + P(Lose1 and Win2) + P(Win1 and Win2) = 0.1389 + 0.1389 + 0.0278 = 0.3056 or 1 – P(Lose 1 and Lose 2) = 1 – 0.6944 = 0.3056

4.47.

To find the correlation, use equation 4-18:

$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

The covariance is found by equation 4-17:

x	P(x)	xP(x)	x - E(x)	y	P(y)	yP(y)	y - E(y)	[x - E(x)][y - E(y)]	P(xy)	[x - E(x)][y - E(y)]P(xy)
100	0.25	25	-125	500	0.25	125	85	(10,625.00)	0.10	(1,062.50)
200	0.40	80	-25	300	0.40	120	-115	2,875.00	0.50	1,437.50
300	0.20	60	75	400	0.20	80	-15	(1,125.00)	0.30	(337.50)
400	0.15	60	175	600	0.15	90	185	32,375.00	0.10	3,237.50
		225				415				3,275.00

The covariance is 3,275 The relationship between the two variables is positive

The standard deviation for variable x is found as follows:

x	P(x)	xP(x)	x-E(x)	$[x-E(x)]^2$	$[x-E(x)]^2P(x)$
100	0.25	25	-125	15625	3906.25
200	0.4	80	-25	625	250
300	0.2	60	75	5625	1125
400	0.15	60	175	30625	4593.75
		225			9875

$\sigma_x = \sqrt{9{,}875} = 99.37$

The standard deviation for y is found using:

y	P(y)	yP(y)	y - E(y)	$[y-E(y)]^2$	$[y-E(y)]^2P(y)$
500	0.25	125	85	7,225	1,806.25
300	0.40	120	-115	13,225	5,290.00
400	0.20	80	-15	225	45.00
600	0.15	90	185	34,225	5,133.75
		415			12,275.00

$\sigma_y = \sqrt{12{,}2751} = 110.79$

The correlation is: $\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$ = (3,275)/[(99.37)(110.79) = 0.2975

The correlation measures the strength of the linear relationship between the two variables. In this case, there is a weak positive linear relationship.

(9)

4-49.

a. To determine the probability you need to divide the number of days by the total of 200.

x	P(x)	xP(x)	x-E(x)	$[x-E(x)]^2$	$[x-E(x)]^2P(x)$
0	0.110	0.000	-2.885	8.3232	0.9156
1	0.100	0.100	-1.885	3.5532	0.3553
2	0.200	0.400	-0.885	0.7832	0.1566
3	0.275	0.825	0.115	0.0132	0.0036
4	0.140	0.560	1.115	1.2432	0.1741
5	0.100	0.500	2.115	4.4732	0.4473
6	0.025	0.150	3.115	9.7032	0.2426
7	0.050	0.350	4.115	16.9332	0.8467
		2.885			3.1418

b. E(x) = 2.885

c. 1.7725

d. The coefficient of variation is:

$$CV = \frac{\sigma}{\mu}(100) = \frac{\sigma_x}{E(x)}(100) = \frac{1.7725}{2.885}(100) = 61.44\%$$

e. Students need to look at the 3[rd] quartile. This is accomplished by determining the cumulative P(x). The 75[th] percentile would be between 3 and 4. Since you want at least 75% you should choose x=4 which means you would need 4(3) = 12 employees.

x	P(x)	Cum P(x)
0	0.110	0.110
1	0.100	0.210
2	0.200	0.410
3	0.275	0.685
4	0.140	0.825
5	0.100	0.925
6	0.025	0.950
7	0.050	1.000

5.3.

a. n = 20; p = .40; binomial

Use binomial table; P(x = 10) = .1171

b. P(7 < x < 12) = P(x = 8) + P(x = 9) + P(x = 10) + P(x = 11)

= .1797 + .1597 + .1171 + .0710

= .5275

c. P(x ≥ 12) = P(x = 12) + P(x = 13) + ... P(x = 20)

= .0355 + .0146 + .0049 + .0013 + .0003

= .0566

(8)

5.5.

a. $C_x^n = \frac{n!}{x!(n-x)!} = \frac{8!}{4!(8-4)!} = 70$ ways

b. $C_6^{10} = \frac{10!}{6!(10-6)!} = 210$ ways

c. $C_3^{10} = \frac{10!}{3!(10-3)!} = 120$ ways

5.17.
n=8, p=0.37

a. P(x=5) = 0.0971

b. P(x<4) = P(x≤3) = 0.6626

c. P(x>2) = P(x≥3) = 1 – P(x≤2) = 1 – 0.3811 = 0.6189

5.28.

a. P(2≤x≤5) = P(x≤5) – P(x≤1) = 0.8576 – 0.1359 = 0.7217

b. P(x=3) = 0.2158

c. P(x≥1) = 1 – P(x=0) = 1 – 0.0302 = 0.9698

5.47.

$\mu = 10.5$

$\sigma = \sqrt{16.7} = 4.087$

a. $z = 3.00$; $z = \frac{x-\mu}{\sigma}$; $3.0o = \frac{x-10.5}{4.087}$; x = 22.76

b. $-1.96 = \frac{x-10.5}{4.087}$; x = 2.49

5.55.

a. P(x>85) = P(z > (85 – 72)/4) = P(z > 3.25) = essentially 0 percent

b. P(x>82) = P(z > (82 – 72)/4) = P(z > 2.5) = 0.5 – 0.4938 = 0.0062

c. The number of fliers is actually a discrete variable. The normal distribution assumes that the value can take on an infinite number of possible outcomes.

9

5.69.

a. P(x>50) = (60-50)/(60-20) = 0.25

b. P(x = 45) = 0; you cannot find the probability of a specific value in a continuous distribution.

c. P(25<x<35) = (35.25)/(60-20) = 0.25

d. P(x<34) = (34-20)/(60-20) = 0.35

6.3.

a. $$\mu = \frac{\sum x}{N} = \frac{1,954}{11} = 177.64$$

$$\bar{x} = \frac{\sum x}{n} = \frac{650}{3} = 216.67$$

Sampling error = 216.67 - 177.64 = 39.03

b. The smallest 3 values that could be in a sample from the population are:

100 100 105 $$\bar{x} = \frac{\sum x}{n} = \frac{305}{3} = 101.67$$

Sampling error = 101.67 - 177.64 = -75.97

The largest 3 values that could be in a sample from the population are:

400 330 200 $$\bar{x} = \frac{\sum x}{n} = \frac{930}{3} = 310.0$$

Sampling error = 310.0 – 177.64 = 132.36

The range in potential sampling error is -75.97 to 132.36

c. The smallest 5 values are:

100 100 105 120 129 $$\bar{x} = \frac{\sum x}{n} = \frac{554}{5} = 110.8$$

Sampling error = 110.8 – 177.64 = -66.84

The largest 5 values are:

400 330 200 190 150 $$\bar{x} = \frac{\sum x}{n} = \frac{1.270}{5} = 254.0$$

Sampling error = 254.0 – 177.64 = 76.36

It is seen that the range of potential sampling error is greater when the sample is size five compared with a sample of size 3. In general, increasing the sample size will reduce the range for potential sampling error.

(10)

6.7. $\bar{x} = \frac{\sum x}{n} = \frac{170.6}{10} = 17.06$

Sampling error = 17.06 – 16.9 = 0.16 gallon

6.11.

$\mu = 1{,}000 \quad \sigma = 200$

a. $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{970 - 1{,}000}{\frac{200}{\sqrt{5}}} = -.34$; P(z < -.34) = .50 - .1331 = .3669

b. $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{970 - 1{,}000}{\frac{200}{\sqrt{10}}} = -.47$; P(z < -.47) = .50 - .1808 = .3192

c. The probability of extreme sampling error is reduced when the sample size is increased since the sampling distribution is less variable.

6.31.

a. $\sigma = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{.65(1-.65)}{100}} = .0477$

$z = \frac{\bar{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$; $z = \frac{.63 - .65}{\sqrt{\frac{.65(1-.65)}{100}}} = -.42$; P(z < -.42) = .50 - .1628 = .3372

b. $z = \frac{.63 - .65}{\sqrt{\frac{.65(1-.65)}{200}}} = -.59$; P(z < -.59) = .50 - .2224 = .2776

When the sample size is increased, the spread in the sampling distribution is reduced making the probability of a value between any two point less than it would have been for a smaller sample size. Less chance of extreme sampling error.

6.35.

a. $z = \frac{.42 - .40}{\sqrt{\frac{.40(1-.40)}{1000}}} = 1.29$; P(z < 1.29) = .50 + .4015 = .9015

b. $z = \frac{.44 - .40}{\sqrt{\frac{.40(1-.40)}{1000}}} = 2.58$; P(z > 2.58) = .50 - .4951 = .0049

(11)

7.9.

Since we don't know the population standard deviation and since the sample size is small, we must assume that the population is approximately normally distributed.

The interval estimate is given by: $\bar{x} \pm t\frac{s}{\sqrt{n}}$; $92.2 \pm 2.2622(15.562/\sqrt{10})$

92.2 ± 11.1326; 81.0674 ----------- 103.3326

7.15.

a. $1.2 \pm 1.96(0.5/\sqrt{200})$; 1.1307 ----- 1.2693

b. $1.2 \pm 1.645(0.5/\sqrt{200})$; 1.1418 ----- 1.2582

c. The interval in b is more precise and because it is narrower it may be more useful in decision making. However, intervals formed with a 0.90 confidence coefficient contain the true mean a smaller proportion of the time.

7.27.

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = 246.667;\ n = \frac{z^2s^2}{e^2} = \frac{(1.645^2)(246.667^2)}{60^2} = 45.73 = 46$$

Note, the 10 values in the pilot sample can be used leaving an additional 36 items needed in the sample.

7.36.

a. $n = (2.58)^2(21)^2/(5)^2 = 116.96$ or 117

b. $n = (1.645)^2(21)^2/(5)^2 = 47.7 = 48$

c. $n = (1.96)^2(21)^2/(8)^2 = 26.5 = 27$

7.37.

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = 1.91;\ n = \frac{z^2s^2}{e^2} = \frac{(2.33^2)(1.91^2)}{.10^2} = 1,,980.5 = 1,981$$

7.43.

$\bar{p} = 55/300 = .1833$; $0.1833 \pm 1.645(\sqrt{[(0.1833)(1-0.1833)]/300})$; $.1833 \pm .0367$;

.1466 ----------- .2200

7.55.

a. Use 0.50 for the population proportion to provide a conservatively large sample sample size.

$$n = \frac{z^2p(1-p)}{e^2} = \frac{1.96^2(.50)(1-.50)}{.03^2} = 1,067.1 = 1,068$$

b $\bar{p} = .18$; $0.18 \pm 1.96(\sqrt{[(0.18)(1-0.18)]/1,068})$; $.18 \pm .0230$; .1570 ----- .2030

c. The reason is that the sample size was computed based on the assumption that the population proportion is 0.50. However, the interval was computed using the sample proportion equal to 0.18. This implies that the population is less variable than assumed when the sample size was computed. They would have been wise to have selected a pilot sample to get a feel for what the population proportion might be before taking a sample as large as 1,068.

(12)

8.1.

a.

$H_o: \mu \le 20$
$H_A: \mu > 20$

b.

$H_o: \mu = 50$
$H_A: \mu \ne 50$

c.

$H_o: \mu \ge 35$
$H_A: \mu < 35$

d.

$H_o: \mu \le 87$
$H_A: \mu > 87$

e

$H_o: \mu \le 6$
$H_A: \mu > 6$

8.5.

a. Even though the population standard deviation is unknown, since n = 100 is large, we can use the standard normal distribution to obtain the critical value.

$\bar{x}_\alpha = 4{,}000 - 1.645(205/\sqrt{100})$; $\bar{x}_\alpha = 3966.2775$

If $\bar{x} < 3966.2775$ reject H_o
If $\bar{x} \ge 3966.2775$ do not reject H_o

If p-value < .05, reject H_o
If p-value $\ge$ .05, do not reject H_o

b. Since 3980 > 3966.2775 do not reject H_o

p-value = P(z < -.9756) = .50 - .3365 = .1635

Since p-value = .1635 $\ge$ 0.05 do not reject H_o

c. The two research hypotheses that could have produced the null and alternative hypotheses are:

The population mean is less than 4,000.
The population mean is at least 4,000.

8.9.

a. p-value = 2 * P(z > 2.36) = 2 * (.50 - .4909) = 0.0182

b. p-value = P(z < -1.85) = .50 - .4678 = 0.0322

c. p-value = P(z > .84) = .50 - .2995 = 0.2005

d. p-value = P(z > -2.06) = .50 + .4803 = 0.9803

8.23.

a. $z_\alpha = 1.645;\ p + z\left(\sqrt{\frac{p(1-p)}{n}}\right) = 0.40 + 1.645\left(\sqrt{\frac{0.40(1-0.40)}{150}}\right) = 0.4658$

b. $z_\alpha = -1.28;\ p + z\left(\sqrt{\frac{p(1-p)}{n}}\right) = 0.70 - 1.28\left(\sqrt{\frac{0.70(1-0.70)}{200}}\right) = 0.6585$

c. $z_\alpha = 1.645;\ p \pm z\left(\sqrt{\frac{p(1-p)}{n}}\right) = 0.85 \pm 1.645\left(\sqrt{\frac{0.85(1-0.85)}{100}}\right) = 0.7913$ and 0.9087

8.27.

a. $\bar{p}_\alpha = p + z\sqrt{\frac{p(1-p)}{n}} = .24 + 1.645\sqrt{\frac{.24(1-.24)}{100}}$; .3103

If p > .3103, reject the null hypothesis
Otherwise, do not reject

Since .27 < .3103, do not reject the null hypothesis

b. $z = \frac{\bar{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{.27 - .24}{\sqrt{\frac{.24(1-.24)}{100}}} = .7024$

If z > 1.645, reject the null hypothesis
Otherwise, do not reject

Since .7024 < 1.645, do not reject

8.31.

a. H_0: $p \geq 0.01$
H_a: $p < 0.01$
$\bar{p} = 6/800 = 0.0075$
$z = (0.0075 - 0.01) / \sqrt{(0.01)(1-0.01)/800} = -0.7107$

Decision Rule:

If z < -1.645 reject H_0, otherwise do not reject

Since z = -0.7107 > -1.645 do not reject and conclude that the percentage of lost luggage is 1% or more.

b.

$0.0075 \pm 1.96(\sqrt{(0.0075)(1-0.0075)/800})$; .0075 ± .006 ; 0.0015 --------- 0.0135