

Statistics Exam Review

CHAPTER 1

Name: Key

Write the letter of the response that best answers each problem.

1. A consumer research company wants to estimate the average cost of an airline ticket for a round trip from New York to Los Angeles. A random sample of 50 airfares had an average price of \$438. Identify the variable.

- (a) Simple random sample (b) Airline fare
(c) Consumer research company (d) City (e) \$438

1. B

2. For the information in parts A. through E., choose the highest level of measurement (or cannot determine):

- (a) Ratio (b) Interval
(c) Ordinal (d) Nominal (e) Cannot determine

- A. Temperature inside refrigerators
B. Horsepower of racecar engines
C. Marital status of school board members
D. Rating of television programs (poor, fair, good, excellent)
E. Age of children enrolled in a daycare

2. A. B
B. A
C. D
D. C
E. A

3. Consider the following study:

Students in a limnology class took water samples from a lake to determine the temperature at different depths. Which of the following techniques for gathering data do you think was used?

- (a) Double-blind experiment (b) Experiment
(c) Systematic sampling (d) Simulation (e) Observational study

3. E

4. Label each of the following numerical measures as a statistic or a parameter.

- (a) The average score on the GRE for all U.S. students
(b) The average score on the GRE for all Rutgers applicants
(c) The average score on the GRE for a random sample of California residents

4. (a) Statistic
(b) Parameter
(c) Statistic

5. Identify each of the following samples by naming the sampling technique used.

- (a) Cluster (b) Convenience
 (c) Simple random (d) Stratified (e) Systematic

A. Every tenth customer entering a health club is asked to select his or her preferred method of exercise.

5. A. e

B. Group the subscribers of a magazine into three different income categories and then select a random sample from each group. Ask each subscriber to name their favorite feature.

B. d

C. Take a sample of six zip codes from the Minneapolis metropolitan region and use all the car dealerships in the selected areas. Determine the number of new cars sold each month at each dealership.

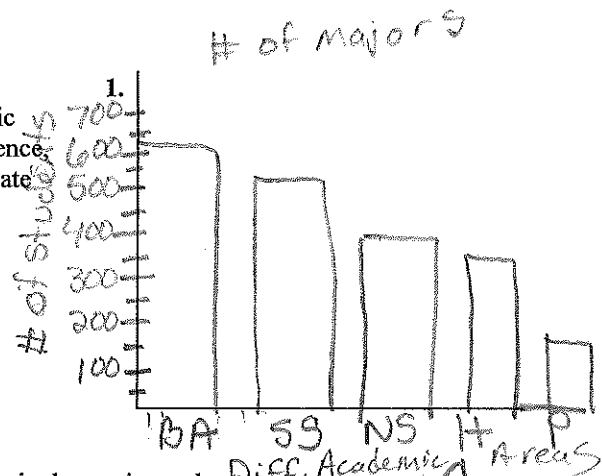
C. a

D. Use a random number table to select a sample of books at the library and determine the number of pages in each book.

D. c

CHAPTER 2

1. The Dean's Office at Hendrix College gave the following information about numbers of majors in different academic areas: Humanities, 372; Natural Science, 415; Social Science, 511; Business Administration, 619; Philosophy, 196. Create a Pareto chart representing this information.



2. _____ identify the frequency of events or categories in decreasing order of frequency of occurrence.

2. c

- (a) Time-series graphs (b) Bar graphs (c) Pareto charts
 (d) Stem-and-leaf displays (e) Circle graphs

3. _____ are useful for quantitative or qualitative data. With qualitative data, the frequency or percentage of occurrence can be displayed. With quantitative data, the measurement itself can be displayed.

3. b

- (a) Time-series graphs (b) Bar graphs (c) Pareto charts
 (d) Stem-and-leaf displays (e) Circle graphs

4. Professor Hill in the Music Department kept a list of the number of students visiting his office each week for two semesters (30 weeks). The results were

$\triangle 15$ $\square 23$ $\triangle 17$ $\circ 13$ $\square 3$ $\circ 9$ $\square 7$ $\square 6$ $\square 8$ $\circ 11$
 $\triangle 16$ $\square 32$ $\triangle 27$ $\square 4$ $\square 20$ $\square 3$ $\triangle 28$ $\square 5$ $\square 6$ $\circ 11$
 $\triangle 20$ $\circ 12$ $\square 8$ $\circ 10$ $\square 25$ $\circ 10$ $\square 8$ $\triangle 15$ $\circ 11$ $\circ 9$

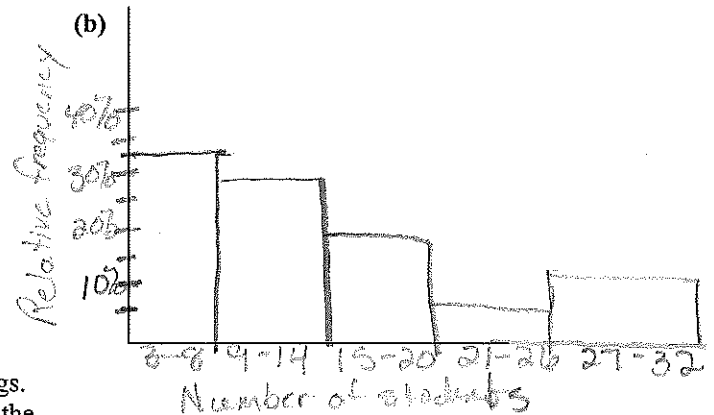
$$\frac{29}{5} = 5.8 \text{ CW}$$

- (a) Make a frequency table with five classes, showing class boundaries, class midpoints, frequencies, relative frequencies, and cumulative frequencies.

	Class Boundaries	Midpoint	Frequency	Relative Frequency	Cumulative Frequency
2.5 - 8.5	3 - 8		10	$\frac{10}{30} = \frac{1}{3} \approx 33.3\%$	33.3%
8.5 - 14.5	9 - 14		9	$\frac{9}{30} = \frac{3}{10} = 30\%$	30%
14.5 - 20.5	15 - 20		6	$\frac{6}{30} = \frac{1}{5} = 20\%$	20%
20.5 - 26.5	21 - 26		2	$\frac{2}{30} = \frac{1}{15} \approx 6.67\%$	6.67%
26.5 - 32.5	27 - 32		3	$\frac{3}{30} = \frac{1}{10} = 10\%$	10%

Students visit per week for 30 weeks

- (b) Draw a relative frequency histogram from the information in part (a).



5. Jim is a taxi driver who keeps a record of his meter readings. The results for the past twenty meter readings (rounded to the nearest dollar) are given below.

~~15~~ ~~10~~ ~~9~~ ~~21~~ 19 17 ~~8~~ 35 ~~22~~ 33
 46 ~~5~~ 24 ~~37~~ 51 49 57 42 ~~12~~ 16

Make a stem-and-leaf display of the data.

5. Past 20 meter Readings

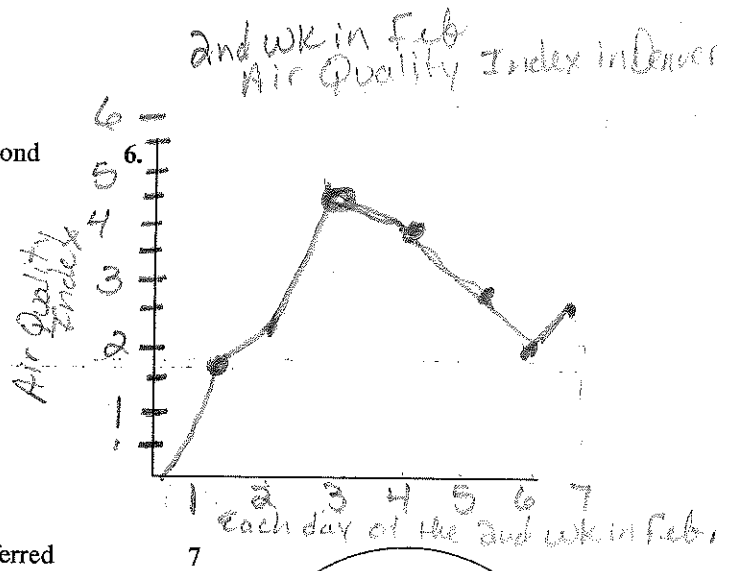
```

0 | 5, 7, 8, 9
1 | 2, 5, 6, 7, 9
2 | 1, 2, 4
3 | 3, 5, 7
4 | 2, 6, 9
5 | 1, 7
  
```

6. The Air Quality Index in Denver for each day of the second week of February is shown below.

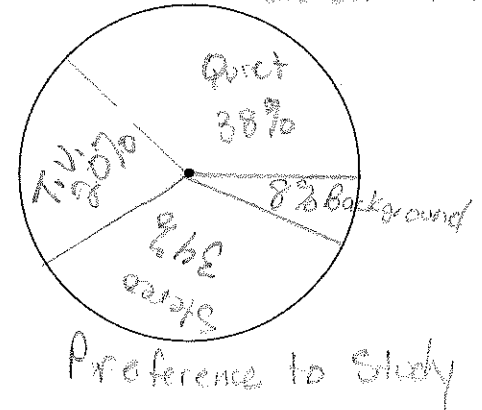
1.7 2.4 5.3 4.1 3.2 2.0 2.5

Make a time-series graph for these data.



7. A survey of 100 students was taken to see how they preferred to study. The survey showed that 38 students liked it quiet, 20 students liked the television on, 34 students liked the stereo on, and 8 students liked background noise such as in a coffee shop. Make a circle graph to display this information.

Quiet	38%	137°
T.V.	20%	72°
Stereo	34%	122°
B.G.	8%	29°



CHAPTER 3

1. A random sample of 18 airline carry-on luggage bags had the following weights (rounded to the nearest pound).

12 25 10 38 12 19 8 12 17
41 7 22 10 19 12 16 5 14

Find the mean, median, and mode of these weights.

1. $\bar{x} = 16.61$ Med = 13 mode = 12

2. A sample of 7 northern pike from Taltson Lake had the following lengths (rounded to the nearest inch).

21 27 46 35 41 36 25

- (a) Find the range.
(b) Find the sample mean.
(c) Find the sample variance.
(d) Find the sample standard deviation.

2. (a) 25
(b) 33
(c) 81.667
(d) 9.037

3. A random sample of receipts for individuals eating at an Applebee's Restaurant had a sample mean of $\bar{x} = \$10.38$ and a sample standard deviation $s = \$2.17$.

- (a) Compute the coefficient of variation for this data.
 (b) Use Chebyshev's theorem to find the smallest interval centered on the mean in which we can expect at least 75% of meal receipts to fall.

$$10.38 - 2(2.17) = 6.04$$

$$10.38 + 2(2.17) = 14.72$$

3. (a) $\frac{2.17}{10.38} \cdot 100 = 20.91\%$

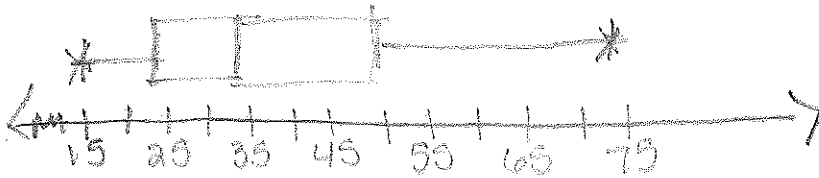
(b) $6.04 - 14.72$

4. A random sample of 27 skiers at a Vail, Colorado ski resort gave their ages. The results were

18	25	32	16	41	52	29	58	23
<u>62</u>	47	56	19	22	38	15	46	33
49	52	37	26	72	44	19	24	29

- (a) Give the five number summary.
 (b) Create a box-and-whisker plot for the given data.

Vail, Colorado skiers Ages



- (c) Find the interquartile range.

4. (a) $\text{Min } 15, Q_1 23, \text{Med } 33, Q_3 49, \text{Max } 72$

(b)

(c) $49 - 23 = 26$

Statistics Exam Review 2
CHAPTER 4

Name: Key

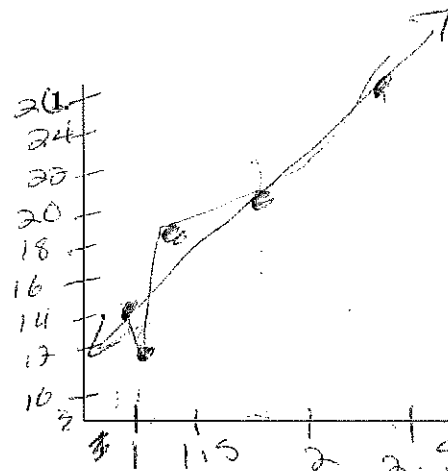
For the given data, solve the following problems.

Do higher paid chief executive officers control bigger companies? Here, x = annual CEO salary (\$ millions) and y = annual company revenue (\$ billions). The following data are based on information from *Forbes* magazine and represents a sample of top US executives.

x (\$ millions)	0.8	1.0	1.1	1.7	2.3
y (\$ billions)	14	11	19	20	25

1. Create a scatter diagram. Using only the scatter diagram, would you estimate the correlation coefficient to be positive, close to zero, or negative? Explain your answer.

Positive



2. For the given data compute each of the following.

(a) \bar{x} and \bar{y}

2. (a) $\bar{x} = 1.38$ $\bar{y} = 17.8$

(b) $\sum x$, $\sum y$, $\sum x^2$, $\sum y^2$, $\sum xy$

(b) $\sum x = 6.9$ $\sum y = 89$ $\sum x^2 = 11.03$

- (c) The slope b and y -intercept a of the least squares line; write out the equation for the least squares line.

$\sum y^2 = 1703$ $\sum xy = 134.6$
 $b = 7.8$ $a = 7.02$

- (d) Graph the least squares line on your scatter plot from problem 1.

(d) _____

3. Compute the sample correlation coefficient r . Compute the coefficient of determination. Give a brief explanation of the meaning of the coefficient of determination in the context of this problem.

3. $r = 0.880$ $r^2 = 0.77570$

4. If a CEO has an annual salary of \$1.5 million, what is his or her annual company revenue as predicted by the least squares line?

4. $x = 1.5$ millions
 $y = 18.735$ billion

Write the letter of the response that best answers each problem.

Does the weight of a vehicle affect the gas mileage? The following random sample was collected where x = weight of a vehicle (in hundreds of pounds) and y = miles per gallon.

x (100 lbs)	26	35	29	39	20
y (mpg)	22.0	16.1	18.8	15.7	23.4

1 5. Based on a scatter diagram and the context of the problem, would you estimate the linear correlation coefficient to be

- (a) close to -1 (b) close to 0 and negative
(c) close to 1 (d) close to 0 and positive (e) Cannot determine

5. A

2 6. What is the equation for the least squares line?

- (a) $y = -32.55x + 0.448$ (b) $y = -32.55x - 0.448$
(c) $y = -32.55x + 0.448$ (d) $y = -0.448x + 32.55$ (e) $y = 0.448x - 32.55$

6. D

3 7. Compute the coefficient of determination.

- (a) -0.941 (b) 0.941
(c) -0.970 (d) 0.970 (e) 0.965

7. B

4 8. If a vehicle weighs 2200 pounds, what does the least-squares line predict for the miles per gallon?

- (a) 22.7 (b) 66.0
(c) 42.4 (d) 22.0 (e) Cannot determine

8. A

CHAPTER 5

- A random sample of 317 mp3 music players had 19 that were defective.
 - How would you estimate the probability that a new mp3 player is defective? What is your estimate?
 - What is your estimate for the probability that a new mp3 player is not defective?
 - Either an mp3 player is defective or not. What is the sample space in this problem. Do the probabilities assigned to the sample space add up to one?

1. (a) $\frac{19}{317} = .0599 = 6\%$

(b) $1 - .0599 = 94.01\%$

(c) Defective, Not Defective
 Ups! $94.01 + 5.99 = 100\%$

- If you roll a single fair die and count the number of dots on top, what is the probability of getting a number less than 3 on a single throw?

2. $\frac{2}{6} = \frac{1}{3}$

- You roll two fair dice, a blue one and a yellow one.

- Find $P(\text{even number on the blue die and 3 on the yellow die})$.
- Find $P(3 \text{ on the blue die and even number on the yellow die})$.
- Find $P(\text{even number on the blue die and even number on the yellow die})$.

3. (a) $\frac{1}{12}$

(b) $\frac{1}{12}$

(c) $\frac{1}{4}$

- An urn contains 12 balls identical in every respect except color. There are 3 red balls, 7 green balls, and 2 blue balls.

- You draw two balls from the urn, but replace the first ball before drawing the second. Find the probability that the first ball is red and the second is green.
- Repeat part (a), but do not replace the first ball before drawing the second.

4. (a) $\frac{2}{144} = 0.149$

(b) $\frac{2}{132} = 0.159$

- Robert is applying for a bank loan to open up a pizza franchise. He must complete a written application, and then be interviewed by bank officers. Past records for this bank show that the probability of being approved on the written application is 0.63. The probability of being approved by the interview committee is 0.85, given the candidate has been approved on the written application. What is the probability Robert is approved on both the written application and the interview?

5. $(0.63)(0.85) =$
 0.5355
 53.55%

6. A hair salon completed a survey of 360 customers about satisfaction with service and type of customer. A walk-in customer is one who has seen no ads and not been referred. The other customers either saw a TV ad or were referred to the salon (but not both). The results follow.

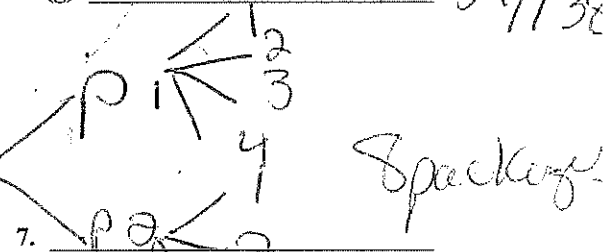
	Walk-In	TV Ad	Referred	Total
Not Satisfied	21	9	5	35
Neutral	18	25	37	80
Satisfied	36	43	59	138
Very Satisfied	28	31	48	107
Total	103	108	149	360

Assume the table represents the entire population of customers. Find the probability that a customer is

- (a) Not satisfied
 (b) Not satisfied and a walk-in
 (c) Not satisfied, given referred
 (d) Referred
 (e) Very satisfied, given referred
 (f) Very satisfied and saw a TV ad
 (g) Are the events not satisfied and referred independent? Explain your answer.

6. (a) $\frac{35}{360}$
 (b) $\frac{21}{360}$
 (c) $\frac{5}{149}$
 (d) $\frac{149}{360}$
 (e) $\frac{48}{107}$
 (f) $\frac{31}{360}$
 (g) $\frac{149}{360} \neq \frac{59}{138}$

7. A computer package comes with two different choices of printers and four different choices of monitors. If a store wants to display each package combination that is for sale, how many packages must be displayed? Make a tree diagram showing the outcomes for selecting printer and monitor.



8. In how many ways can the 40 members of a 4H club select a president, a vice-president, a secretary, and a treasurer?

8. $P_{40,4} = 2,193,360$

9. In how many different ways can a person choose three movies to see in a theater playing 11 movies.

9. $C_{11,3} = 165$

CHAPTER 6

1. Sam is a representative who sells large appliances such as refrigerators and stoves. Let x = number of appliances Sam sells on a given day. Let f = frequency (number of days) per which he sells x appliances. For a random sample of 240 days, Sam had the following sales record.

x	0	1	2	3	4	5	6	7
f	9	72	63	41	28	14	8	5

0.38, 0.3, 0.263, 0.171, 0.117, 0.058, 0.033, 0.021

Assume the sales record is representative of the population of all sales days.

- (a) Use the relative frequency to find $P(x)$ for $x = 0$ to 7.

- (b) Create a histogram of the probability distribution of part (a).

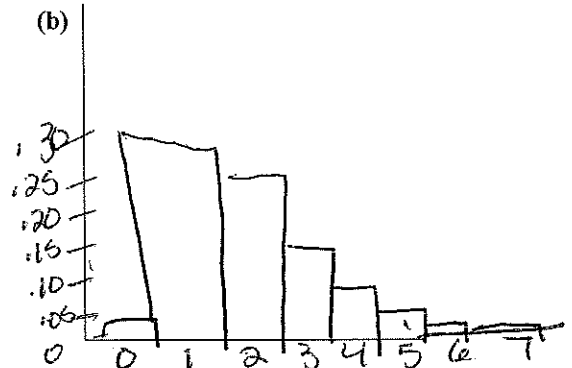
- (c) Compute the probability that x is between 2 and 5 (including 2 and 5).
 (d) Compute the probability that x is less than 3.
 (e) Compute the expected value of the x distribution.
 (f) Compute the standard deviation of the x distribution.

2. The director of a health club conducted a survey and found that 23% of members used only the pool for workouts. Based on this information, what is the probability that for a random sample of 10 members, 4 used only the pool for workouts?

3. Of those mountain climbers who attempt Mt. McKinley, only 65% reach the summit. In a random sample of 11 mountain climbers who attempt to climb Mt. McKinley, what is the probability of each of the following?

- (a) All 11 reach the summit.
 (b) At least 9 reach the summit.
 (c) No more than 3 reach the summit.
 (d) Exactly 9, 10, or 11 reach the summit.

1. (a) _____



(c) $P(2 \leq x \leq 5) = 0.609$

(d) $P(x < 3) = 0.601$

(e) $\mu = 2.44$

(f) $\sigma = 1.828$
 $\sigma = 1.574$

2. $P(4) = 0.122$

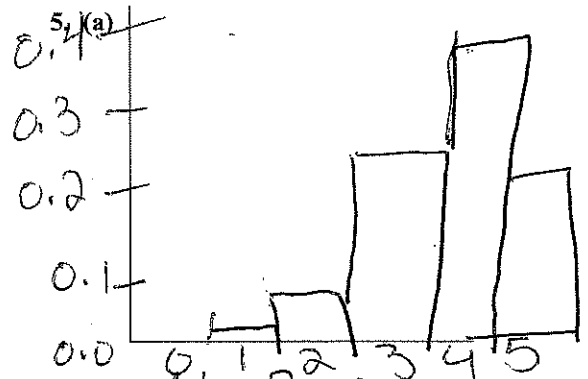
3. (a) $P(11) = 0.0088$
 (b) $P(F \geq 9) = 0.200$
 (c) $P(F \leq 3) = 0.012$
 (d) $P(9 \leq r \leq 11) = 0.200$

4. A coach found that about 12% of all hockey games end in overtime. What is the expected number of games ending in overtime if a random sample of 50 hockey games are played?

4. 6

5. The probability that a semi-truck exceeds the speed limit on I-80 between Cheyenne and Rock Springs, Wyoming is about 75%. Suppose a random sample of 5 trucks are observed.

(a) Make a histogram showing the probability that $r = 0, 1, 2, 3, 4, 5$ trucks are exceeding the speed limit.



(b) Find the mean μ of this probability distribution.

(b) $\mu = 3.75$

(c) Find the standard deviation σ of the probability distribution.

(c) $\sigma = 0.968$

6. Records show the probability of catching at least one northern pike over 40 inches at Taltson Lake (Canada) is about 15% for each full day a person spends fishing. A person spends 7 days fishing. Let x = the number of days the person catches at least one northern pike over 40 inches.

(a) Find the mean μ of this probability distribution.

6. (a) $\mu = 1.05$

(b) Find the standard deviation σ of the probability distribution.

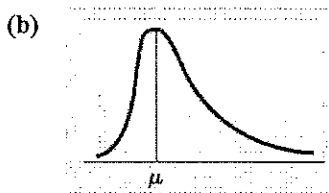
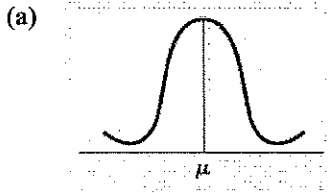
(b) $\sigma = 0.945$

(c) Using Chebyshev's theorem, would it be unusual for a person to catch a 40 inch northern pike on 4 of the 7 days?

(c) Yes since $4 > \mu + 2.5\sigma$
 $= 1.05 + (2.5)(0.945) =$
 3.4125

CHAPTER 7

1. Each of the following curves fails to be a normal curve. Give reasons why these curves are not normal curves.



1. (a) Tails should get closer to the x-axis

(b) Must be symmetrical

2. According to the empirical rule, for a distribution that is symmetrical and bell-shaped, approximately _____ of the data values will lie within one standard deviation on each side of the mean.

2. 68%

3. Assuming that the weights of newborn babies are normally distributed with mean 6.5 pounds and standard deviation 1.2 pounds, how many babies in a group of 80 do we expect to weigh more than 8.9 pounds?

3. 1.82 ≈ 2 babies

4. Let x be a random variable that represents the length of time it takes a student to write a term paper for Dr. Adam's sociology class. After interviewing many students, it was found that x has an approximate normal distribution with mean $\mu = 6.8$ hours and standard deviation $\sigma = 2.1$ hours.

Convert each of the following x intervals to standardized z intervals.

(a) $x \leq 7.5$

4. (a) $z \leq 0.33$

(b) $5 \leq x \leq 8$

(b) $-0.86 \leq z \leq 0.57$

(c) $x \geq 4$

(c) $z \geq -1.33$

Convert each of the following z intervals to raw score x intervals.

(d) $z \geq -2$

(d) $x \geq 2.6$

(e) $0 \leq z \leq 2$

(e) $6.8 \leq x \leq 11$

(f) $z \leq 3$

(f) $x \leq 13.1$

5. Operating temperatures of two models of portable electric generators follow a normal distribution. For generator I, the mean temperature is $\mu_1 = 148^\circ\text{F}$ with standard deviation $\sigma_1 = 25^\circ\text{F}$. For generator II, the mean temperature is $\mu_2 = 143^\circ\text{F}$ with standard deviation $\sigma_2 = 8^\circ\text{F}$. At peak power demand, generator I was operating at 166°F , and generator II was operating at 165°F .

(a) At peak power output, both generators are operating at about the same temperature. Relative to the operating characteristics, is one a lot hotter than the other? Explain.

(b) Convert the peak power temperature for each generator to standard z units. Then locate both z scores under a standard normal curve. Could one generator be near a melt down? Which one? Explain your answer.

6. Weights of a certain model of fully loaded gravel trucks follow a normal distribution with mean $\mu = 6.4$ tons and standard deviation $\sigma = 0.3$ ton. What is the probability that a fully loaded truck of this model is

(a) less than 6 tons?

(b) more than 7 tons?

(c) between 6 and 7 tons?

7. Quality control studies for Speedy Jet Computer Printers show the lifetime of the printer follows a normal distribution with mean $\mu = 4$ years and standard deviation $\sigma = 0.78$ years. The company will replace any printer that fails during the guarantee period. How long should Speedy Jet printers be guaranteed if the company wishes to replace no more than 10% of the printers?

5. (a) Generator II is hotter
See z values

(b) $z = 0.72$ Gen. I
 $z = 2.75$ Gen. II

Gen II could be near a melt down since it is hotter!

6. (a) $P(X < 6) = 0.0912$
(b) $P(X > 7) = 0.0228$
(c) $P(6 \leq X \leq 7) = 0.8860$

7. 3 years