The Normal Distribution

Objectives

After completing this chapter, you should be able to

1 Identify distributions as symmetric or skewed.

2 Identify the properties of a normal distribution.

3 Find the area under the standard normal distribution, given various z values.

4 Find probabilities for a normally distributed variable by transforming it into a standard normal variable.

5 Find specific data values for given percentages, using the standard normal distribution.

6 Use the central limit theorem to solve problems involving sample means for large samples.

7 Use the normal approximation to compute probabilities for a binomial variable.

Outline

Introduction

6–1 Normal Distributions

6–2 Applications of the Normal Distribution

6–3 The Central Limit Theorem

6–4 The Normal Approximation to the Binomial Distribution

Summary

Review

- **Chapter 2: Distribution of data**

- **Chapter 3: Measures of data sets, including measures of center and variation**

- **Chapter 4: Principles of probability**

- **Chapter 5: Discrete probability distributions**

- **Chapter 6: Continuous probability distributions**

6.1 Normal Distributions

- Many continuous variables have distributions that are bell-shaped and are called **approximately normally distributed variables**.

- The theoretical curve, called the **bell curve** or the **Gaussian distribution**, can be used to study many variables that are not normally distributed but are approximately normal.

Normal Distributions

The mathematical equation for the normal distribution is:

$$y = \frac{e^{-(X-\mu)^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}}$$

where

$e \approx 2.718$

$\pi \approx 3.14$

$\mu = population\ mean$

$\sigma = population\ standard\ deviation$

Normal distribution is determined by the fixed values of the mean μ and standard deviation σ.
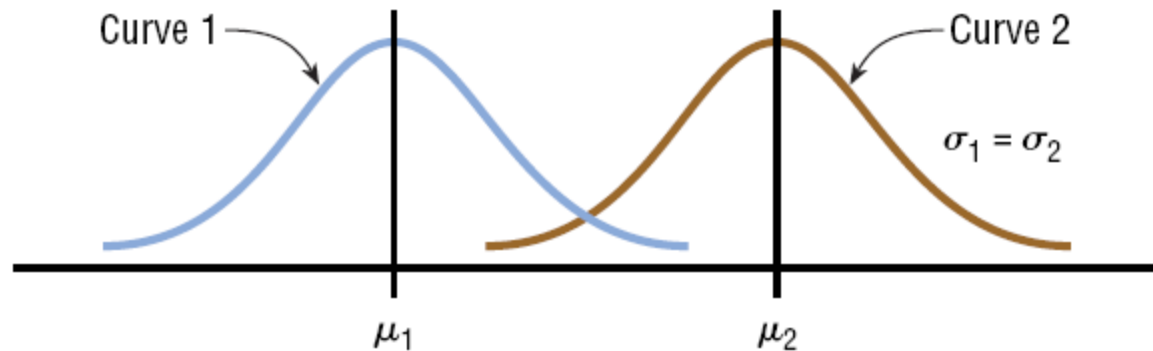
Normal Distributions

- The shape and position of the normal distribution curve depend on two parameters, the **mean** and the **standard deviation**.

- Each normally distributed variable has its own normal distribution curve, which depends on the values of the variable's mean and standard deviation.
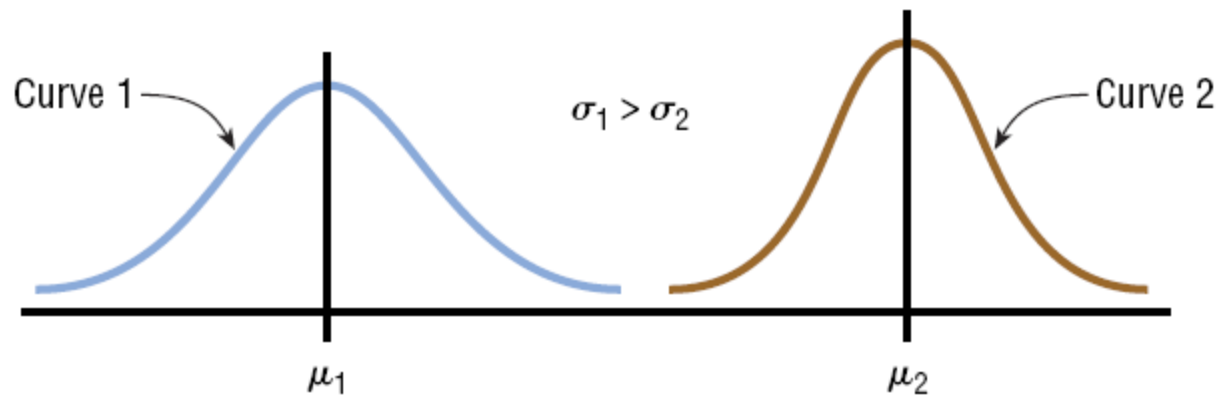

Curve 2

Curve 1

$\sigma_1 > \sigma_2$

$\mu_1 = \mu_2$
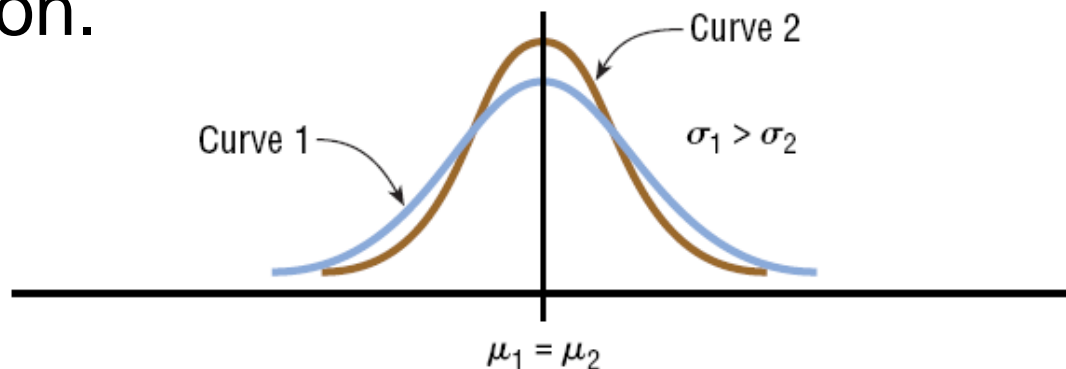
(a) Same means but different standard deviations

Normal Distributions

(b) Different means but same standard deviations

(c) Different means and different standard deviations

Normal Distribution Properties

- The normal distribution curve is bell-shaped.

- The mean, median, and mode are equal and located at the center of the distribution.

- The normal distribution curve is **unimodal** (i.e., it has only one mode).

- The curve is symmetrical about the mean, which is equivalent to saying that its shape is the same on both sides of a vertical line passing through the center.
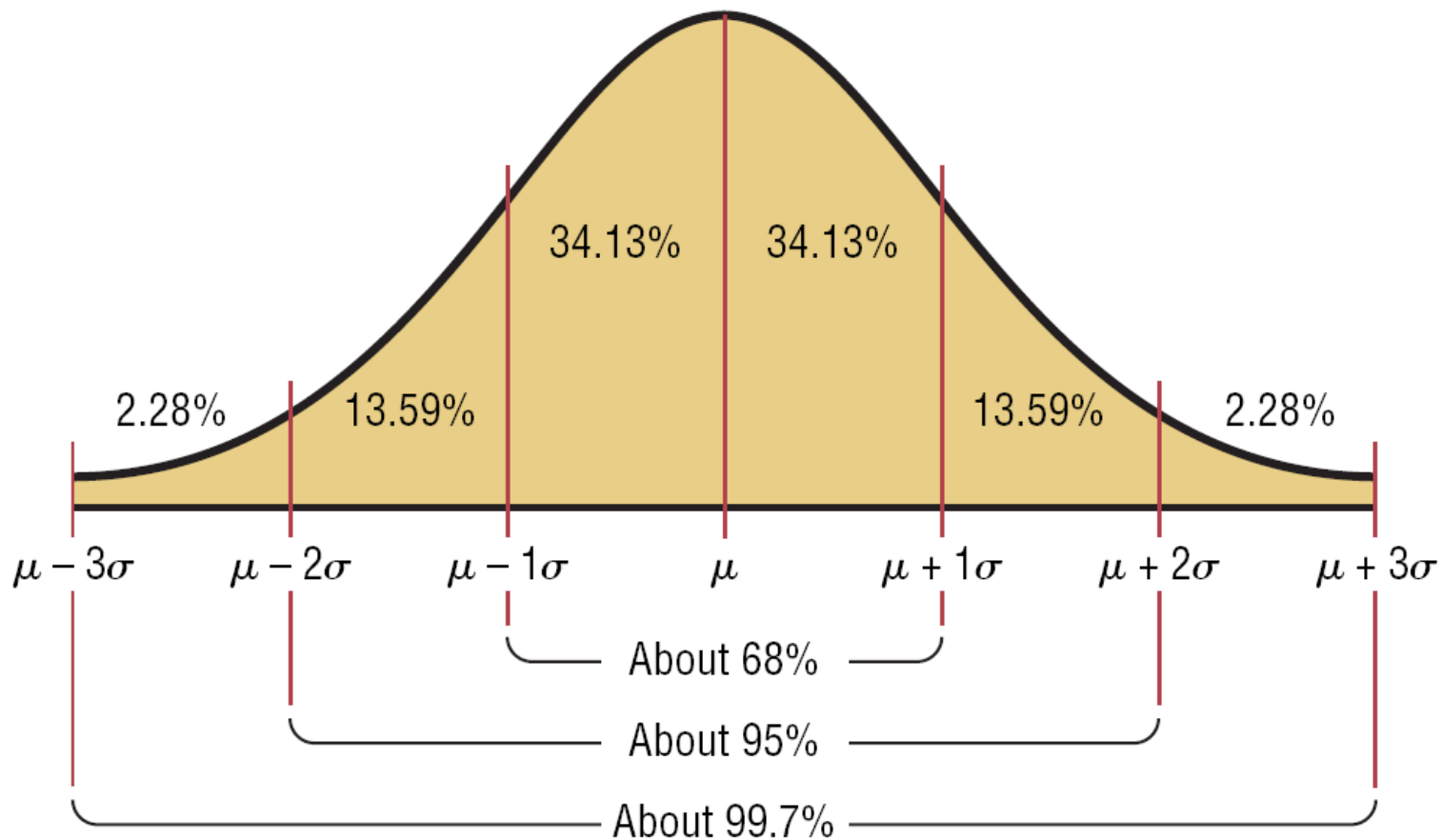
Normal Distribution Properties

- The curve is continuous—i.e., there are no gaps or holes. For each value of *X*, here is a corresponding value of *Y*.

- The curve never touches the *x* axis. Theoretically, no matter how far in either direction the curve extends, it never meets the *x* axis—but it gets increasingly closer.

Normal Distribution Properties

■ The total area under the normal distribution curve is equal to 1.00 or 100%.

■ The area under the normal curve that lies within

□ one standard deviation of the mean is approximately 0.68 (68%).

□ two standard deviations of the mean is approximately 0.95 (95%).

□ three standard deviations of the mean is approximately 0.997 (99.7%).

Normal Distribution Properties

Standard Normal Distribution

■ Since each normally distributed variable has its own mean and standard deviation, the shape and location of these curves will vary. In practical applications, one would have to have a table of areas under the curve for each variable. To simplify this, statisticians use the standard normal distribution.

■ The **standard normal distribution** is a normal distribution with a mean of 0 and a standard deviation of 1.

Develop the skill to find areas (or probabilities or relative frequencies) corresponding to various regions under the graph of the standard normal distribution. Find z-scores that correspond to area under the graph.
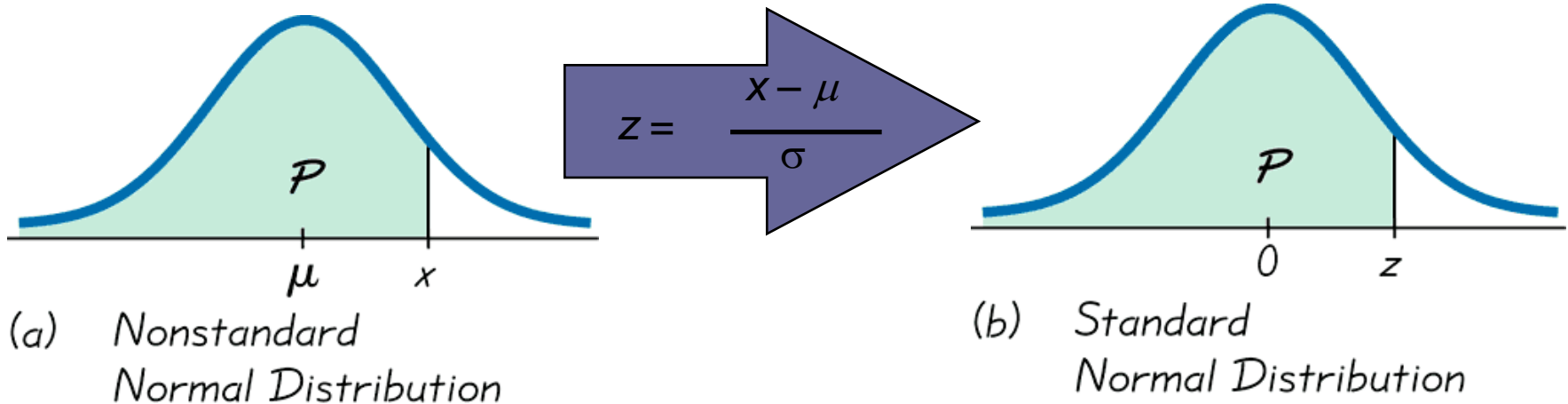
z value (Standard Value)

The *z* value is the number of standard deviations that a particular *X* value is away from the mean. The formula for finding the z value is:

$$z = \frac{\text{value - mean}}{\text{standard deviation}} \quad \text{vs.} \quad y = \frac{e^{-(X-\mu)^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}}$$

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

Round *z* scores to 2 decimal places

Converting to a Standard Normal Distribution

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

$\mathcal{P}$

μ x

(a)   Nonstandard Normal Distribution

$\mathcal{P}$

0 z

(b)   Standard Normal Distribution

Area under the Standard Normal Distribution Curve

1. To the left of any *z* value:

Look up the *z* value in the table and use the area given.
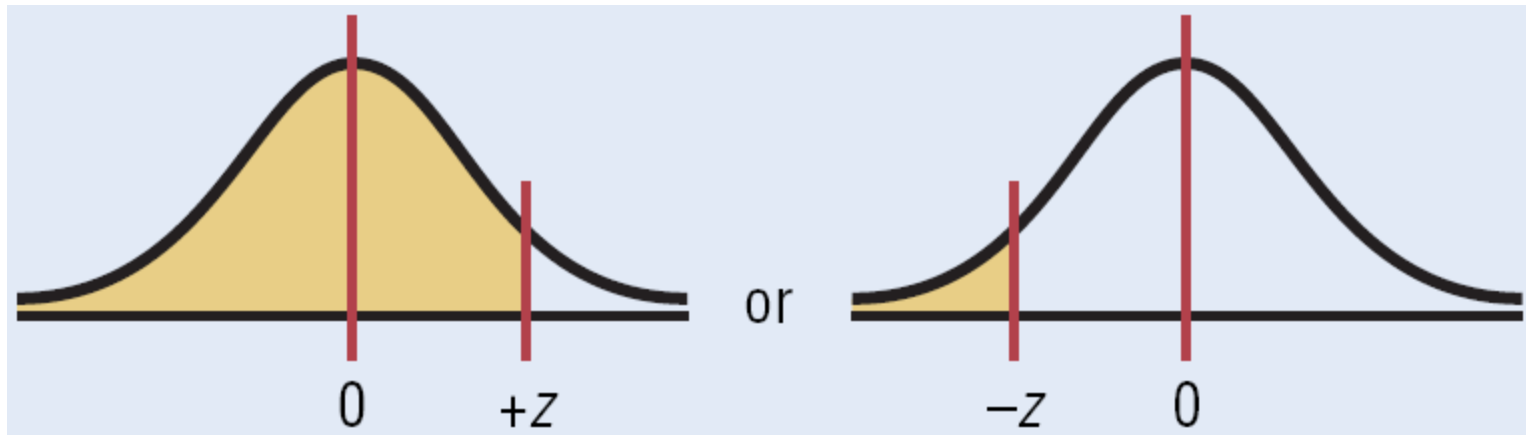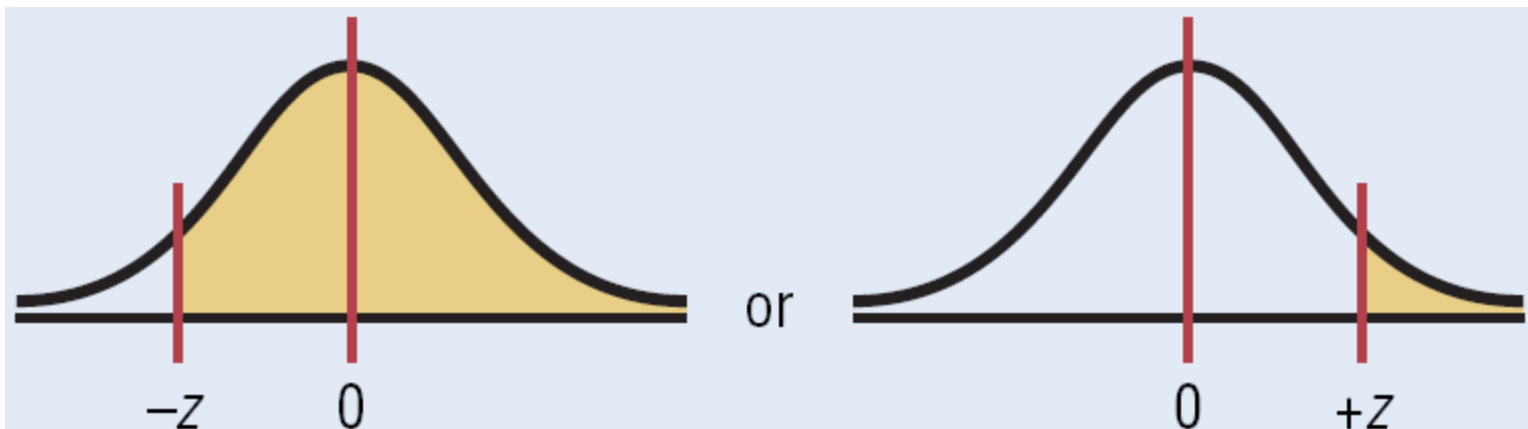
Area under the Standard Normal Distribution Curve

2. To the right of any *z* value:

Look up the *z* value and subtract the area from 1.

Area under the Standard Normal Distribution Curve

3. Between two *z* values:

Look up both *z* values and subtract the corresponding areas.

Table E — The Standard Normal Distribution

Cumulative Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233

Using Table E

1. It is designed only for the *standard* normal distribution, which has a mean of 0 and a standard deviation of 1.

2. It is on two pages, with one page for *negative z*-scores and the other page for *positive z*-scores.

3. Each value in the body of the table is a *cumulative area from the left* up to a vertical boundary above a specific *z*-score.

Using Table E

4. When working with a graph, avoid confusion between *z*-scores and areas.

    z Score
    Distance along horizontal scale of the standard normal distribution; refer to the leftmost column and top row of Table E.

    Area
    Region under the curve; refer to the values in the body of Table E.

5. The part of the z-score denoting hundredths is found across the top.

Notation

$$P(a < z < b)$$

denotes the probability that the *z* score is between *a* and *b.*

$$P(z > a)$$

denotes the probability that the *z* score is greater than *a.*

$$P(z < a)$$

denotes the probability that the *z* score is less than *a.*
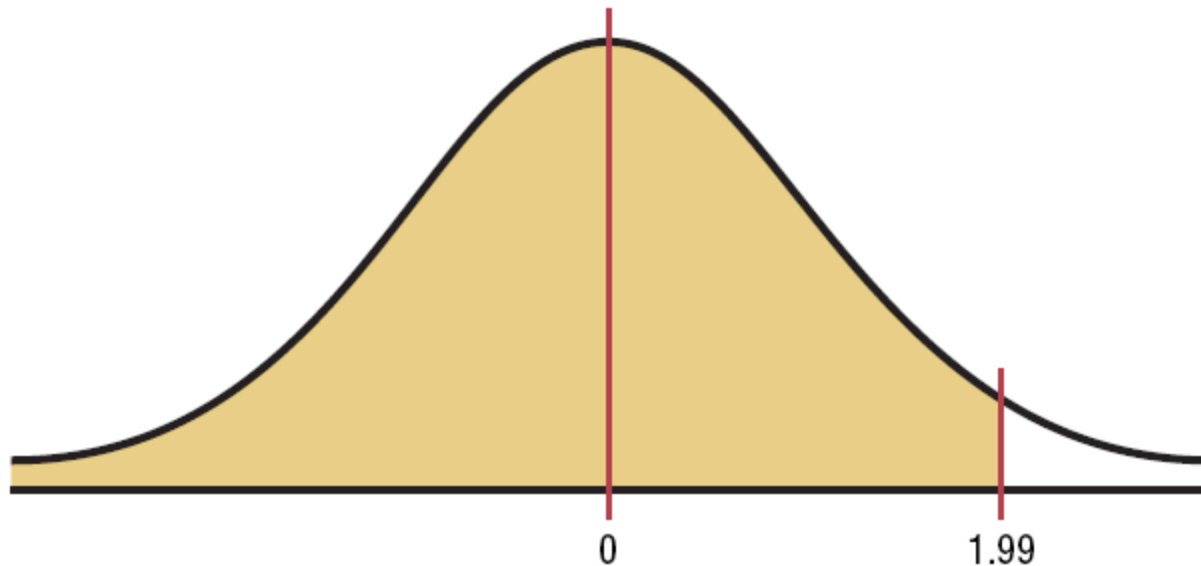
Chapter 6
Normal Distributions

Section 6-1

Example 6-1

Page #306

Example 6-1: Area under the Curve
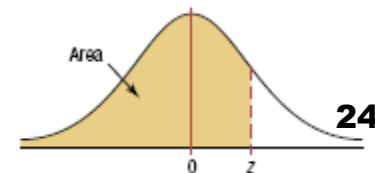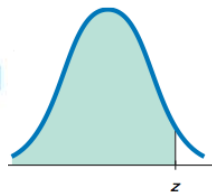
Find the area to the left of $z = 1.99$.

The value in the 1.9 row and the .09 column of Table E is .9767. The area is .9767.

Cumulative Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952

Gives the cumulative area from the left up to a vertical line above a specific value of z.

Bluman, Chapter 6, 03/2010

Area

Chapter 6
Normal Distributions

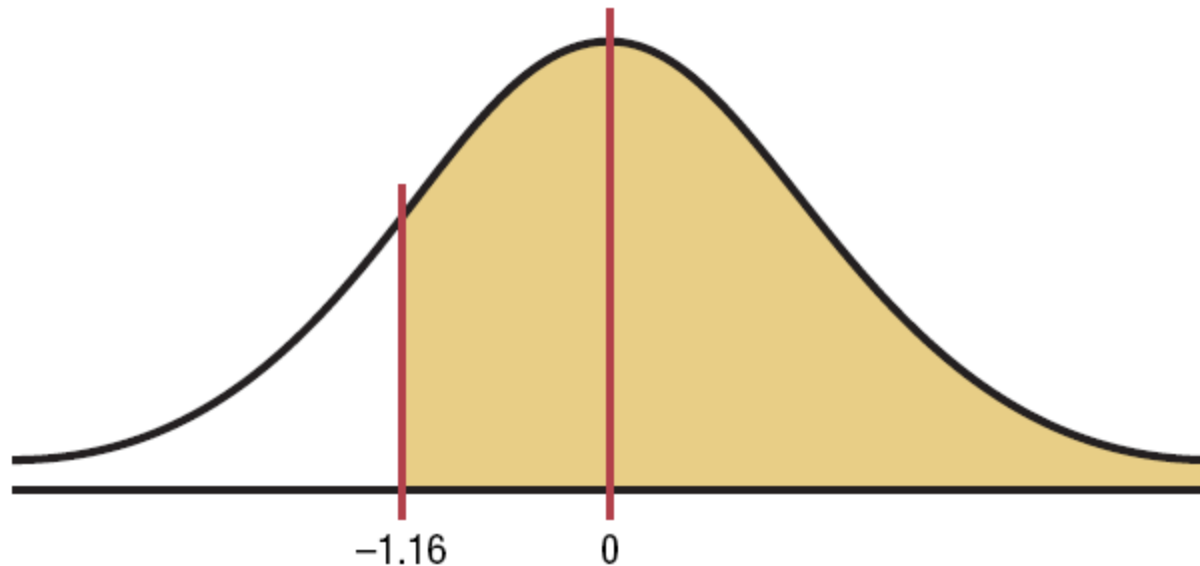# Section 6-1

Example 6-2

Page #306

Example 6-2: Area under the Curve
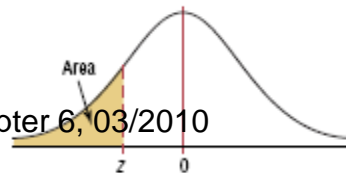
Find the area to right of $z = -1.16$.

The value in the -1.1 row and the .06 column of Table E is .1230. The area is 1 - .1230 = .8770.

Table E — The Standard Normal Distribution

Cumulative Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611

Area

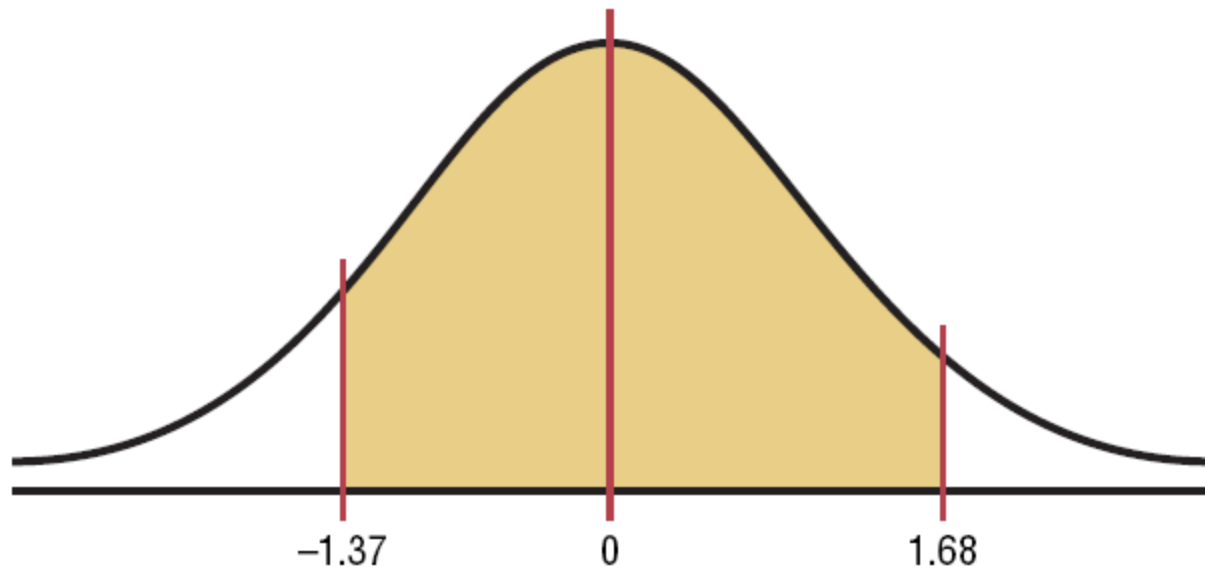# Chapter 6
Normal Distributions

Section 6-1

Example 6-3

Page #307

Example 6-3: Area under the Curve

Find the area between $z = 1.68$ and $z = -1.37$.

The values for $z = 1.68$ is .9535 and for $z = -1.37$ is .0853. The area is .9535 - .0853 = .8682.

Chapter 6
Normal Distributions

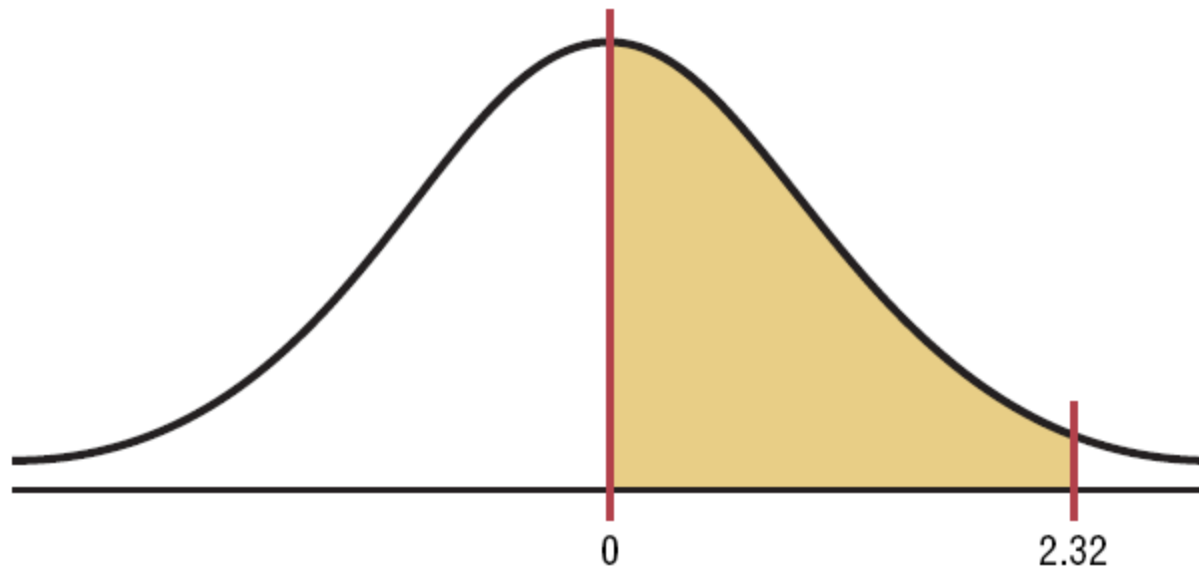A normal distribution curve can be used as a probability distribution curve for normally distributed variables.

Section 6-1

Example 6-4

Page #308

Example 6-4: Probability

a. Find the probability: $P(0 < z < 2.32)$

The values for $z = 2.32$ is .9898 and for $z = 0$ is .5000. The probability is .9898 - .5000 = .4898.
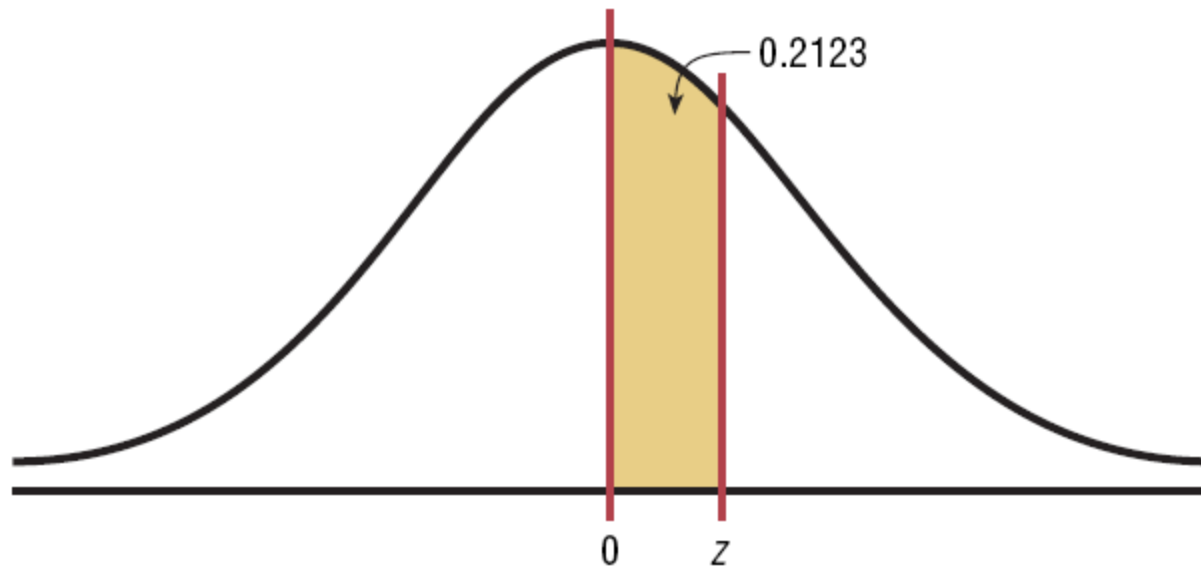
Chapter 6
Normal Distributions

Section 6-1

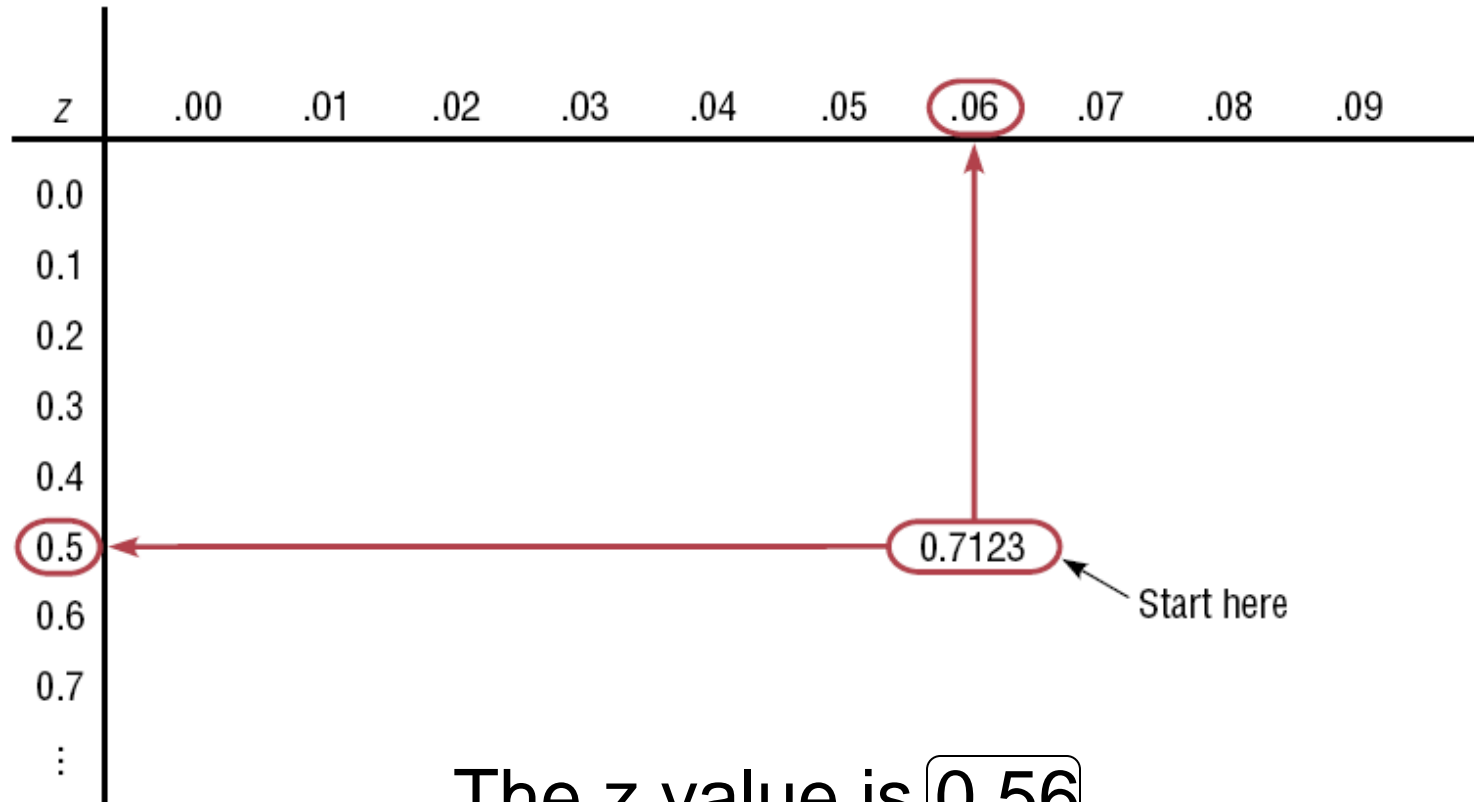## Example 6-5

Page #309

Example 6-5: Probability

Find the *z* value such that the area under the standard normal distribution curve between 0 and the z value is 0.2123.

Add .5000 to .2123 to get the cumulative area of .7123. Then look for that value inside Table E.

Example 6-5: Probability

Add .5000 to .2123 to get the cumulative area of .7123. Then look for that value inside Table E.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0										
0.1										
0.2										
0.3										
0.4										
0.5								0.7123		
0.6										
0.7										

Start here

The z value is 0.56.

6.2 Applications of the Normal Distributions

- The standard normal distribution curve can be used to solve a wide variety of practical problems. The only requirement is that the variable be normally or approximately normally distributed.

- For all the problems presented in this chapter, you can assume that the variable is normally or approximately normally distributed.

Applications of the Normal Distributions

- To solve problems by using the standard normal distribution, transform the original variable to a standard normal distribution variable by using the *z* value formula.

- This formula transforms the values of the variable into standard units or *z* values. Once the variable is transformed, then the Procedure Table and Table E in Appendix C can be used to solve problems.

Chapter 6
Normal Distributions

Section 6-2

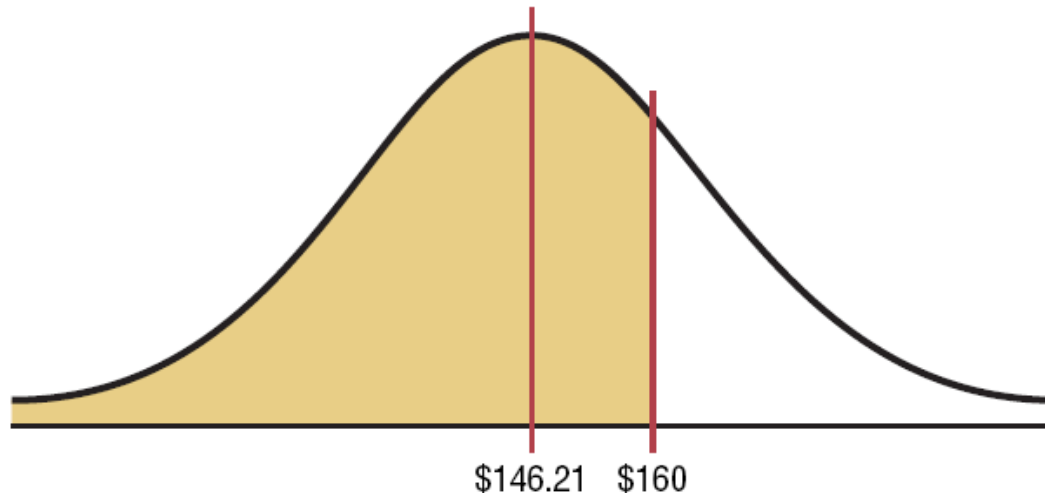Example 6-6

Page #317

Example 6-6: Holiday Spending

A survey by the National Retail Federation found that women spend on average $146.21 for the Christmas holidays. Assume the standard deviation is $29.44. Find the percentage of women who spend less than $160.00. Assume the variable is normally distributed.

Step 1: Draw the normal distribution curve.

$146.21 $160

Example 6-6: Holiday Spending

Step 2: Find the z value corresponding to $160.00.

$$z = \frac{X - \mu}{\sigma} = \frac{160.00 - 146.21}{29.44} = 0.47$$
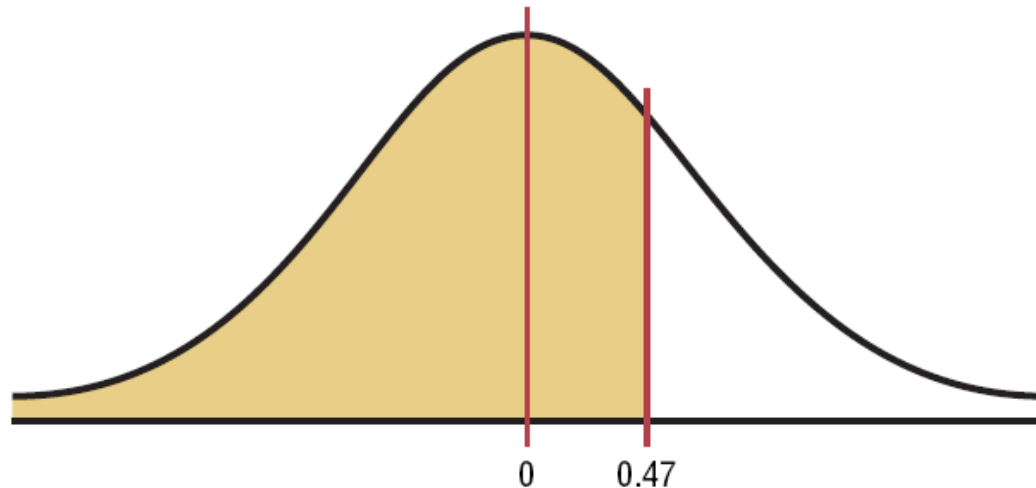
Step 3: Find the area to the left of $z = 0.47$.

Table E gives us an area of .6808.

68% of women spend less than $160.

Chapter 6
Normal Distributions

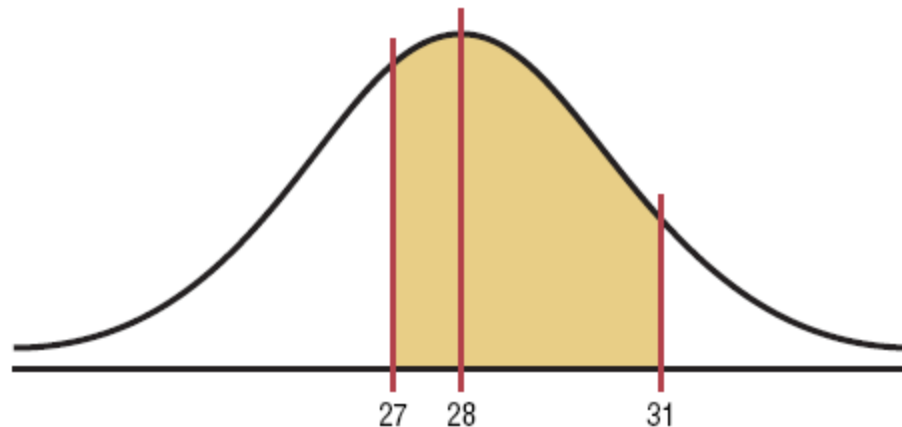# Section 6-2

Example 6-7a

Page #317

Example 6-7a: Newspaper Recycling

Each month, an American household generates an average of 28 pounds of newspaper for garbage or recycling. Assume the standard deviation is 2 pounds. If a household is selected at random, find the probability of its generating between 27 and 31 pounds per month. Assume the variable is approximately normally distributed.

Step 1: Draw the normal distribution curve.

27 28   31

Example 6-7a: Newspaper Recycling

Step 2: Find *z* values corresponding to 27 and 31.

$$z = \frac{27-28}{2} = -0.5 \qquad z = \frac{31-28}{2} = 1.5$$
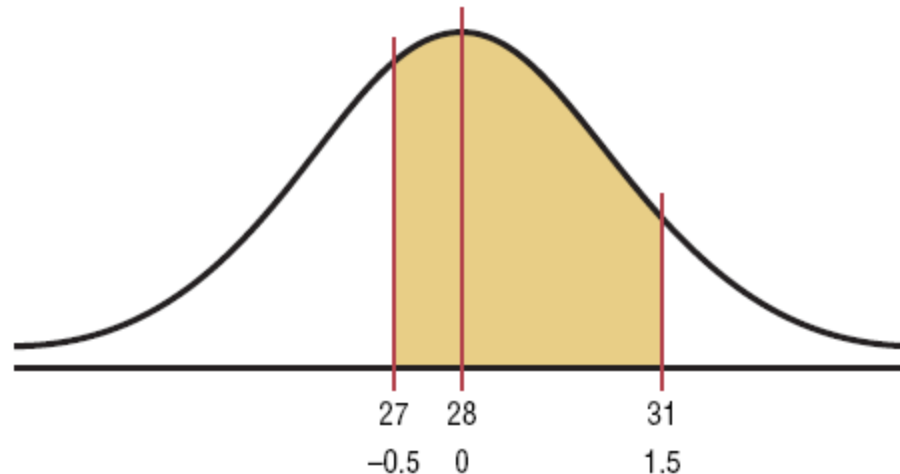
Step 3: Find the area between $z = $ -0.5 and $z = $ 1.5.

Table E gives us an area of .9332 - .3085 = .6247. The probability is 62%.

Chapter 6
Normal Distributions

Section 6-2

Example 6-8

Page #318

Example 6-8: Emergency Response

The American Automobile Association reports that the average time it takes to respond to an emergency call is 25 minutes. Assume the variable is approximately normally distributed and the standard deviation is 4.5 minutes. If 80 calls are randomly selected, approximately how many will be responded to in less than 15 minutes?

Step 1: Draw the normal distribution curve.

15   25

Example 6-8: Emergency Response

Step 2: Find the *z* value for 15.

$$z = \frac{15 - 25}{4.5} = -2.22$$

Step 3: Find the area to the left of $z = -2.22$. It is 0.0132.

Step 4: To find how many calls will be made in less than 15 minutes, multiply the sample size 80 by 0.0132 to get 1.056. Hence, approximately 1 call will be responded to in under 15 minutes.

Chapter 6
Normal Distributions

Formula for Finding X

When one must find the value of X, the following formula can be used:

$$X = z \cdot \sigma + \mu$$

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

Section 6-2
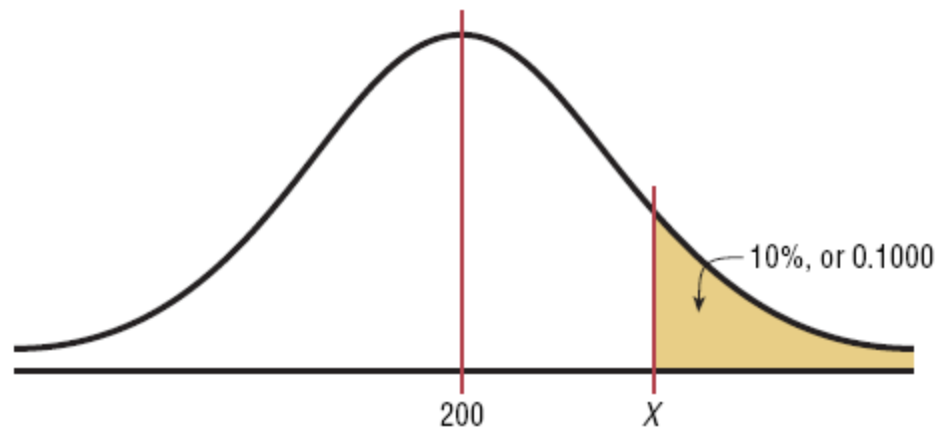
Example 6-9
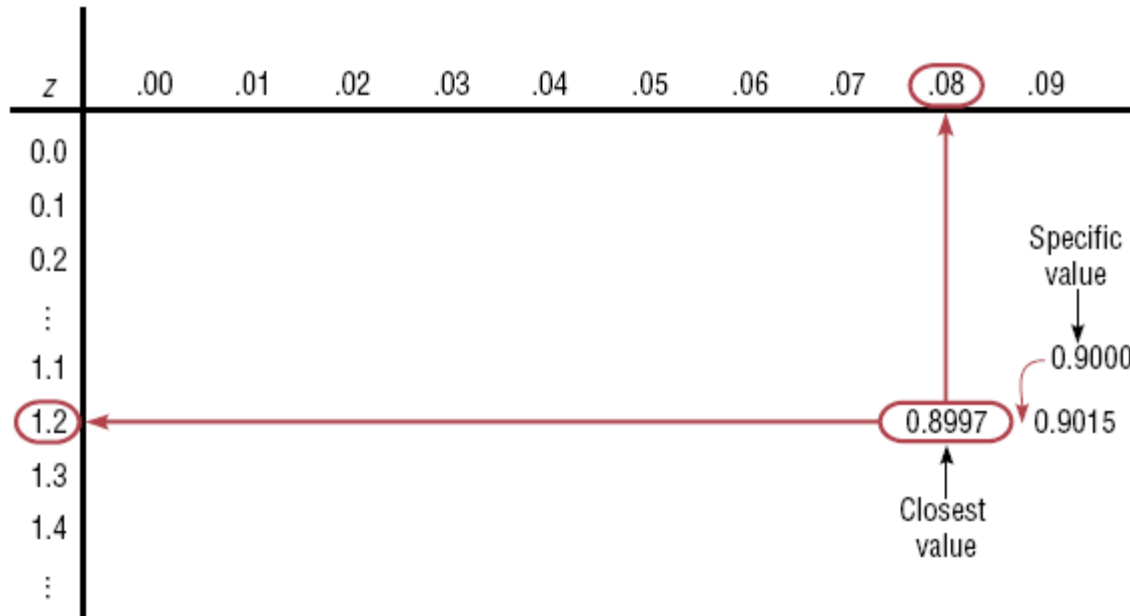
Page #319

Example 6-9: Police Academy

To qualify for a police academy, candidates must score in the top 10% on a general abilities test. The test has a mean of 200 and a standard deviation of 20. Find the lowest possible score to qualify. Assume the test scores are normally distributed.

Step 1: Draw the normal distribution curve.

10%, or 0.1000

200   X

Example 6-8: Newspaper Recycling

Step 2: Subtract 1 - 0.1000 to find area to the left, 0.9000. Look for the closest value to that in Table E.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0										
0.1										
0.2										
⋮										
1.1										
1.2									0.8997	0.9015
1.3										
1.4										
⋮										

Specific value
0.9000

Closest value

Step 3: Find X.

X = Z * σ + μ = 1.28 (20) + 200 = 225.6

The cutoff, the lowest possible score to qualify, is 226.

Chapter 6
# Normal Distributions

Section 6-2

Example 6-10

Page #321

# Example 6-10: Systolic Blood Pressure

For a medical study, a researcher wishes to select people in the middle 60% of the population based on blood pressure. If the mean systolic blood pressure is 120 and the standard deviation is 8, find the upper and lower readings that would qualify people to participate in the study.

Step 1: Draw the normal distribution curve.
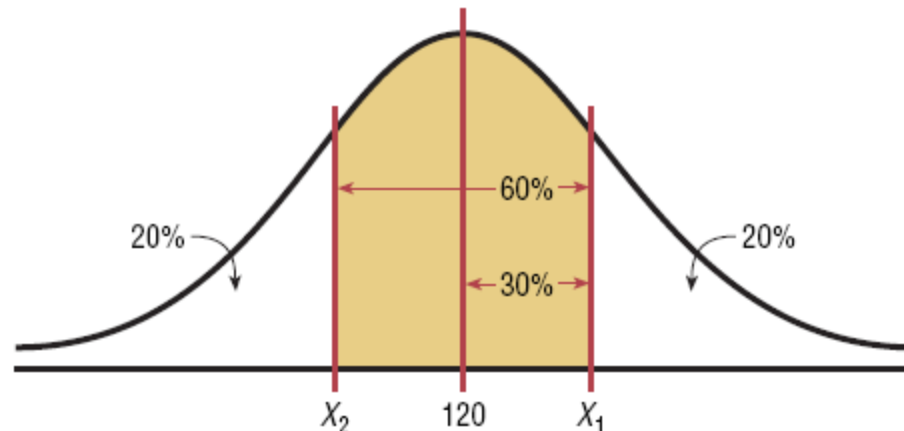
Example 6-10: Systolic Blood Pressure

$$X = z \cdot \sigma + \mu$$
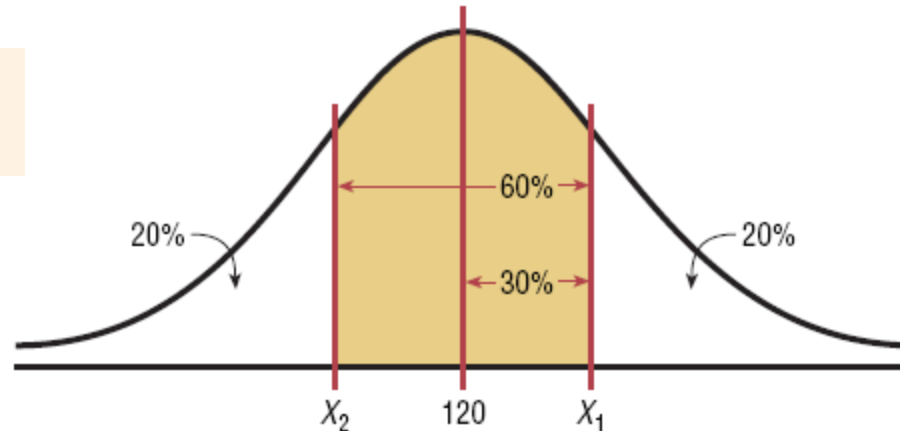
Area to the left of the positive z: $0.5000 + 0.3000 = 0.8000$.

Using Table E, $z \approx 0.84$. $X = 0.84(8) + 120 = 126.72$

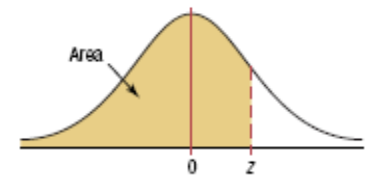Area to the left of the negative z: $0.5000 - 0.3000 = 0.2000$.

Using Table E, $z \approx -0.84$. $X = -0.84(8) + 120 = 113.28$

The middle 60% of readings are between 113 and 127.

Cumulative Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952

Bluman, Chapter 6, 03/2010

Table E — The Standard Normal Distribution

Cumulative Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

For z values less than −3.49, use 0.0001.

Area

z   0

Applications with Normal Distributions
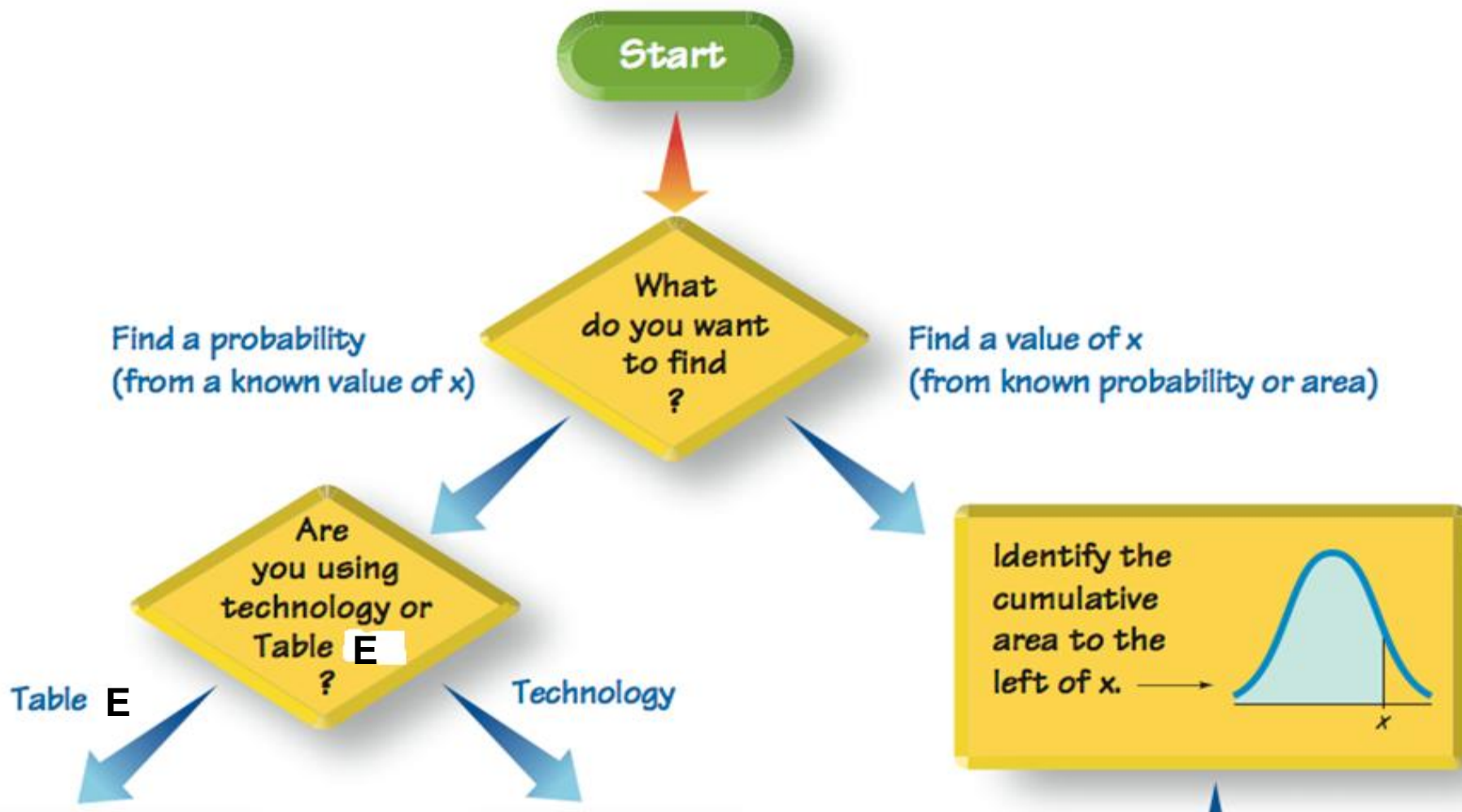
Find a probability (from a known value of x)

Table **E**

?

Technology
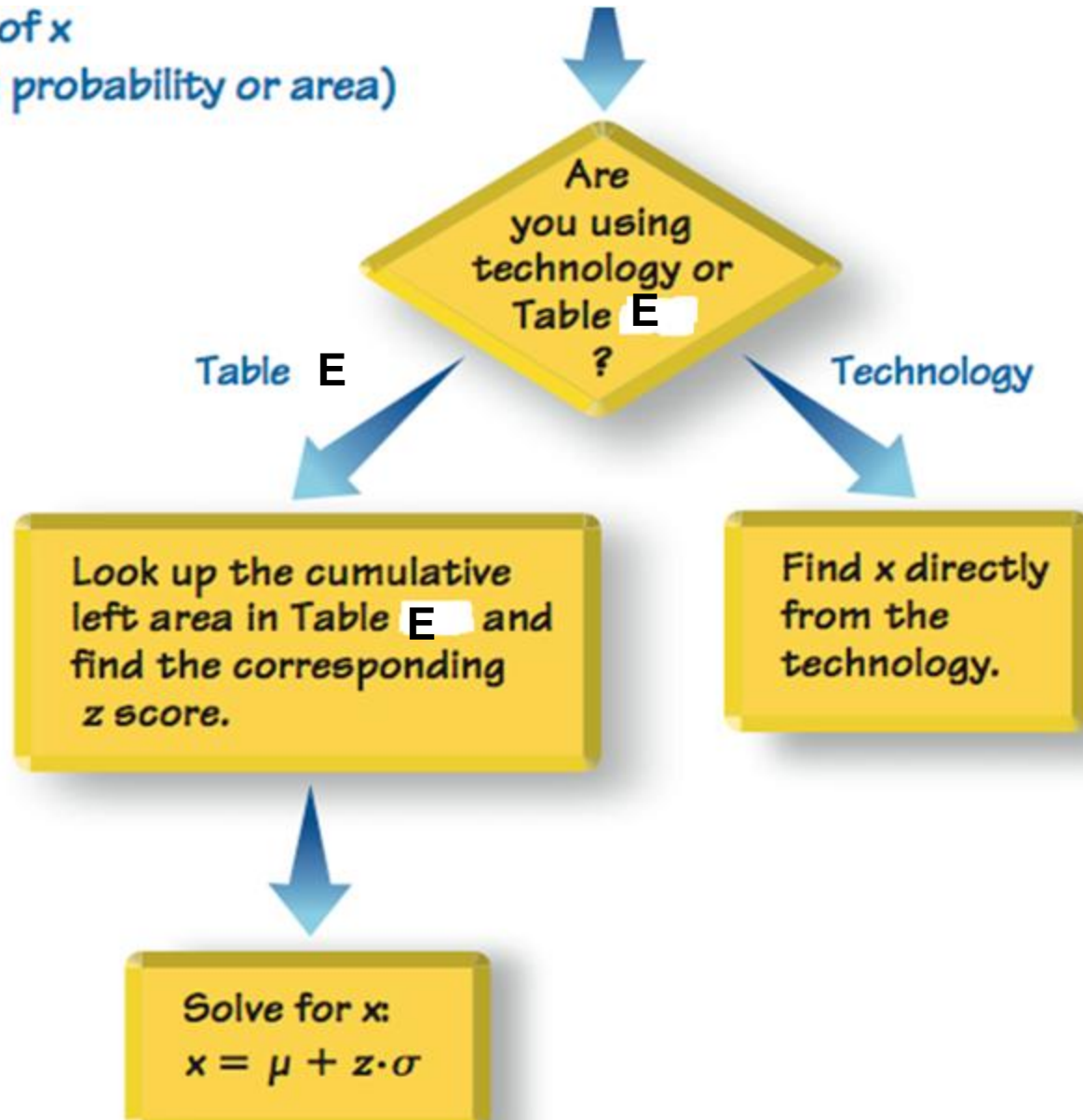
Convert to the standard normal distribution by finding z:

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

Find the probability by using the technology.

Look up z in Table **E** and find the cumulative area to the left of z.

Find a value of x
(from known probability or area)

Are you using technology or Table **E**?

Table **E**

Technology

Look up the cumulative left area in Table **E** and find the corresponding z score.

Find x directly from the technology.

Solve for x:
$$x = \mu + z \cdot \sigma$$

Normal Distributions

- A normally shaped or bell-shaped distribution is only one of many shapes that a distribution can assume; however, it is very important since many statistical methods require that the distribution of values (shown in subsequent chapters) be normally or approximately normally shaped.

- There are a number of ways statisticians check for normality. We will focus on three of them.

Checking for Normality

- Histogram

- Pearson's Index PI of Skewness

- Outliers

- Other Tests
  - Normal Quantile Plot
  - Chi-Square Goodness-of-Fit Test
  - Kolmogorov-Smikirov Test
  - Lilliefors Test

Chapter 6
Normal Distributions

Section 6-2

Example 6-11

Page #322

Example 6-11: Technology Inventories

A survey of 18 high-technology firms showed the number of days' inventory they had on hand. Determine if the data are approximately normally distributed.

5 29 34 44 45 63 68 74 74

81 88 91 97 98 113 118 151 158

Method 1: Construct a Histogram.

Class	Frequency
5–29	2
30–54	3
55–79	4
80–104	5
105–129	2
130–154	1
155–179	1
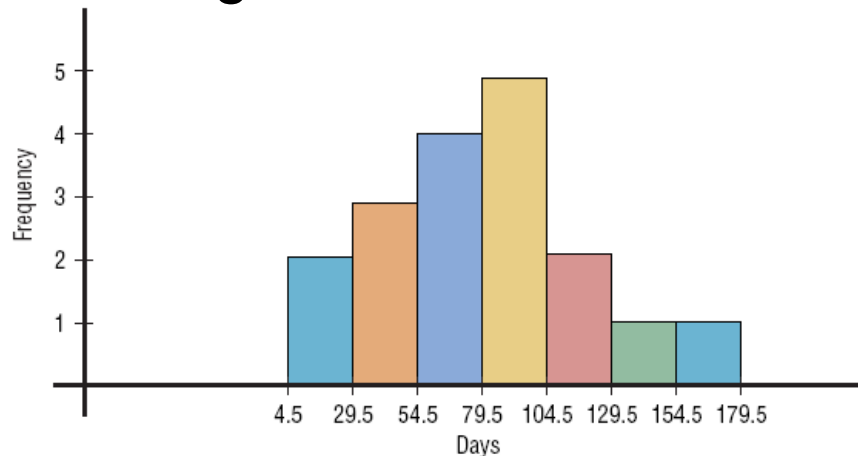
The histogram is approximately bell-shaped.

Example 6-11: Technology Inventories

Method 2: Check for Skewness.

$$\bar{X} = 79.5, MD = 77.5, s = 40.5$$

$$\text{PI} = \frac{3(\bar{X} - MD)}{s} = \frac{3(79.5 - 77.5)}{40.5} = 0.148$$

The PI is not greater than +1 or less than -1, so it can be concluded that the distribution is not significantly skewed.

Method 3: Check for Outliers.

	lowest	Q2	MD	Q3	highest
Five-Number Summary:	5 -	45 -	77.5 -	98 -	158

IQR = Q3 − Q1 = 98 − 45 = 53

Q1 − 1.5(IQR) = 45 − 1.5(53) = -34.5

Q3 + 1.5(IQR) = 98 + 1.5(53) = 177.5

No data below -34.5 or above 177.5, so no outliers.

Example 6-11: Technology Inventories

A survey of 18 high-technology firms showed the number of days' inventory they had on hand. Determine if the data are approximately normally distributed.

5 29 34 44 45 63 68 74 74

81 88 91 97 98 113 118 151 158

Conclusion:

- The histogram is approximately bell-shaped.
- The data are not significantly skewed.
- There are no outliers.

Thus, it can be concluded that the distribution is approximately normally distributed.

6.3 The Central Limit Theorem

In addition to knowing how individual data values vary about the mean for a population, statisticians are interested in knowing how the means of samples of the same size taken from the same population vary about the population mean.

Distribution of Sample Means

- A ***sampling distribution of sample means*** is a distribution obtained by using the means computed from random samples of a specific size taken from a population.

- ***Sampling error*** is the difference between the sample measure and the corresponding population measure due to the fact that the sample is not a perfect representation of the population.

Properties of the Distribution of Sample Means

- The mean of the sample means will be the same as the population mean.

- The standard deviation of the sample means will be smaller than the standard deviation of the population, and will be equal to the population standard deviation divided by the square root of the sample size.

The Central Limit Theorem

- As the sample size *n* increases, the shape of the distribution of the sample means taken with replacement from a population with mean μ and standard deviation σ will approach a normal distribution.

- The mean of the sample means equals the population mean. $\mu_{\bar{X}} = \mu$.

- The standard deviation of the sample means is called the **standard error of the mean** (S.E.)

$$\sigma_{\bar{X}} = \sigma / \sqrt{n}.$$

The Central Limit Theorem

- The central limit theorem can be used to answer questions about sample means in the same manner that the normal distribution can be used to answer questions about individual values.

- A new formula must be used for the z values:

$$z = \frac{\overline{X} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

Students sometimes have difficulty deciding whether to use

$$z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \qquad \text{or} \qquad z = \frac{X - \mu}{\sigma}$$

The formula

$$z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

should be used to gain information about a sample mean, as shown in this section. The formula

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

is used to gain information about an individual data value obtained from the population. Notice that the first formula contains $\overline{X}$, the symbol for the sample mean, while the second formula contains X, the symbol for an individual data value.

Chapter 6
Normal Distributions

Section 6-3

Example 6-13

Page #334

Example 6-13: Hours of Television

A. C. Neilsen reported that children between the ages of 2 and 5 watch an <u>average</u> of 25 hours of television per week. Assume the variable is normally distributed and the standard deviation is 3 hours. If <u>20</u> children between the ages of 2 and 5 are randomly selected, find the probability that the <u>mean</u> of the number of hours they watch television will be greater than 26.3 hours.
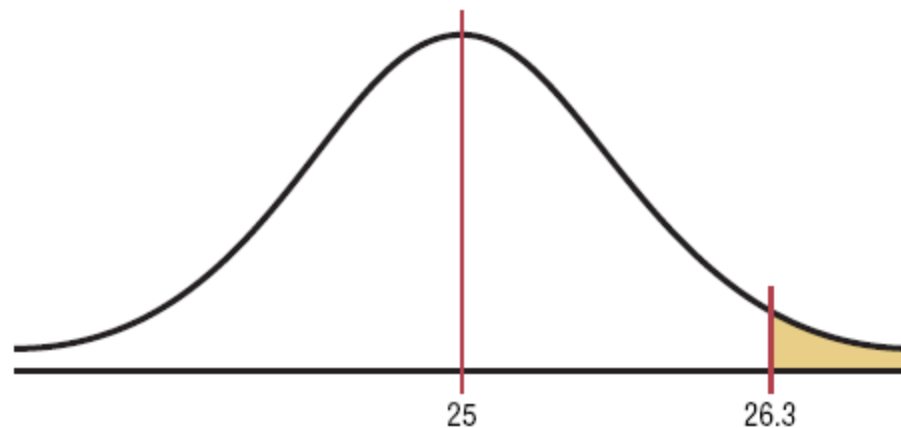
Example 6-13: Hours of Television
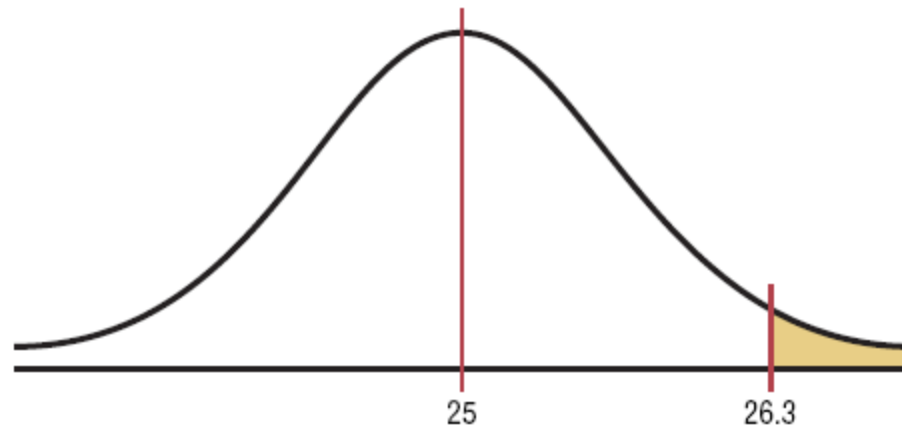
Since we are calculating probability for a sample mean, we need the Central Limit Theorem formula

$$z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{26.3 - 25}{3 / \sqrt{20}} = 1.94$$

The area is $1.0000 - 0.9738 = 0.0262$. The probability of obtaining a sample mean larger than 26.3 hours is 2.62%.

Chapter 6
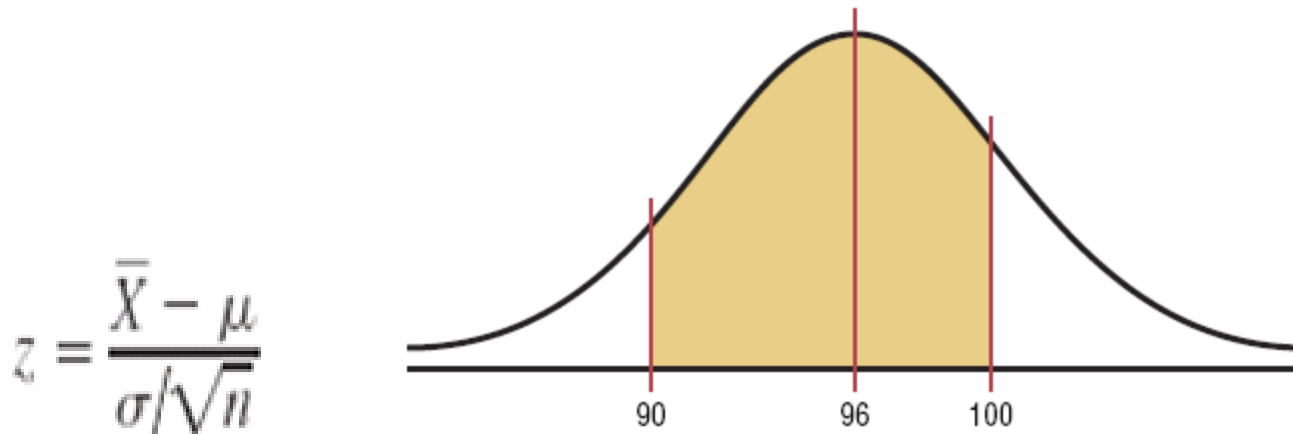# Normal Distributions

Section 6-3

Example 6-14

Page #335

Example 6-14: Vehicle Age

The <u>average</u> age of a vehicle registered in the United States is 8 years, or <u>96</u> months. Assume the standard deviation is 16 months. If a random sample of <u>36</u> vehicles is selected, find the probability that the <u>mean</u> of their age is between 90 and 100 months.

Since the sample is 30 or larger, the normality assumption is not necessary.
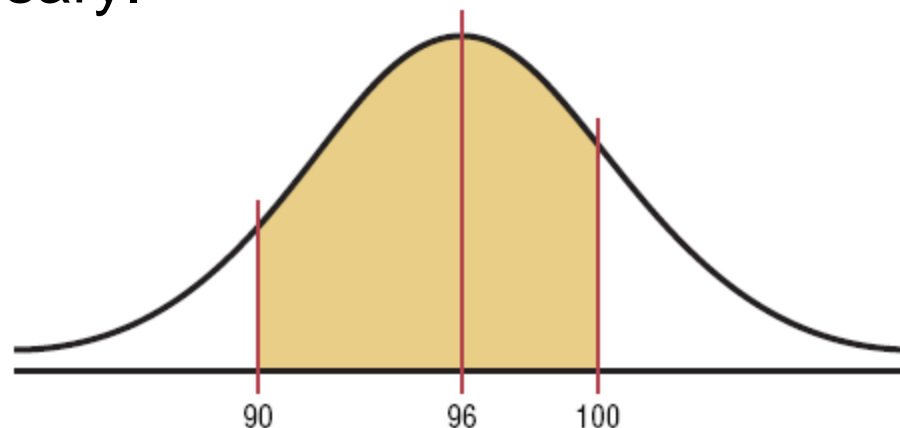
Example 6-14: Vehicle Age

$$z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

$$z = \frac{90 - 96}{16 / \sqrt{36}} = -2.25 \qquad z = \frac{100 - 96}{16 / \sqrt{36}} = 1.50$$

Table E gives us areas 0.9332 and 0.0122, respectively.

The desired area is 0.9332 - 0.0122 = 0.9210.

The probability of obtaining a sample mean between 90 and 100 months is 92.1%.
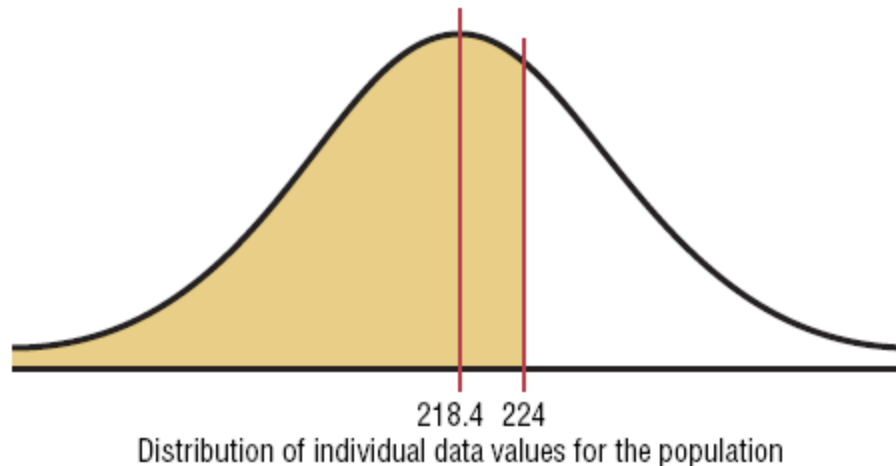
Chapter 6
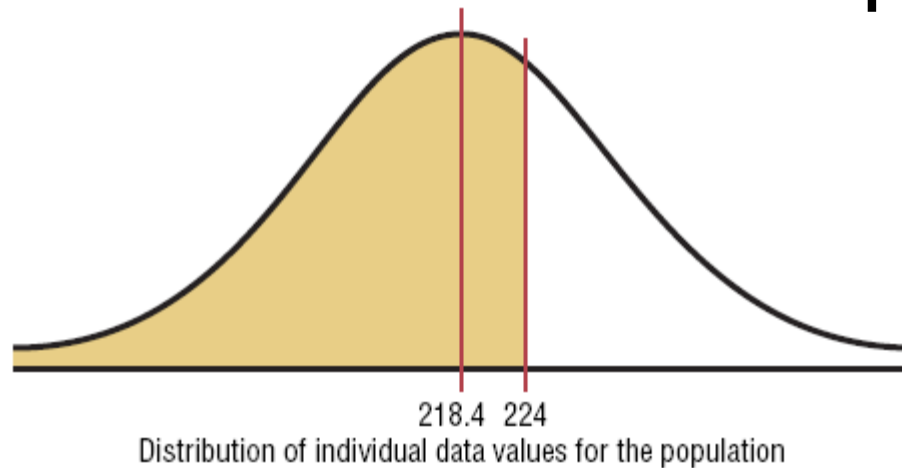# Normal Distributions

Section 6-3

Example 6-15

Page #336

Example 6-15: Meat Consumption

The <u>average</u> number of pounds of meat that a person consumes per year is 218.4 pounds. Assume that the standard deviation is 25 pounds and the distribution is approximately normal.

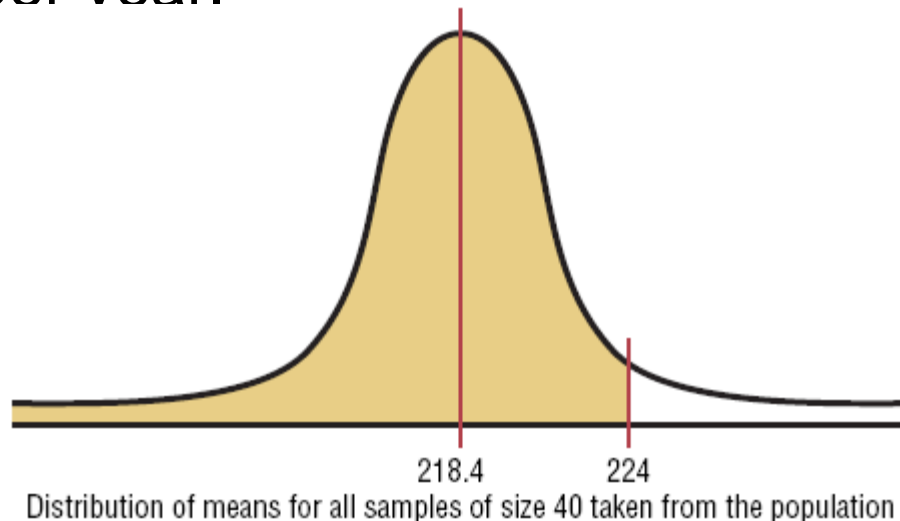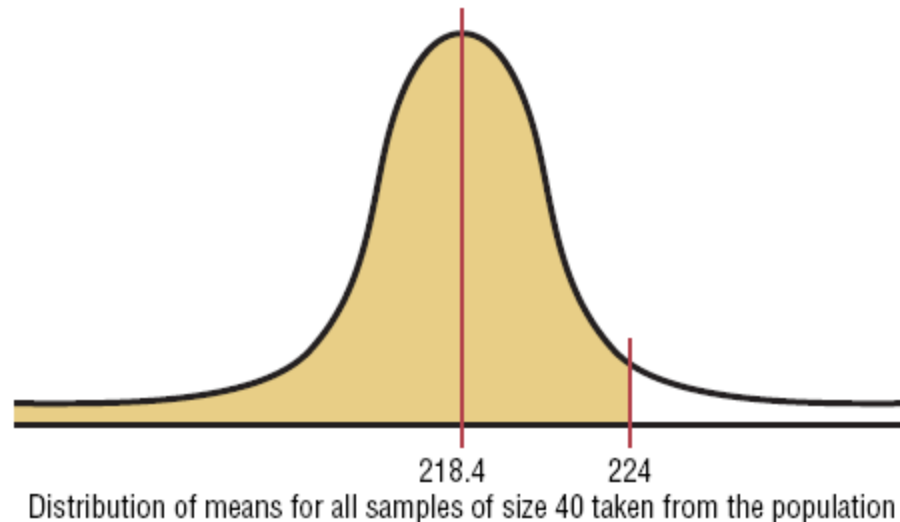   a. Find the probability that a person selected at random consumes less than 224 pounds per year.

218.4 224

Distribution of individual data values for the population

Example 6-15: Meat Consumption

218.4 224
Distribution of individual data values for the population

$$z = \frac{X - \mu}{\sigma} = \frac{224 - 218.4}{25} = 0.22$$

The area to the left of $z = 0.22$ is 0.5871. Hence, the probability of selecting an individual who consumes less than 224 pounds of meat per year is 0.5871, or 58.71%.

Example 6-15: Meat Consumption

The <u>average</u> number of pounds of meat that a person consumes per year is 218.4 pounds. Assume that the standard deviation is 25 pounds and the distribution is approximately normal.

b. If a sample of <u>40</u> individuals is selected, find the probability that <u>mean</u> of the sample will be less than 224 pounds per vear.

218.4   224
Distribution of means for all samples of size 40 taken from the population

Example 6-15: Meat Consumption

218.4   224

Distribution of means for all samples of size 40 taken from the population

$$z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} = \frac{224 - 218.4}{25 / \sqrt{40}} = 1.42$$

The area to the left of $z = 1.42$ is 0.9222. Hence, the probability that the mean of a sample of 40 individuals is less than 224 pounds per year is 0.9222, or 92.22%.

Finite Population Correction Factor

- The formula for standard error of the mean is accurate when the samples are drawn with replacement or are drawn without replacement from a very large or infinite population.

- A **correction factor** is necessary for computing the standard error of the mean for samples drawn without replacement from a finite population.

Finite Population Correction Factor

- The correction factor is computed using the following formula:

$$\sqrt{\frac{N-n}{N-1}}$$

  where *N* is the population size and *n* is the sample size.

- The standard error of the mean must be multiplied by the correction factor to adjust it for large samples taken from a small population.

Finite Population Correction Factor

$$\frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

- The standard error for the mean must be adjusted when it is included in the formula for calculating the *z* values.

$$\frac{\overline{X} - \mu}{\dfrac{\sigma}{\sqrt{n}} \cdot \sqrt{\dfrac{N-n}{N-1}}}$$

6.4 The Normal Approximation to the Binomial Distribution

A normal distribution is often used to solve problems that involve the binomial distribution since when *n* is large (say, 100), the calculations are too difficult to do by hand using the binomial distribution.
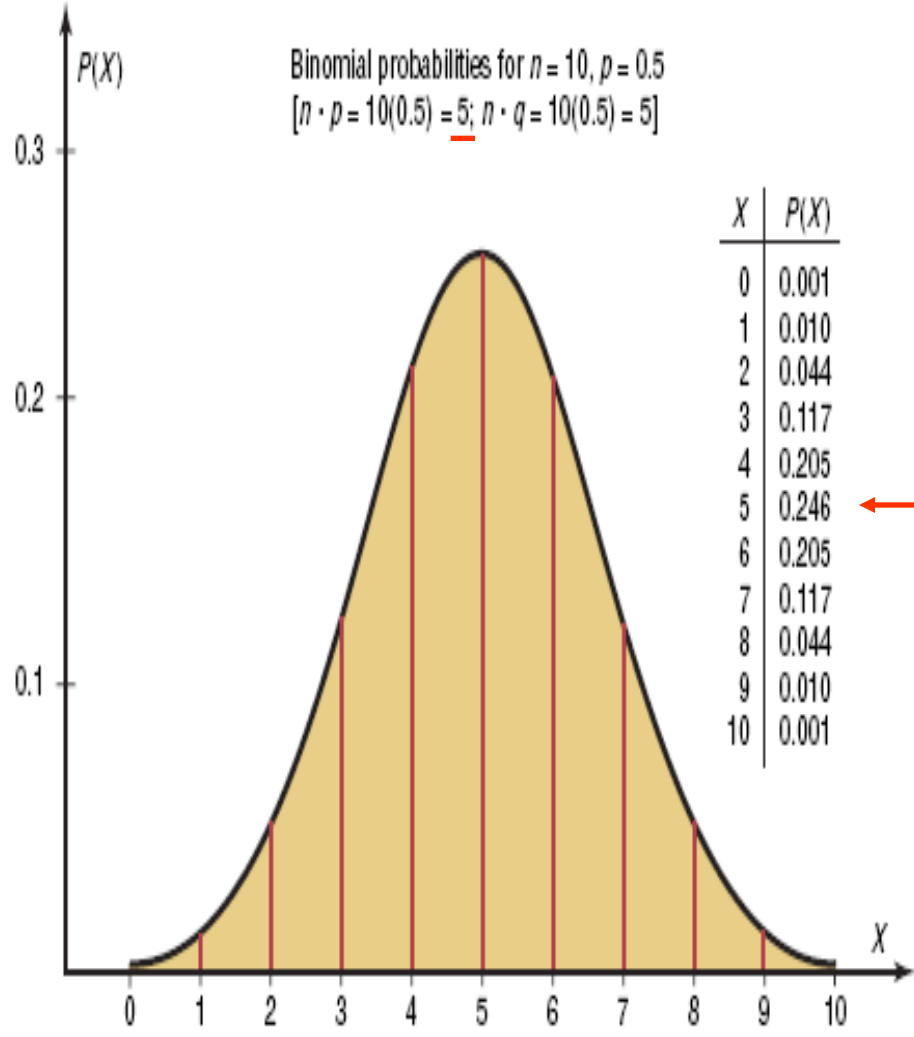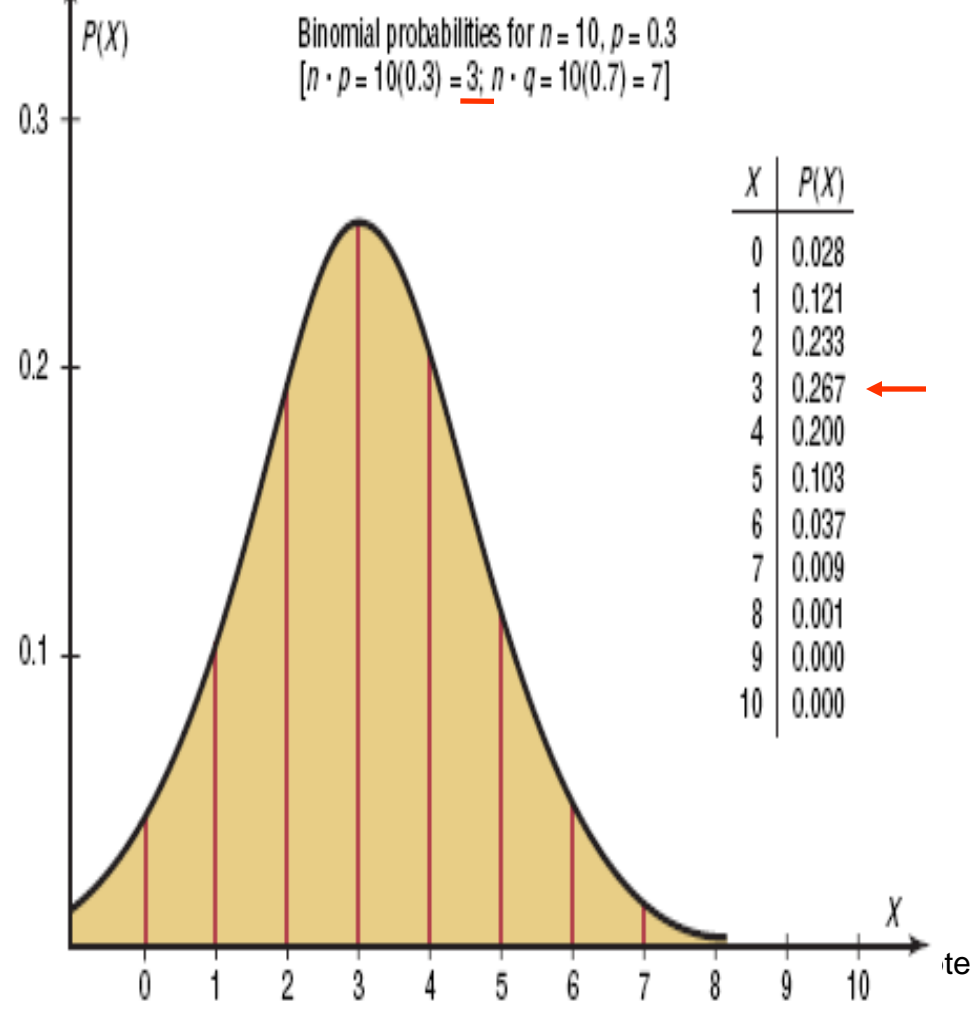
The Normal Approximation to the Binomial Distribution

- The normal approximation to the binomial is appropriate when $np \geq 5$ and $nq \geq 5$.

- In addition, a **correction for continuity** may be used in the normal approximation to the binomial.

- The continuity correction means that for any specific value of X, say 8, the boundaries of X in the binomial distribution (in this case, 7.5 to 8.5) must be used.

http://www.ruf.rice.edu/~lane/stat_sim/binom_demo.html
http://www.stat.wvu.edu/SRS/Modules/NormalApprox/normalapprox.html

Normal distribution should not
be used as an approximation below (n*p = μ)

Binomial probabilities for $n = 10$, $p = 0.3$
$[n \cdot p = 10(0.3) = 3; n \cdot q = 10(0.7) = 7]$

X	P(X)
0	0.028
1	0.121
2	0.233
3	0.267
4	0.200
5	0.103
6	0.037
7	0.009
8	0.001
9	0.000
10	0.000

Binomial probabilities for $n = 10$, $p = 0.5$
$[n \cdot p = 10(0.5) = 5; n \cdot q = 10(0.5) = 5]$

X	P(X)
0	0.001
1	0.010
2	0.044
3	0.117
4	0.205
5	0.246
6	0.205
7	0.117
8	0.044
9	0.010
10	0.001

The Normal Approximation to the Binomial Distribution

Binomial	**Normal**	
When finding:	Use:	Notes:
$P(X = a)$	$P(a - 0.5 < X < a + 0.5)$	L& H boundary of a
$P(X \geq a)$	$P(X > a - 0.5)$	L boundary of a
$P(X > a)$	$P(X > a + 0.5)$	L boundary of a+1
$P(X \leq a)$	$P(X < a + 0.5)$	U boundary of a
$P(X < a)$	$P(X < a - 0.5)$	U boundary of a-1

For all cases, $\mu = np$, $\sigma = \sqrt{npq}$, $np \geq 5$, $nq \geq 5$

Binomial   **Normal**

X = <u>at least</u> 8

(X >= 8 includes 8 and above)

X = <u>more than</u> 8

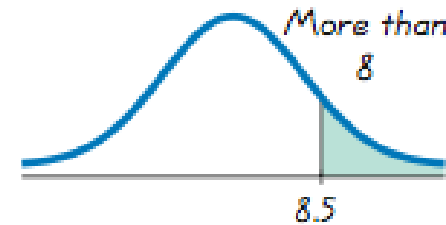(X >8 doesn't include 8.
9, 10, …)

X = <u>at most</u> 8

(X <=8 includes 8 and below)

X = <u>fewer than</u> 8

(X <8 doesn't include 8.
7, 6, …)

X = <u>exactly</u> 8
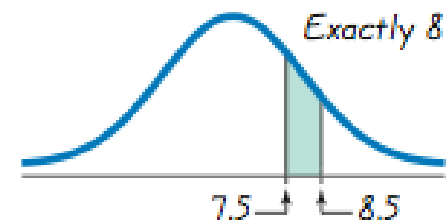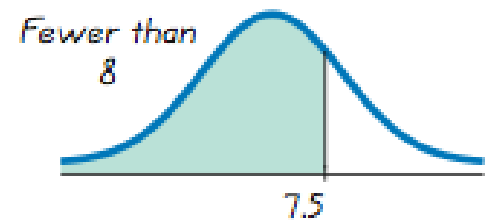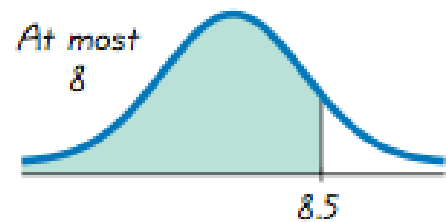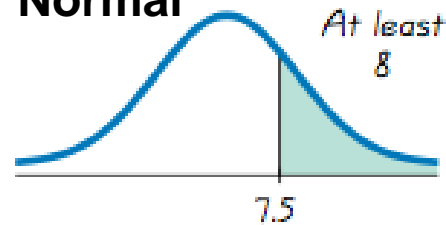
The Normal Approximation to the Binomial Distribution

Procedure Table

Step 1: Check to see whether the normal approximation can be used.

Step 2: Find the mean μ and the standard deviation σ.

Step 3: Write the problem in probability notation, using X.

Step 4: Rewrite the problem by using the continuity correction factor, and show the corresponding area under the normal distribution.

Step 5: Find the corresponding z values.

Step 6: Find the solution.

Chapter 6
Normal Distributions

Section 6-4

Example 6-16

Page #343

Example 6-16: Reading While Driving

A magazine reported that 6% of American drivers read the newspaper while driving. If 300 drivers are selected at random, find the probability that exactly 25 say they read the newspaper while driving.

Here, $p = 0.06$, $q = 0.94$, and $n = 300$.

Step 1: Check to see whether a normal approximation can be used.

$np = (300)(0.06) = 18$ and $nq = (300)(0.94) = 282$

Since $np \geq 5$ and $nq \geq 5$, we can use the normal distribution.

Step 2: Find the mean and standard deviation.

$\mu = np = (300)(0.06) = 18$

$\sigma = \sqrt{npq} = \sqrt{300(0.06)(0.94)} = 4.11$

Example 6-16: Reading While Driving

Step 3: Write in probability notation. $\mathrm{P}(X=25)$

Step 4: Rewrite using the continuity correction factor.

$$\mathrm{P}(24.5 < X < 25.5)$$

Step 5: Find the corresponding *z* values.

$$z = \frac{24.5 - 18}{4.11} = 1.58, \quad z = \frac{25.5 - 18}{4.11} = 1.82$$

Step 6: Find the solution

The area between the two *z* values is
0.9656 - 0.9429 = 0.0227, or 2.27%.

Hence, the probability that exactly 25 people read the newspaper while driving is 2.27%.

Chapter 6
Normal Distributions

Section 6-4

Example 6-17

Page #343

Example 6-17: Widowed Bowlers

Of the members of a bowling league, 10% are widowed. If 200 bowling league members are selected at random, find the probability that 10 or more will be widowed.

Here, $p = 0.10$, $q = 0.90$, and $n = 200$.

Step 1: Check to see whether a normal approximation can be used.

$np = (200)(0.10) = 20$ and $nq = (200)(0.90) = 180$

Since $np \geq 5$ and $nq \geq 5$, we can use the normal distribution.

Step 2: Find the mean and standard deviation.

$\mu = np = (200)(0.10) = 20$

$\sigma = \sqrt{npq} = \sqrt{200(0.10)(0.90)} = 4.24$

Example 6-17: Widowed Bowlers

Step 3: Write in probability notation. $P(X \geq 10)$

Step 4: Rewrite using the continuity correction factor.

$P(X > 9.5)$

Step 5: Find the corresponding z values.

$$z = \frac{9.5 - 20}{4.24} = -2.48$$

Step 6: Find the solution

The area to the right of the z value is
1.0000 - 0.0066 = 0.9934, or 99.34%.

The probability of 10 or more widowed people in a random sample of 200 bowling league members is 99.34%.