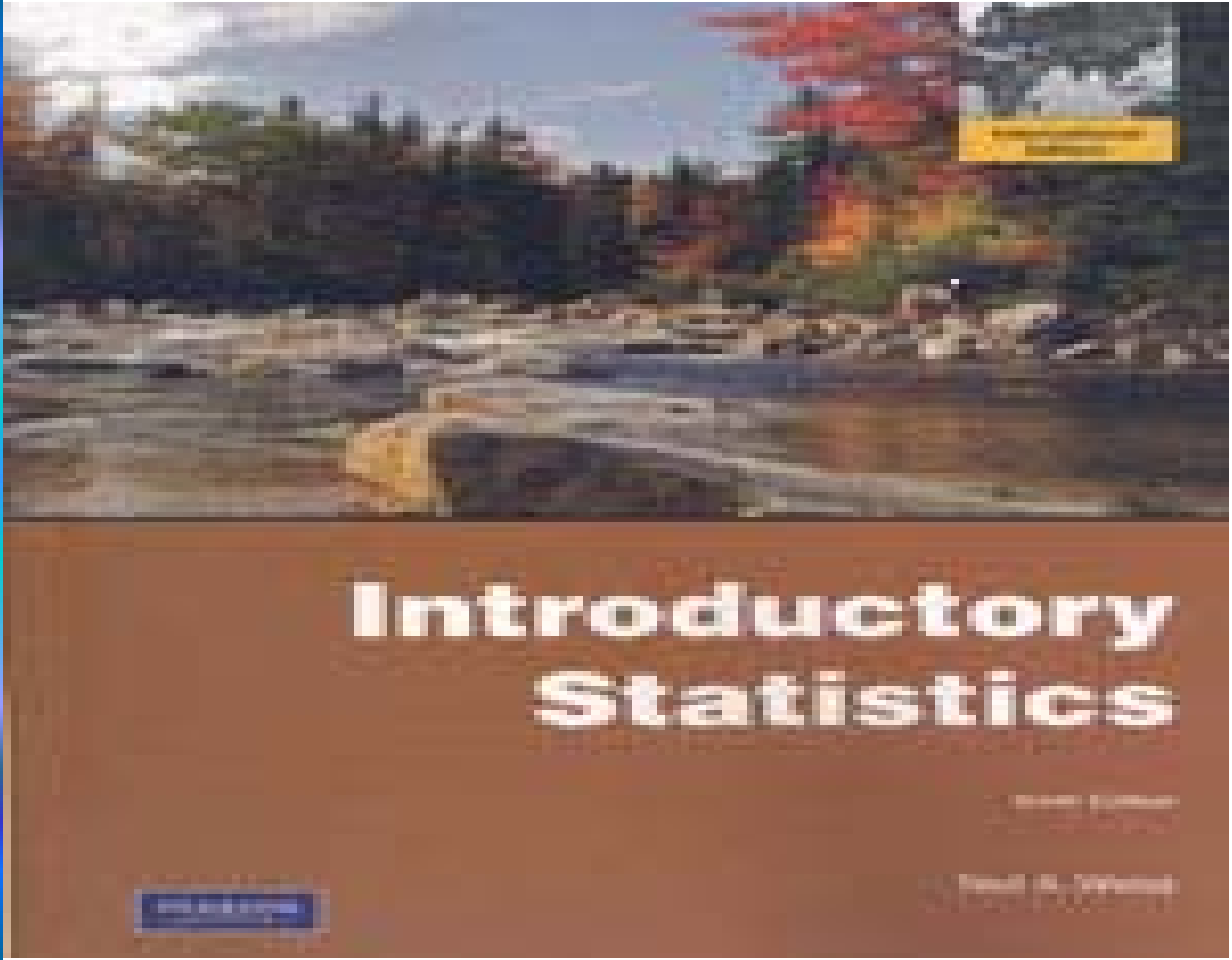

Introductory
Statistics

Chapter 8

Confidence Intervals for One Population Mean

Next Goals

- Statistical confidence
- Confidence intervals
- Behavior of confidence intervals
- Choosing the sample size

Motivations

- Distinguish chance variations from permanent features of a phenomenon.

- Give SAT test to a SRS of 500 California seniors, mean $\bar{X} = 461$,

  What does it say about the mean SAT score of 250,000 HS seniors in CA?

- Is 12 /20 vs. 8/20 improvements in treatment and control group a strong enough evidence in favor of a drug?

Section 8.1

Estimating a Population Mean

Definition 8.1

Point Estimate

A **point estimate** of a parameter is the value of a statistic used to estimate the parameter.

Interval Estimates

- An Interval Estimate states the range within which a population parameter probably lies.

  - The interval within which a population parameter is expected to occur is called a confidence interval.

  - The two confidence intervals that are used extensively are the 95% and the 99%.

Standard Error of the Sample Means

The standard error of the sample mean is the standard deviation of the sampling distribution of the sample means.

In the case where we know about the standard deviation of the population, σ, then

- It is computed by $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$
- $\sigma_{\bar{x}}$ is the symbol for the standard error of the sample mean.
- n is the size of the sample.

Assumptions Made

- Methods of formal inference rely on the assumption that the data come from properly **randomized** experiment (e.g. SRS).

- Statistics gives methods that give correct results a prescribed (high) percentage of times (if repeated many times).

- Most prominent: **confidence intervals** and **tests of significance.**

Example

We observe 15 plots of corn with yields (in bushels):

138, 139.1, 113, 132.5, 140.7, 109.7, 118.9, 134.8, 109.6, 127.3,115.6, 130.4, 130.2, 111.7, 105.5

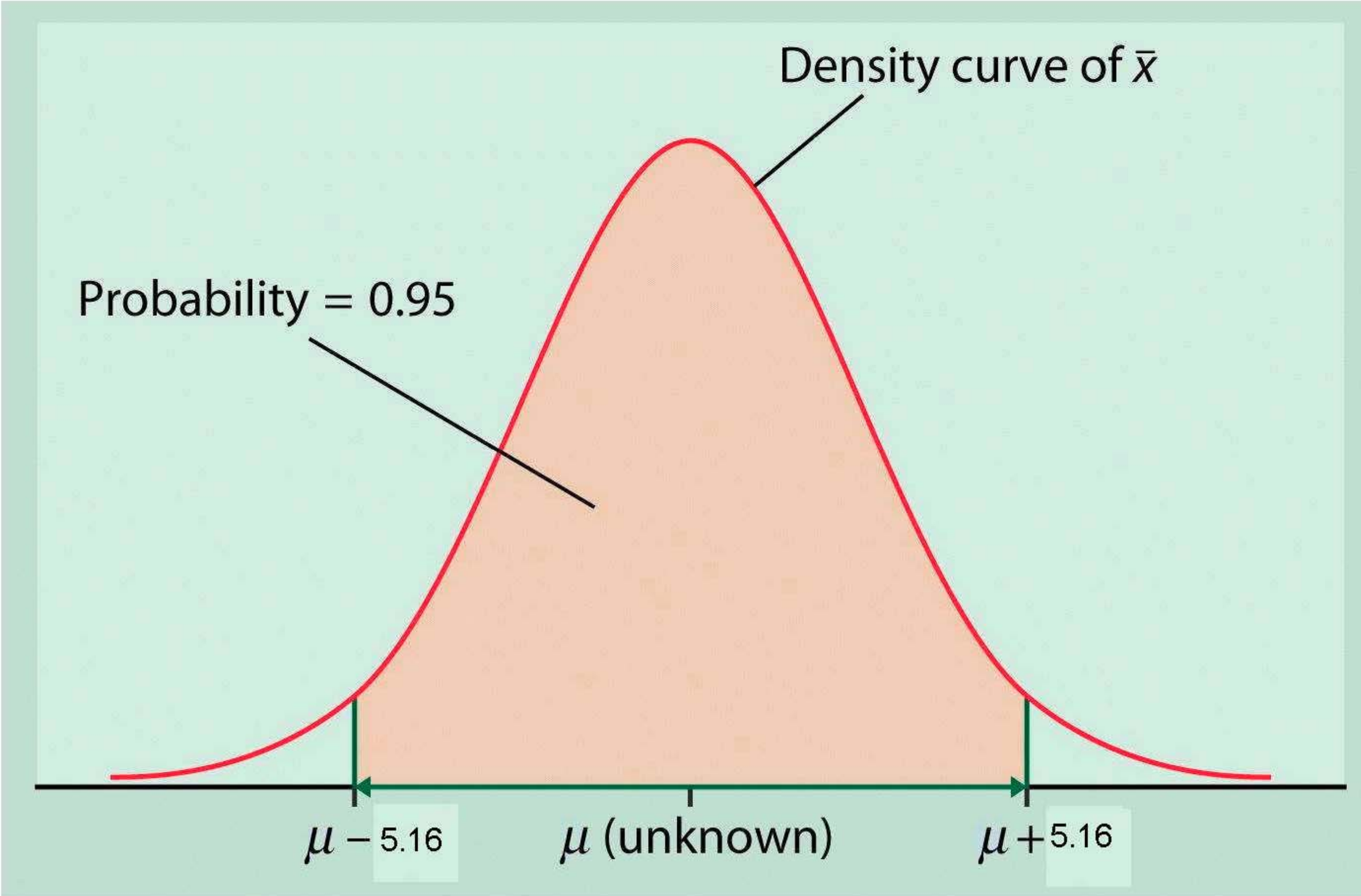
sample mean= 123.8

What can be said about (population) mean yield of this variety of corn?

Example: Statistical confidence --

- Assume that yield is N(μ, σ) with unknown μ and σ=10 (just assume σ is known).
- Then $\bar{x} \sim N(\mu, \sigma/\sqrt{15}) = N(\mu, 2.58)$
- 68-95-99.7 rule : 95% of times sample mean is within 2×2.58=5.16 from μ
- Thus 95% of times

$$\mu - 5.16 < \bar{x} < \mu + 5.16$$


Density curve of $\bar{x}$
Probability = 0.95
$\mu - 5.16$
μ (unknown)
$\mu + 5.16$

Put differently: 95% of times

$$\bar{x} - 5.16 < \mu < \bar{x} + 5.16$$

- The *random* interval *covers* the unknown (but nonrandom) population parameter μ 95% of times.
- Here is our confidence: 95%. It is in method rather than this particular interval.

Example cont.: (mean yield of corn estimated)

μ is between $\bar{x} \pm 5.16 = 123.8 \pm 5.16 = (118.64,\ 128.96)$

- This particular confidence interval may contain μ or not…
- However, such systematic **method** gives intervals covering the population mean μ in 95% of cases.

Confidence intervals:

- Typically: **estimate ± margin of error**
- Always an interval of the form (a, b) with endpoints calculated from data
- **Confidence level**, C, gives probability that such interval(s) will cover the true value of the parameter

Definition 8.2

Confidence-Interval Estimate

Confidence interval (CI): An interval of numbers obtained from a point estimate of a parameter.

Confidence level: The confidence we have that the parameter lies in the confidence interval (i.e., that the confidence interval contains the parameter).

Confidence-interval estimate: The confidence level and confidence interval.

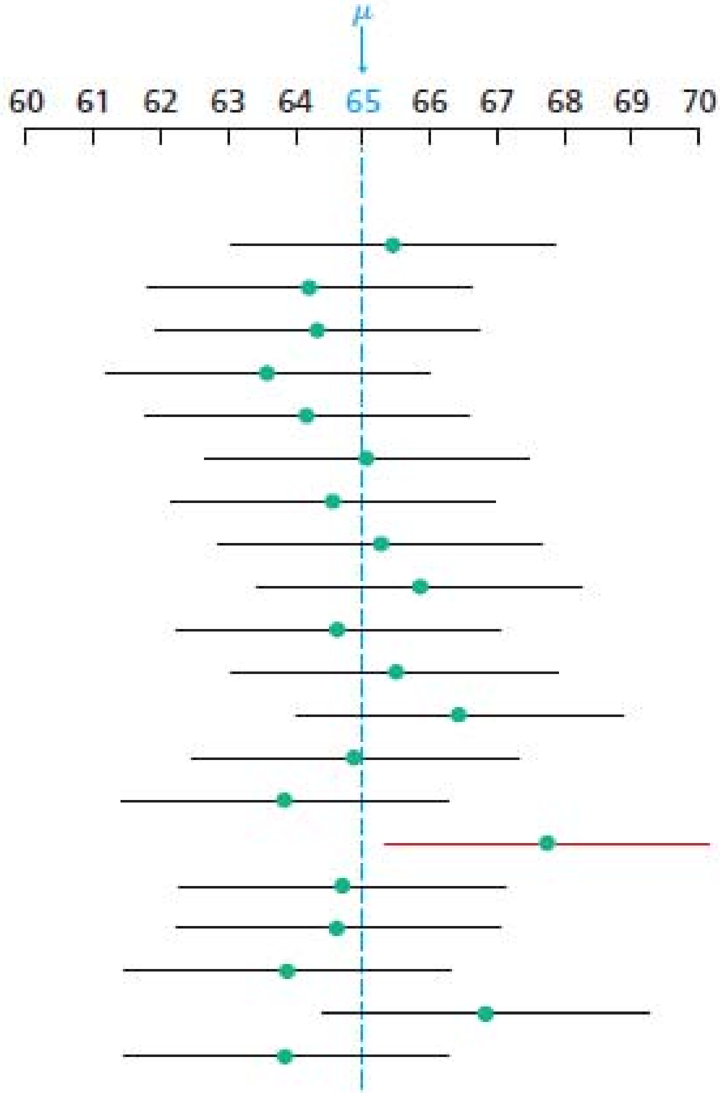
Example 8.3 Continuation of Example 8.1 (p326)

Price of New Mobile Homes, Consider again the prices of new mobile homes.

Figure 8.2

Twenty confidence intervals for the mean price of all new mobile homes, each based on a sample of 36 new mobile homes

Sample	$\bar{x}$	95.44% CI	μ in CI?
1	65.45	63.06 to 67.85	yes
2	64.21	61.81 to 66.61	yes
3	64.33	61.93 to 66.73	yes
4	63.59	61.19 to 65.99	yes
5	64.17	61.77 to 66.57	yes
6	65.07	62.67 to 67.47	yes
7	64.56	62.16 to 66.96	yes
8	65.28	62.88 to 67.68	yes
9	65.87	63.48 to 68.27	yes
10	64.61	62.22 to 67.01	yes
11	65.51	63.11 to 67.91	yes
12	66.45	64.05 to 68.85	yes
13	64.88	62.48 to 67.28	yes
14	63.85	61.45 to 66.25	yes
15	67.73	65.33 to 70.13	no
16	64.70	62.30 to 67.10	yes
17	64.60	62.20 to 67.00	yes
18	63.88	61.48 to 66.28	yes
19	66.82	64.42 to 69.22	yes
20	63.84	61.45 to 66.24	yes

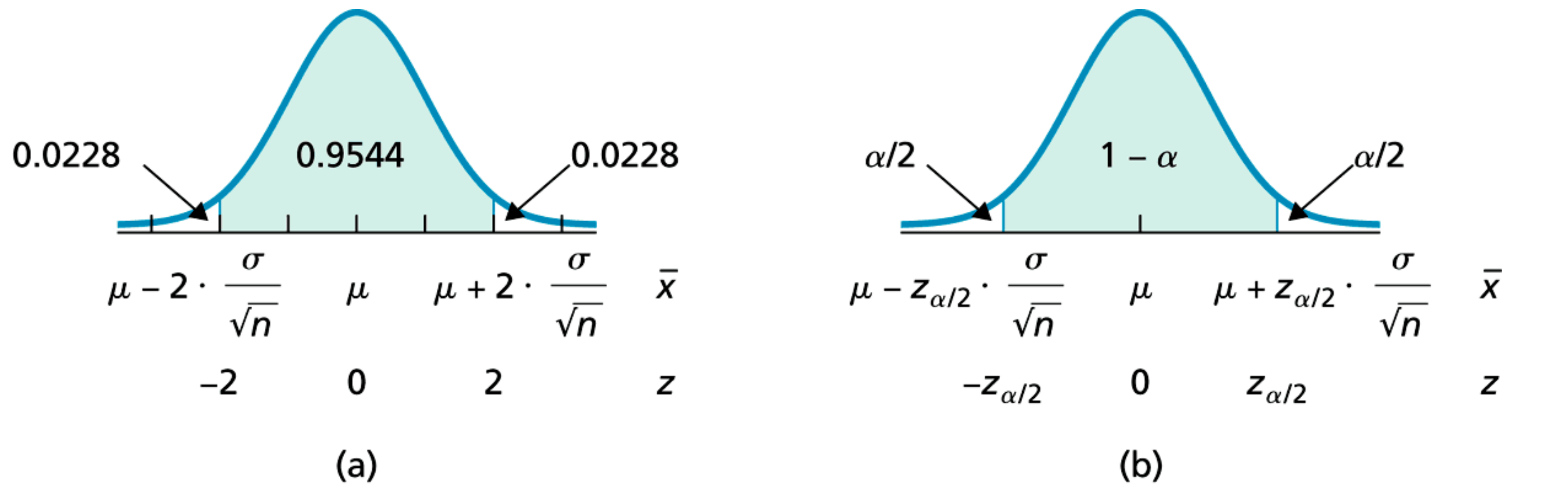


Section 8.2

Confidence Intervals for One Population Mean when Sigma Is Known

Figure 8.3

(a) 95.44% of all samples have means within 2 standard deviations of μ; (b) 100(1 – α)% of all samples have means within $z_{\alpha/2}$ standard deviations of μ



Procedure 8.1

One-Mean z-Interval Procedure

Purpose To find a confidence interval for a population mean, μ

Assumptions

1. Simple random sample

2. Normal population or large sample

3. σ known

Step 1 **For a confidence level of $1 - \alpha$, use Table II to find $z_{\alpha/2}$.**

Step 2 **The confidence interval for μ is from**

$$\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \quad \text{to} \quad \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}},$$

where $z_{\alpha/2}$ is found in Step 1, n is the sample size, and $\bar{x}$ is computed from the sample data.

Step 3 **Interpret the confidence interval.**

Note: The confidence interval is exact for normal populations and is approximately correct for large samples from nonnormal populations.

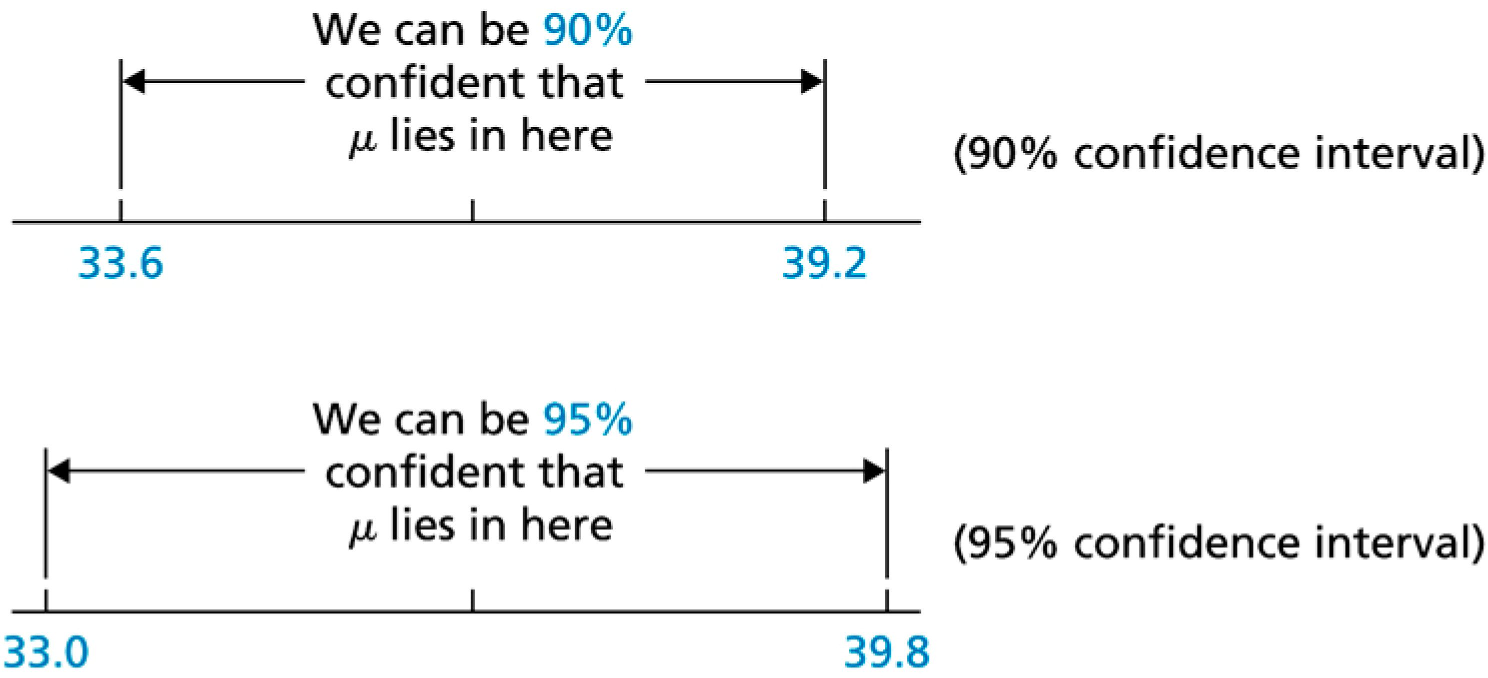
Table 8.3, Page 331

Ages, in years, of 50 randomly selected people in the civilian labor force

22	58	40	42	43
32	34	45	38	19
33	16	49	29	30
43	37	19	21	62
60	41	28	35	37
51	37	65	57	26
27	31	33	24	34
28	39	43	26	38
42	40	31	34	38
35	29	33	32	33

Figure 8.5

90% and 95% confidence intervals for μ, using the data in Table 8.3



Z Values for Some of the More Common Levels of Confidence

Confidence Level	Z Value
90%	1.645
95%	1.96
98%	2.33
99%	2.575

Section 8.3
Margin of Error

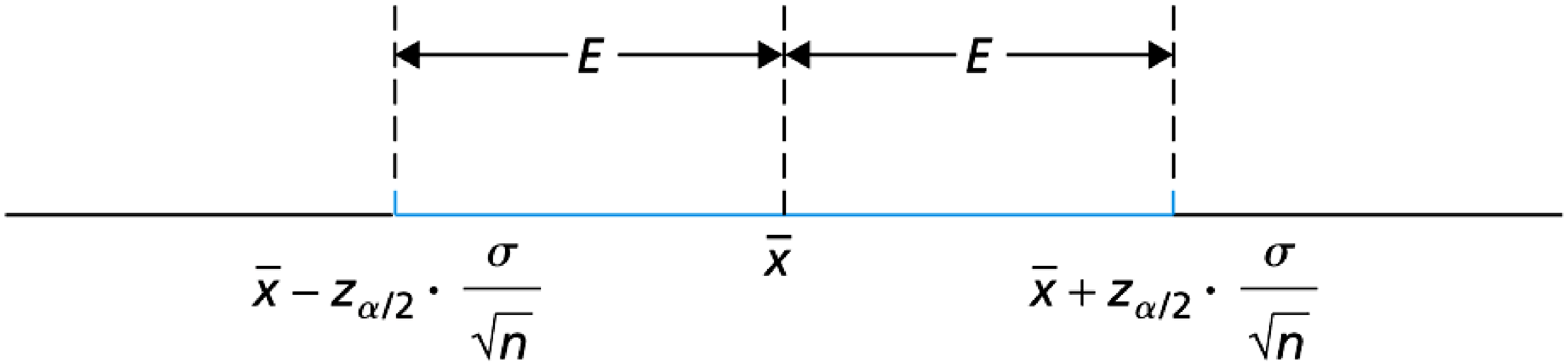
Definition 8.3 & Figure 8.7

Margin of Error for the Estimate of μ

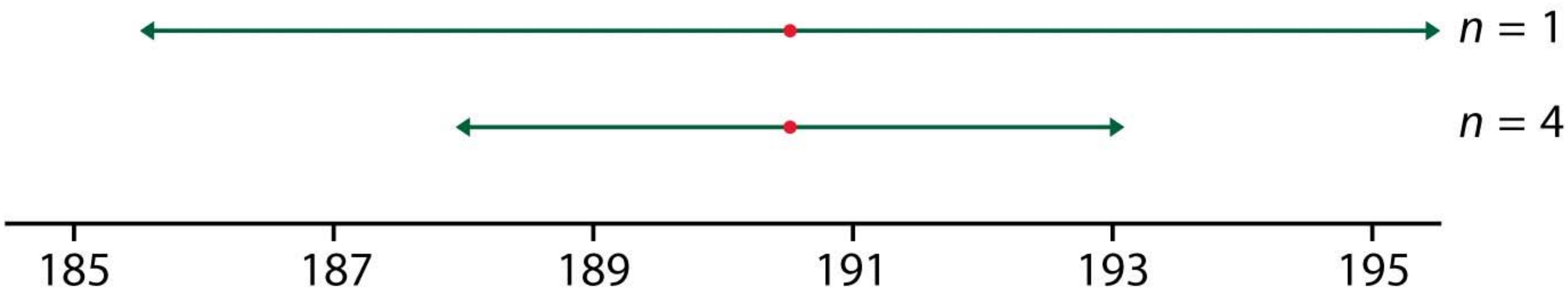
The **margin of error** for the estimate of μ is

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}.$$

Figure 8.7 illustrates the margin of error.



Width of the confidence interval decreases with sample size:



- More sample size-> narrower interval
- Less sample size-> wider interval

What else may we observe?

To decrease margin of error $z^* \dfrac{\sigma}{\sqrt{n}}$

You may do (some of) the following

- Lower confidence level C
- Increase sample size n
- Reduce σ

Example: The weights of full boxes of a certain kind of cereal are normally distributed with a standard deviation of 0.27 oz. A sample of 18 randomly selected boxes produced a mean weight of 9.87 oz. Find a 95% confidence interval for the true mean weight of a box of this cereal.

Solution:

1. *Describe the population parameter of concern*
   The mean, μ, weight of all boxes of this cereal

2. *Specify the confidence interval criteria*

   a. Check the assumptions
      The weights are normally distributed, the distribution of $\bar{x}$ is normal

   b. Identify the probability distribution and formula to be used
      Use the standard normal variable z with $\sigma = 0.27$

   c. Determine the level of confidence, $1 - \alpha$
      The question asks for 95% confidence: $1 - \alpha = 0.95$

Solution Continued

3. *Collect and present information*

   The sample information is given in the statement of the problem

   Given:

$$n = 18; \quad \bar{x} = 9.87$$

4. *Determine the confidence interval*

   a. Determine the confidence coefficient using Z Table.

Solution Continued

b. Find the maximum error of estimate
   Use the maximum error part of the formula for a CI

$$E = z(\alpha/2)\frac{\sigma}{\sqrt{n}} = 1.96\frac{0.27}{\sqrt{18}} = 0.1247$$

c. Find the lower and upper confidence limits
   Use the sample mean and the maximum error:

$$\bar{x} - z(\alpha/2)\frac{\sigma}{\sqrt{n}} \quad \text{to} \quad \bar{x} + z(\alpha/2)\frac{\sigma}{\sqrt{n}}$$

$$9.87 - 0.1247 \quad \text{to} \quad 9.87 + 0.1247$$

$$9.7453 \quad \text{to} \quad 9.9947$$

$$9.75 \quad \text{to} \quad 10.00$$

5. *State the confidence interval*

9.75 to 10.00 is a 95% confidence interval for the true mean weight, μ, of cereal boxes

Example: A random sample of the test scores of 100 applicants for clerk-typist positions at a large insurance company showed a mean score of 72.6. Determine a 99% confidence interval for the mean score of all applicants at the insurance company. Assume the population standard deviation of test scores is 10.5.

Solution:

1. *Parameter of concern*

   The mean test score, μ, of all applicants at the insurance company

2. *Confidence interval criteria*

   a. Assumptions: The distribution of the variable, test score, is not known. However, the sample size is large enough (n = 100) so that the CLT applies

   b. Probability distribution: standard normal variable z with $\sigma = 10.5$

   c. The level of confidence: 99%, or $1 - \alpha = 0.99$

Solution Continued

3. *Sample information*
   Given: $n = 100$ and $\bar{x} = 72.6$

4. *The confidence interval*
   a. Confidence coefficient: $z(\alpha/2) = z(0.005) = 2.58$

   b. Maximum error: $E = z(\alpha/2)\,(\sigma / \sqrt{n}) = (2.58)(10.5 / \sqrt{100}) = 2.709$

   c. The lower and upper limits:
   $72.6 - 2.709 = 69.891$ to $72.6 + 2.709 = 75.309$

5. *Confidence interval*
   With 99% confidence we say, "The mean test score is between 69.9 and 75.3", or "69.9 to 75.3 is a 99% confidence interval for the true mean test score"

Note: The *confidence* is in the *process*. 99% confidence means: if we conduct the experiment over and over, and construct lots of confidence intervals, then 99% of the confidence intervals will contain the true mean value μ.

Choosing sample size:

To have a desired margin of error $m = z^* \dfrac{\sigma}{\sqrt{n}}$

Take sample (at least) of $n = \left(\dfrac{z^* \sigma}{m} \right)^2$

Formula 8.1

Sample Size for Estimating μ

The sample size required for a $(1 - \alpha)$-level confidence interval for μ with a specified margin of error, E, is given by the formula

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2,$$

rounded up to the nearest whole number.

Selecting a Sample Size *n*

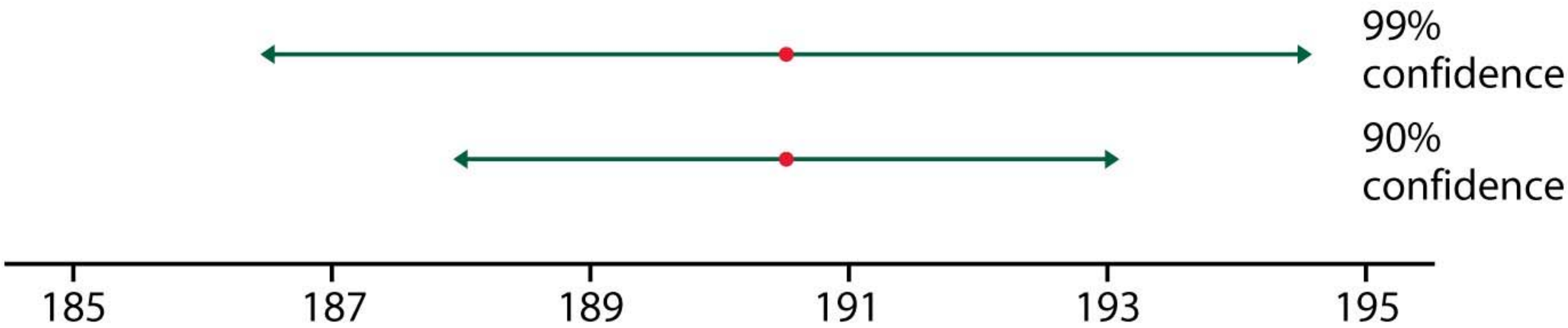
There are 3 factors that determine the size of a sample, none of which has any direct relationship to the size of the population. They are:

- The level of confidence selected.
- The maximum allowable error.
- The variation in the population.

The level of confidence selected

At what value of the confidence level is best suited for a particular study? Any value chose from 0% to 100% can be used, but it depends largely on the person whom is conducting the experiment or the survey (many criteria need to be considered here).

Width of the confidence interval increases with confidence level:



- More confidence -> Wider Interval
- Less confidence-> Narrower Interval

The maximum allowable error

The maximum allowable error, denoted by *E*, is the amount of error those conducting the study are willing to tolerate. The smaller the value of *E* that we expect, the larger the sample size that is needed, which will means a higher cost.

The variation in the population

If the population is widely dispersed or there are large variation between the population members, then we will need a larger sample size to capture all those diversities in a population. However, it the population members are close together and very similar (homogeneous), then we only a small sample size to capture the population estimates.

Steps that we can take to assist in finding *n*

Here three suggestions for finding *n*:

1. Use a comparable study (data or information gain from studies conducted on similar types of study).
2. Use a range-based approach (find the two extreme values in the population and assumed a normal distribution).
3. Conduct a pilot study (use a small sample set that is representative of the population, and get an estimate from the study).

Determining Sample Size when Estimating μ

- **Z formula**

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

- **Error of Estimation (tolerable error)**

$$E = \overline{X} - \mu$$

- **Estimated Sample Size**

$$n = \frac{Z_{\frac{\alpha}{2}}^{2}\sigma^{2}}{E^{2}} = \left(\frac{Z_{\frac{\alpha}{2}}\sigma}{E}\right)^{2}$$

- **Estimated σ**

$$\sigma \approx \frac{1}{4} range$$

EXAMPLE

A consumer group would like to estimate the mean monthly electricity charge for a single family house in July within \$5 using a 99 percent level of confidence. Based on similar studies the standard deviation is estimated to be \$20.00. How large a sample is required?

$$n = \left(\frac{(2.58)(20)}{5} \right)^2 = 107$$

Sample Size When Estimating μ: Example

$$E = 1, \quad \sigma = 4$$

$$90\% \text{ confidence } \Rightarrow Z = 1.645$$

$$n = \frac{Z^2_{\frac{\alpha}{2}} \sigma^2}{E^2}$$

$$= \frac{(1.645)^2 (4)^2}{1^2}$$

$$= 43.30 \text{ or } 44$$

Solution for Demonstration Problem

$$E = 2, \quad range = 25$$

$$95\% \text{ confidence } \Rightarrow Z = 1.96$$

$$estimated\ \sigma: \quad \frac{1}{4} range = \left(\frac{1}{4}\right)(25) = 6.25$$

$$n = \frac{Z^2 \sigma^2}{E^2}$$

$$= \frac{(1.96)^2 (6.25)^2}{2^2}$$

$$= 37.52 \quad \text{or} \quad 38$$

Section 8.4

Confidence Intervals for One Population Mean When Sigma Is Unknown

Standard Error of the Sample Means

- If σ is not known and ($n \geqq 30$ or population is known to be normal), the standard deviation of the sample, designated s, is used to approximate the population standard deviation. The formula for the standard error is:

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

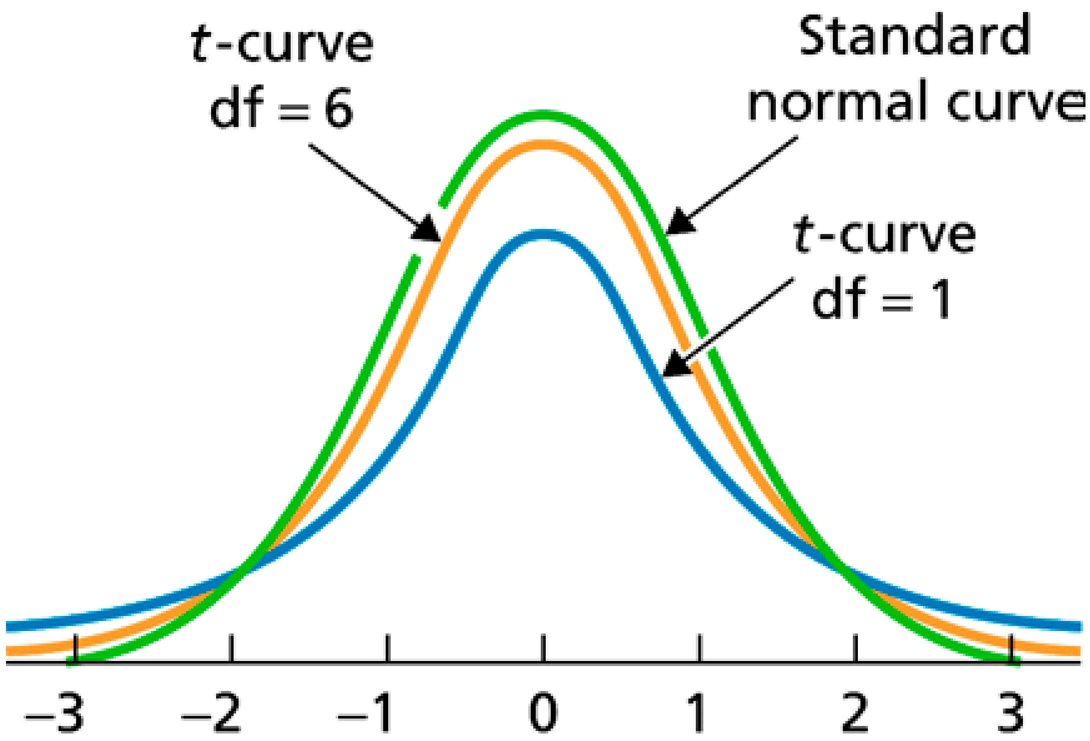
Point and Interval Estimates

- If the population standard deviation is unknown and the distribution is normal, then we use the *t* distribution.

$$\overline{X} \pm t \frac{s}{\sqrt{n}}$$

Figure 8.8

Standard normal curve and two t-curves



Key Fact 8.6

Basic Properties of *t*-Curves

Property 1: The total area under a *t*-curve equals 1.

Property 2: A *t*-curve extends indefinitely in both directions, approaching, but never touching, the horizontal axis as it does so.

Property 3: A *t*-curve is symmetric about 0.

Property 4: As the number of degrees of freedom becomes larger, *t*-curves look increasingly like the standard normal curve.

Table 8.4

Values of t_α

df	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$	df
⋮	⋮	⋮	⋮	⋮	⋮	⋮
12	1.356	1.782	2.179	2.681	3.055	*12*
13	1.350	1.771	2.160	2.650	3.012	*13*
14	1.345	1.761	2.145	2.624	2.977	*14*
15	1.341	1.753	2.131	2.602	2.947	*15*
⋮	⋮	⋮	⋮	⋮	⋮	⋮

TABLE D *t* distribution critical values

	Upper tail probability *p*											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
z^*	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level *C*											

Procedure 8.2

One-Mean *t*-Interval Procedure

Purpose To find a confidence interval for a population mean, μ

Assumptions

1. Simple random sample
2. Normal population or large sample
3. σ unknown

Step 1 **For a confidence level of $1-\alpha$, use Table IV to find $t_{\alpha/2}$ with df $= n-1$, where n is the sample size.**

Step 2 **The confidence interval for μ is from**

$$\bar{x} - t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \quad \text{to} \quad \bar{x} + t_{\alpha/2} \cdot \frac{s}{\sqrt{n}},$$

where $t_{\alpha/2}$ is found in Step 1 and $\bar{x}$ and s are computed from the sample data.

Step 3 **Interpret the confidence interval.**

Note: The confidence interval is exact for normal populations and is approximately correct for large samples from nonnormal populations.

EXAMPLE

The Dean of the Business School wants to estimate the mean number of hours worked per week by students. A sample of 49 students showed a mean of 24 hours with a standard deviation of 4 hours. Do we know what is the population mean? Else, can we 95% confident that the population mean is in what range?

In this case, the value of the *population mean* is not known. Our best estimate of this value is the sample mean of 24.0 hours.

Example *continued*

Furthermore, σ is unknown, but sample size is greater than 30. Hence to find the 95 percent confidence interval for the population mean.

$$\bar{X} \pm t_{(48,0.025)} \frac{s}{\sqrt{n}} = 24.00 \pm 2.011 \frac{4}{\sqrt{49}}$$

$$= 24.00 \pm 1.15$$

The confidence limits range from 22.85 to 25.15.

About 95 percent of the similarly constructed intervals include the population parameter.

One-sample *t* Confidence Interval

- Suppose a SRS of size *n* is drawn from population having unknown mean μ. A level C confidence interval for μ is

$$\bar{x} \pm t^* \frac{s}{\sqrt{n}}, \quad \text{or} \quad \left[\bar{x} - t^* \frac{s}{\sqrt{n}}, \bar{x} + t^* \frac{s}{\sqrt{n}}\right]$$

  Here t^* is the value for the t density curve with df=n-1. The area between $-t^*$ and t^* is C.

- The interval is exact for normal population and approximately correct for large *n* in other cases.
- Note the standard error in the denominator.

Table 8.5 & Figure 8.10 Page 346

Example Pickpocket Offenses: The FBI compiles data on robbery and property crimes on publishes the information online. A simple random sample of pickpocket offenses yielded the losses, in $, shown in table below. Use the data to find a 95% confidence interval for the mean loss of all pickpocket offenses.

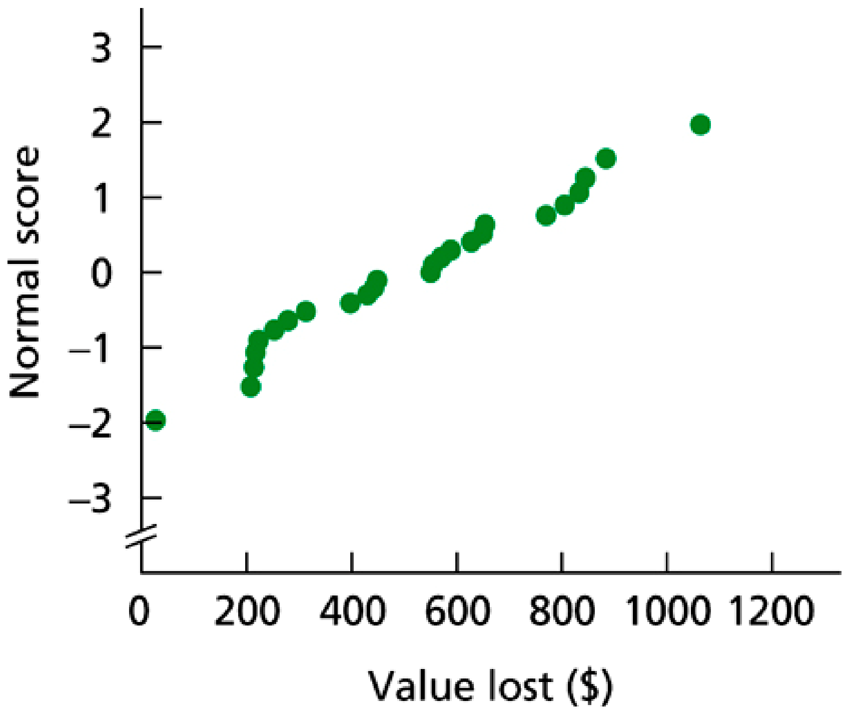
Losses ($) for a sample of 25 pickpocket offenses:

447	207	627	430	883
313	844	253	397	214
217	768	1064	26	587
833	277	805	653	549
649	554	570	223	443

Table 8.5 & Figure 8.10 Page 346

N = 25	Normal Score	N = 25	Normal Score
26	-1.97	554	0.1
207	-1.52	570	0.2
214	-1.26	587	0.3
217	-1.06	627	0.41
223	-0.9	649	0.52
253	-0.76	653	0.63
277	-0.63	768	0.76
313	-0.52	805	0.9
397	-0.41	833	1.06
430	-0.3	844	1.26
443	-0.2	883	1.52
447	-0.1	1064	1.97
549	0		

Normal probability plot of the loss data in Table 8.5



Example

- ✓ **Example:** A study is conducted to learn how long it takes the typical tax payer to complete their federal income tax return. A random sample of 17 income tax filers showed a mean time (in hours) of 7.8 and a standard deviation of 2.3. Find a 95% confidence interval for the true mean time required to complete a federal income tax return. Assume the time to complete the return is normally distributed.

Solution:

1. *Parameter of Interest*
   The mean time required to complete a federal income tax return

2. *Confidence Interval Criteria*
   a. Assumptions: Sampled population assumed normal, σ unknown
   b. Test statistic: t will be used
   c. Confidence level: $1 - \alpha = 0.95$

Solution Continued

3. *The Sample Evidence*: $n = 17, \quad \bar{x} = 7.8, \quad \text{and} \quad s = 2.3$

4. *The Confidence Interval*
   a. Confidence coefficients: $t(\text{df}, \alpha/2) = t(16, 0.025) = 2.12$
   b. Maximum error:
   $$E = t(16, 0.025)\frac{s}{\sqrt{n}} = (2.12) \times \frac{2.3}{\sqrt{17}} = (2.12)(0.5578) = 1.18$$
   c. Confidence limits: $\bar{x} - E \quad \text{to} \quad \bar{x} + E$
   $$7.8 - 1.18 \quad \text{to} \quad 7.8 + 1.18$$
   $$6.62 \quad \text{to} \quad 8.98$$

5. *The Results*:

6.62 to 8.98 is the 95% confidence interval for μ

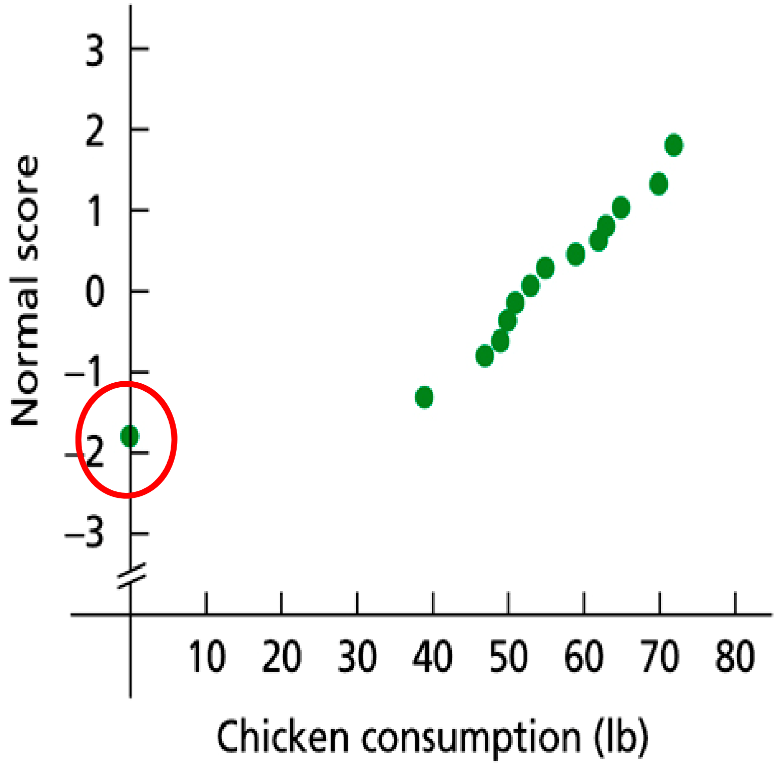
Figure 8.11 Page 347

Example Chicken Consumption: A report publishes data on chicken consumption. Table below shows a year's chicken consumption, in pounds, for 17 randomly selected people. Find a 90% confidence interval for the year's mean chicken consumption.

57	63	63
72	91	0
60	55	73
69	49	61
65	59	82
75	80	

Example 8.11, p347

Figure 8.11

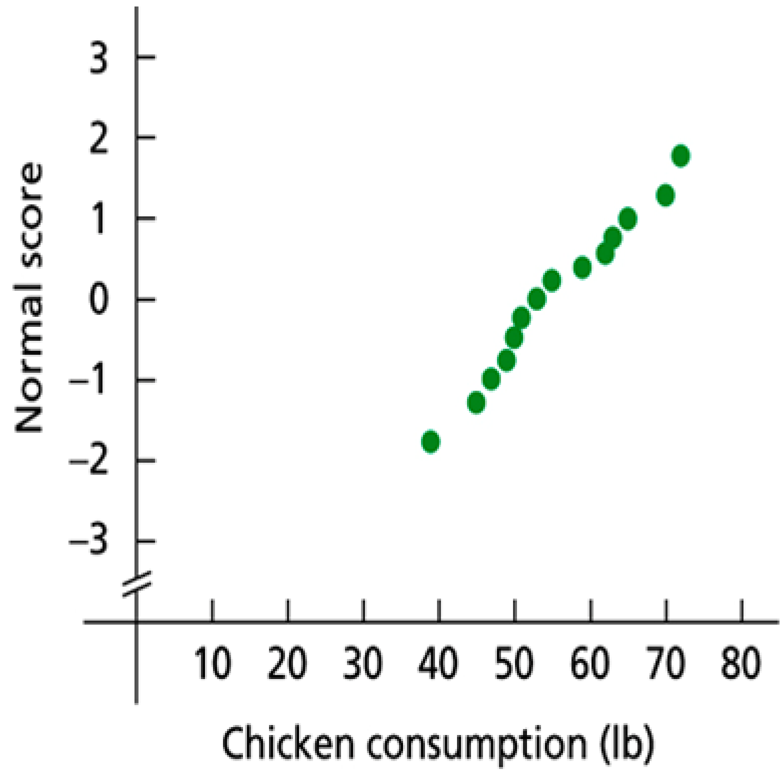


(a)

N = 17	Normal Score
0	-1.8
49	-1.32
55	-1.03
57	-0.8
59	-0.62
60	-0.45
61	-0.29
63	-0.15
63	0
65	0.15
69	0.29
72	0.45
73	0.62
75	0.8
80	1.03
82	1.32
91	1.8

Example 8.11, p347

Figure 8.11

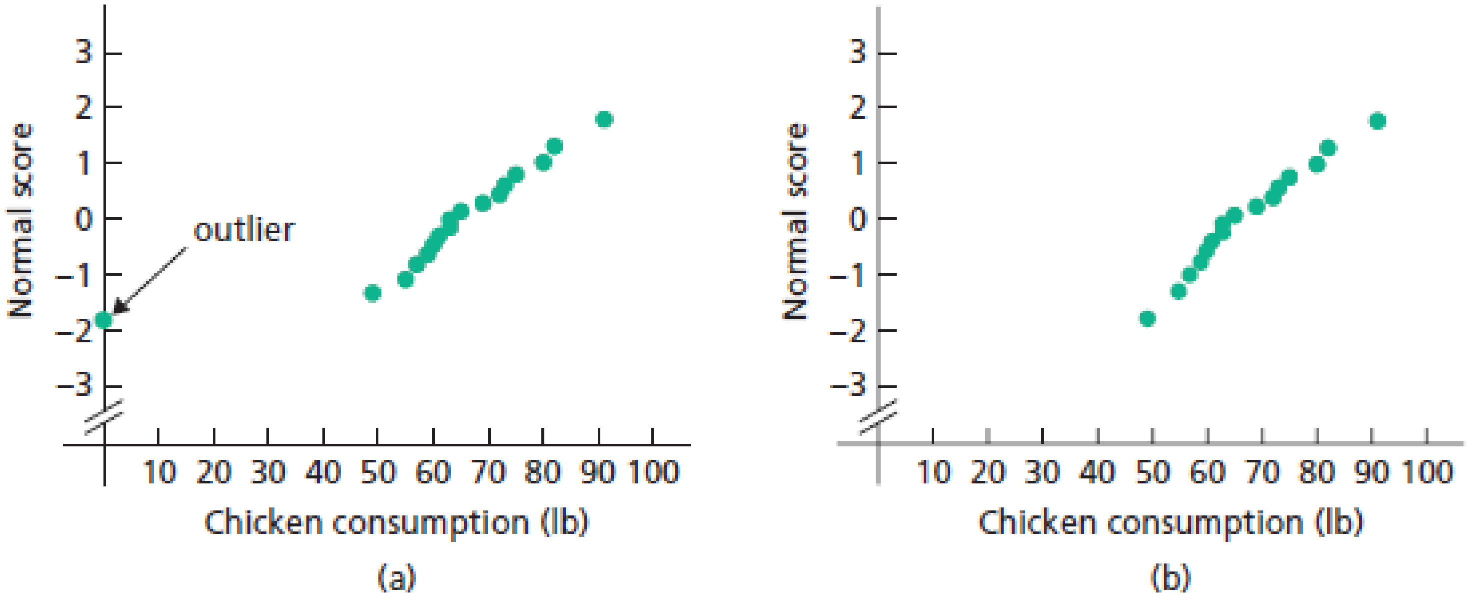


(b)

N = 16	Normal Score
49	-1.77
55	-1.28
57	-0.99
59	-0.76
60	-0.57
61	-0.39
63	-0.23
63	-0.08
65	0.08
69	0.23
72	0.39
73	0.57
75	0.76
80	0.99
82	1.28
91	1.77

Figure 8.11 Page 347

Normal probability plots for chicken consumption:
(a) original data and (b) data with outlier removed



Key Fact 8.1

- Small sample ($n < 15$) – population must be normal.
- Moderate sample ($30 > n \geq 15$) – beware of outliers or the case that population deviate far from normal.
- Large sample ($n \geq 30$) – beware of outliers and if there is justifications to remove the outliers, then we should do so, before the test. Else, we should observe the inferences of outliers.