Chapter Seven
Estimating Population Parameters

Goals of chapter:

After completing this chapter, you should be able to**:**

- Distinguish between a point estimate and a confidence interval estimate.
- Construct and interpret a confidence interval estimate for a single population mean using both the z and t distributions.
- Determine the required sample size to estimate a single population mean within a specified margin of error.
- Form and interpret a confidence interval estimate for a single population proportion.

Point and Interval Estimates

- A point estimate is a *single* number.
  We can estimate a Population Parameter (Unknown) with a Sample Statistic

Mean: μ. The point estimate is the mean $\bar{x} = \frac{\sum x}{n}$.

Variance σ^2. The point estimate is the variance $S^2 = \frac{\sum x^2 - n\bar{x}^2}{n-1}$.

Proportion p . The point estimate is the proportion $\bar{p} = \frac{x}{n}$.

- A *confidence interval* provides additional information about variability.
  An interval estimate provides more information about a population characteristic than does a point estimate. Such interval estimates are called confidence intervals (C.I.).

Confidence Interval Estimate

An interval gives a range of values:

- Takes into consideration variation in the sample statistics from sample to sample.
- Based on observation from 1 sample.
- Gives information about closeness to an unknown population parameter.
- It is stated in terms of confidence level.
- Never 100% sure.

Point Estimate ± (Critical Value)(Standard Error)

Standard Error of the estimate: The standard deviation of the sample statistic, denoted by SE(estimate). It is a standard deviation calculated for a sample statistic and not for some raw data.

Confidence Level (1-α)

Confidence in which the interval will contain the unknown population parameter. It is a percentage (less than 100%).

Example: Suppose confidence level = 95%

Also written (1 - α) = 0.95, then α = 0.05 (significance level).

- A relative frequency interpretation:
  - In the long run, 95% of all the confidence intervals that can be constructed will contain the unknown true parameter.
- A specific interval either will contain or will not contain the true parameter
  - No probability is involved in a specific interval.

Cases for Confidence Interval for the true population mean (μ):

Case I:

A (1-α) 100% Confidence Interval for μ, when (σ known):

Assumptions

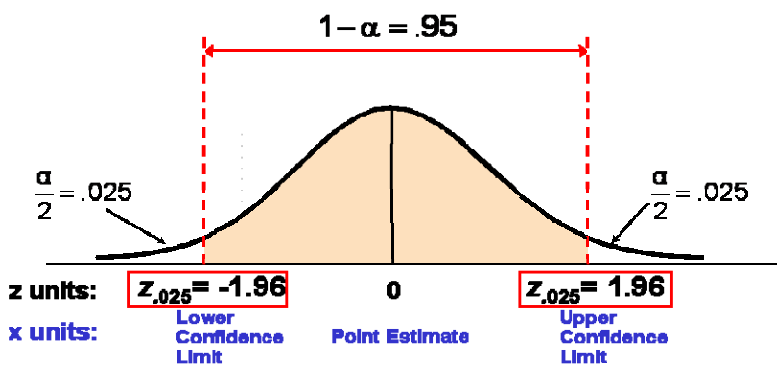
1. Population standard deviation (σ) is known.
2. Population is normally distributed.

$$\overline{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Finding the Critical Value

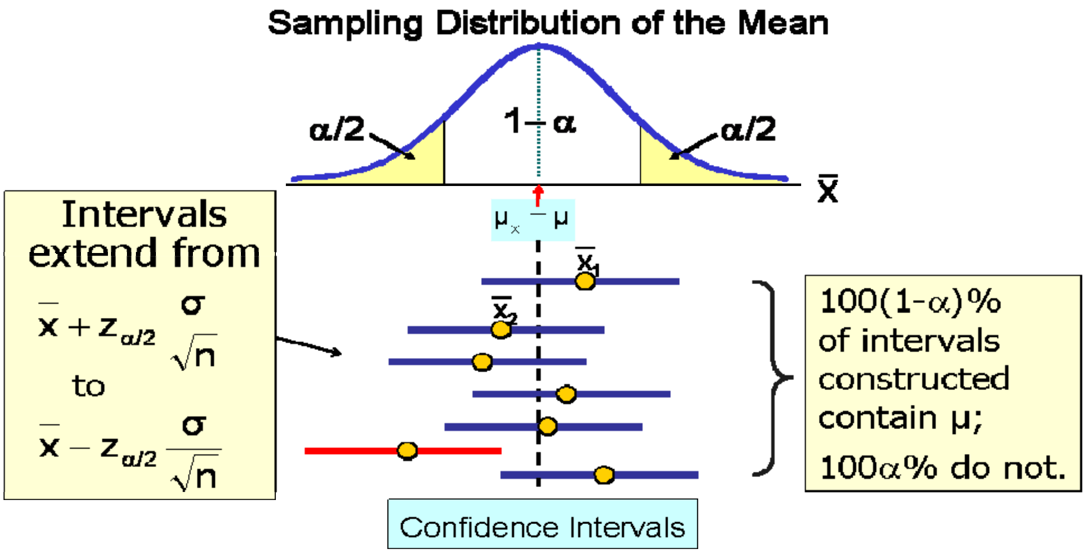
Consider a 95% confidence interval:

$z_{\alpha/2} = \pm 1.96$



Confidence Level	$1-\alpha$	$z_{\alpha/2}$
80%	0.80	1.28
90%	0.90	1.645
95%	0.95	1.96
98%	0.98	2.33
99%	0.99	2.575
99.8%	0.998	3.08
99.9%	0.999	3.27

Interval and Level of Confidence



Margin of Error (*e*): the amount added and subtracted to the point estimate to form the confidence interval

$$e = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Factors Affecting Margin of Error

- **Data variation, σ :** e ↓ as σ ↓
- **Sample size, n :** e ↓ as n ↑
- **Level of confidence, 1 - α :** e ↓ if 1 - α ↓

Example: Determine a 95% confidence interval for the true mean of the population: $n = 11$, $\bar{x} = 2.2$, $\sigma = 0.35$.

$$1-\alpha = 0.95 \Rightarrow \alpha = 0.05 \quad Z_{\frac{\alpha}{2}} = Z_{\frac{0.05}{2}} = 1.96$$

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$= 2.20 \pm 1.96\ (0.35/\sqrt{11})$$

$$= 2.20 \pm 0.2068$$

$$= [1.9932, 2.4068]$$

- We are 95% confident that the true mean resistance is between 1.9932 and 2.4068 ohms.
- Although the true mean may or may not be in this interval, 95% of ALL the intervals formed in this manner will contain the true mean.

Example: Q 7.5 Page 282

Determine the margin of error:

a. Confidence level = 0.98, n = 13 and σ = 15.68

$$e = \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = \pm(2.33)\left(\frac{15.68}{\sqrt{13}}\right) = \pm 10.1328$$

b. Confidence level = 0.99, n = 25 and σ = 3.47

$$e = \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = \pm(2.575)\left(\frac{3.47}{\sqrt{25}}\right) = \pm 1.78705$$

c. Confidence level = 0.98, standard error = 2.356

$$e = \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = \pm(2.33)(2.356) = \pm 5.4895$$

Example: Q 7.6 Page 282

Determine a 99% confidence interval for the true mean of the population

$n = 500, \quad \bar{x} = 1.22, \quad \sigma = 34.6$

$$1-\alpha = 0.99 \Rightarrow \alpha = 0.01 \quad Z_{\frac{\alpha}{2}} = Z_{\frac{0.01}{2}} = 2.757$$

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \Rightarrow 34.6 \pm 2.575\ (1.22/\sqrt{500}) \Rightarrow 34.6 \pm 0.1405$$

$$[34.4595\ ,\ 34.7405]$$

- We are 99% confident that the true mean is between 34.45 and 34.74.
- Although the true mean may or may not be in this interval, 99% of ALL the intervals formed in this manner will contain the true mean.

Case II:

A (1-α) 100% Confidence Interval for μ, when (σ unknown) and the sample size is large

Assumptions

1. Population standard deviation (σ) is unknown.
2. Population is normally distributed.
3. Sample size large $(n \geq 30)$.

$$\overline{X} \pm z_{\alpha/2}\frac{s}{\sqrt{n}}$$

Exemple: A random sample of n = 49 has $\bar{x}$ = 50 and s = 8. Form a 95% confidence interval for μ.

Assumptions:

1. Sample size large $(n \geq 30)$.
2. Population is normally distributed.
3. Population standard deviation σ is unknown.

$$1-\alpha = 0.95 \Rightarrow \alpha = 0.05 \quad Z_{\frac{\alpha}{2}} = Z_{\frac{0.05}{2}} = 1.96$$

$$\overline{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$= 50 \pm 1.96\ (8 / \sqrt{49})$$

$$= 50 \pm 2.24 = [47.76, 52.24]$$

- We are 95% confident that the true mean is between 47.76 and 52.24.
- Although the true mean may or may not be in this interval, 95% of ALL the intervals formed in this manner will contain the true mean.

Example: Q7.14 Page 283

$n = 360,\ \bar{x} = 33.4,\quad s = 11.2$

a. Find 90% C.I for the true mean μ.

Solution:

Assumptions:

1. Sample size large $(n \geq 30)$.
2. Population is normally distributed.
3. Population standard deviation σ is unknown.

$$1 - \alpha = 0.90 \Rightarrow \alpha = 0.1 \quad Z_{\frac{\alpha}{2}} = Z_{\frac{0.1}{2}} = 1.645$$

$$\overline{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$= 33.4 \pm 1.645\ (11.2 / \sqrt{360})$$

$$= 33.4 \pm 0.971 = [32.429\ ,\ 34.371]$$

- We are 90% confident that the true mean number of checks written by ALL customers is between 32.429 and 34.371.
- Although the true mean may or may not be in this interval, 90% of intervals formed in this manner will contain the true mean.

Example: Q7.16 Page 283

$n = 300,\ \bar{x} = \$14.23,\quad s = \$3.0$

a. A 90% C.I for the true average μ

$$1 - \alpha = 0.90 \Rightarrow \alpha = 0.1 \quad Z_{\frac{\alpha}{2}} = Z_{\frac{0.1}{2}} = 1.645$$

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 14.23 \pm 1.645\ (3 / \sqrt{300})$$

$$= 14.23 \pm 0.28493326 = [13.9451\ ,\ 14.5149]$$

We are 90% confident that the true mean is between 13.9451 and 14.5149.

b. The 90% C.I for the total range at $528 is

Lower Limit : $(528)13.9451 = 7363.01$ and Upper Limit : $(528)14.5149 = 7663.86$.

Then the C.I is [7363.01 , 7663.86]

The value 7394 $\in$ [7363.01 , 7663.86], there is no reson to belive that this revenue is out of limit.

Case III:

A (1-α) 100% Confidence Interval for μ, when the sample size is small and (σ is unknown).

Assumptions

1. Sample size is small $(n < 30)$.
2. Population is normally distributed.
3. Population standard deviation σ is unknown.

$$\bar{x} \pm t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$$

Where the margin error: $e = \pm t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}$

Student *t* – distribution:

It has a bell shape like the normal distribution. It has the same properties as the normal distribution, mean = median = mode. And it is depend on its ONLY parameter, namely its dgree of feedom (*d.f* = *n* – 1).

As *n* increases, ***t*** – destribution approaches the ***Z*** – distribution.

Example : Q 7.1 Page 281

NOTE: To solve any part, go to ***t*** – table, the solution is the intersection point between the *conf. Level* with the *d.f.*

Determine the critical values for t - distribution

1. Confidence level $= 0.95$, $n = 26$

$$t_{\frac{\alpha}{2}, n-1} = t_{\frac{0.05}{2}, 25} = 2.0595$$

2. Confidence level $= 0.9$, $n = 31$

$$t_{\frac{\alpha}{2}, n-1} = t_{\frac{0.1}{2}, 30} = 1.6973$$

3. Confidence level $= 0.98$, $n = 15$

$$t_{\frac{\alpha}{2}, n-1} = t_{\frac{0.02}{2}, 14} = 2.6245$$

Example: Q7.5 Page 282 (part b)

Find the margin of error.

Confidence level $= 0.95$, $n = 21$, $\bar{x} = 13.9$ and $s = 2.33$

$$t_{\frac{\alpha}{2}, n-1} = t_{\frac{0.05}{2}, 20} = 2.0860 \implies e = \pm(2.0860)\left(\frac{2.33}{\sqrt{21}}\right) = \pm 1.0606.$$

Example: Q7.10 Page 282

Construct 90% C.I for the population mean.

Sample: 11, 14, 10, 12, 11, 11, 12, 12, 15

<u>Solution:</u>

Assumptions

1. Sample size is small $(n < 30)$.
2. Population is normally distributed.
3. Population standard deviation σ is unknown.

Sample mean : $\bar{x} = \frac{\sum x}{n} = \frac{108}{9} = 12$.

Sample standard deviation :

$$s = \sqrt{\frac{\sum x^2 - n\bar{x}^2}{n-1}} = \sqrt{\frac{1316-(9)(12)^2}{9-1}} = \sqrt{\frac{1316-1296}{8}} = 1.58113883.$$

$$t_{\frac{\alpha}{2},n-1} = t_{\frac{0.1}{2},8} = 1.8595$$

$$\bar{x} \pm t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}$$

$$12 \pm (1.8595)\left(\frac{1.58113883}{\sqrt{9}}\right)$$

$$12 \pm 0.980042551 = [11.019957 \ , 12.98042551]$$

7.2: Determining the Required Minimum Sample Size:

The required sample size can be found to reach a desired margin of error (e) and a level of confidence (1 - α).

The sample size depends on:

1. The cost.
2. The confifence level.
3. The margin of error.
4. The standard deviation.
5.

Required MINIMUM sample size, (σ known):

$$n \geq \frac{z_{\alpha/2}^2 \ \sigma^2}{e^2} = \left(\frac{z_{\alpha/2} \ \sigma}{e}\right)^2$$

Required sample size, (σ unknown):

If σ was unknown, σ can be estimated in the formula of the sample size:

- Use a value for σ that is expected to be at least as large as the true σ.
- Select a <u>**pilot**</u> sample and estimate σ with the sample standard deviation, s.

Then $n \geq \frac{z_{\alpha/2}^2 \ s^2}{e^2} = \left(\frac{z_{\alpha/2} \ s}{e}\right)^2$

Example: If σ = 45, what sample size is needed to be 90% confident of being correct within ± 5?

<u>Solution:</u> $1-\alpha = 0.9 \Rightarrow \alpha = 0.1 \Rightarrow Z_{\frac{\alpha}{2}} = Z_{\frac{0.1}{2}} = 1.645$

$$n \geq \left(\frac{z_{\alpha/2}\,\sigma}{e}\right)^2 = \left(\frac{(1.645)(45)}{5}\right)^2 = 219.19$$

So the required minimum sample size is ***n* = 220 (Always rounded up)**.

Example: Q7.25 Page 288

$\sigma = 40$, $C.I = 95\%$, $e = \pm 2.5$, find the required sample size

Solution: $1-\alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow Z_{\frac{\alpha}{2}} = Z_{\frac{0.05}{2}} = 1.96$

$$n \geq \left(\frac{z_{\alpha/2}\,\sigma}{e}\right)^2 = \left(\frac{(1.96)(40)}{2.5}\right)^2 = 983.4496$$

So the required minimum sample size is ***n* = 994**.

Example: Q7.27 Page 288: (pilot sample of size 10)

$\bar{x} = 2945.2, \quad s = 246.667$, $C.I = 90\%,\ e = \pm 60$

$1-\alpha = 0.9 \Rightarrow \alpha = 0.1 \ \Rightarrow Z_{\frac{\alpha}{2}} = Z_{\frac{0.1}{2}} = 1.645$

$$n \geq \left(\frac{z_{\alpha/2}\,\sigma}{e}\right)^2 = \left(\frac{(1.645)(246.667)}{60}\right)^2 = 45.73$$

So the required sample size is ***n* = 46**. However, the pilot sample is a part of the required sample, so the number of additional ovservation is 36 observations.

7.3: Confidence Intervals for the Population Proportion, *p*:

An interval estimate for the population proportion $p = \frac{X}{N}$ can be calculated by adding an allowance for uncertainty to the sample proportion $\bar{p} = \frac{x}{n}$.

Recall that: The distribution of the sample proportion is approximately normal if the sample size is large $(np \geq 5)$ and $n(1-p) \geq 5$, with

1. Mean $\mu_{\bar{p}} = p$.
2. Standard deviation $\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}$.

But $p = \frac{X}{N}$ is unknown, We will estimate standard deviation with sample data

$$s_{\bar{p}} = \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

Then,

A (1-α) 100% Confidence Interval for the true proportion is given by:

$$\bar{p} \pm z_{\alpha/2}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

where

- z is the standard normal value for the level of confidence desired.

- $\overline{p} = \frac{x}{n}$ is the sample proportion.
- n is the sample size.

Note: Increases in the sample size reduce the width of the confidence interval

Finding the Required Sample Size for proportion problems

1. The margin of error: $e = \pm z_{\alpha/2}\sqrt{\frac{\overline{p}(1-\overline{p})}{n}}$.
2. The sample size: $n \geq \frac{z_{\alpha/2}^2\, p(1-p)}{e^2}$.

p can be estimated with a *pilot sample*, if necessary (or use p = 0.50).

Example: P.P Page 39
A random sample of 100 people shows that 25 are left-handed.
Form a 95% confidence interval for the true proportion of left-handers
Solution:

$$n = 100 \;\; and \;\; x = 25 \qquad then \qquad \overline{p} = \frac{x}{n} = \frac{25}{100} = 0.25$$

$$1-\alpha = 0.95 \;\Rightarrow\; \alpha = 0.05 \;\Rightarrow\; Z_{\frac{0.05}{2}} = 1.96$$

95% confidence interval for the true proportion

$$\overline{p} \pm z_{\alpha/2}\sqrt{\frac{\overline{p}(1-\overline{p})}{n}} = 0.25 \pm 1.96\sqrt{\frac{0.25(0.75)}{100}}$$

$$= 0.25 \;\pm\; 0.0849 = [\,0.1651 \,,\, 0.3349\,]$$

- We are 95% confident that the true percentage of left-handers in the population is between 16.51% and 33.49%.

Example: Q 7.41 Page 294
Compute a 95% C.I for the true proportion, based on sample of size 400, when the sample proportion is 0.3
Solution:

$$n = 400 \;\; and \;\; \overline{p} = 0.3$$

$$1-\alpha = 0.95 \;\Rightarrow\; \alpha = 0.05 \;\Rightarrow\; Z_{\frac{0.05}{2}} = 1.96$$

95% confidence interval for the true proportion

$$\overline{p} \pm z_{\alpha/2}\sqrt{\frac{\overline{p}(1-\overline{p})}{n}} = 0.3 \pm 1.96\sqrt{\frac{0.3(0.7)}{400}}$$

$$= \;0.3 \;\pm\; 0.044909 \;= [0.2550907 \,,\, 0.344909]$$

Example: Q 7.42 Page 294
Detrmine the required sample size needed to estemate a population proportion, when the margin error 0.03, the Confidence level is 95%, and p = 0.5.

Solution:

$1-\alpha = 0.95 \;\Rightarrow\; \alpha = 0.05 \;\Rightarrow\; Z_{\frac{0.05}{2}} = 1.96$ and $e = 0.03$

$$n = \frac{z_{\alpha/2}^2 \; p(1-p)}{e^2} = \frac{(1.96)^2 (0.5)(0.5)}{(0.03)^2} = 1067.11$$

So the required sample size is ***n* = 1068**.

Example: Q 7.46 Page 294

A random sample of size 900 was selected from a population, the sample contained 750 items with a particular attribute.

1. Construct 90% C.I for the true proportion p

   Solution:

   $$n = 900 \quad , x = 750 \quad , \quad \bar{p} = \frac{x}{n} = \frac{750}{900} = 0.833$$

   $$1-\alpha = 0.9 \;\Rightarrow\; \alpha = 0.1 \;\Rightarrow\; Z_{\frac{0.1}{2}} = 1.645$$

95% confidence interval for the true proportion

$$\bar{p} \pm z_{\alpha/2}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} = 0.833 \pm 1.645\sqrt{\frac{0.833(0.167)}{900}}$$

$$= \; 0.833 \;\pm\; 0.02045151 = [0.81254849 \;,\; 0.85345151]$$

2. Refere to part (a). Find the size of the sample would be needed to cut the margin error to half?

   Solution:

   The margin error:

   Since the proportion for the population unknown , use the proportion for the sample

   $$e = \pm\, z_{\alpha/2}\sqrt{\frac{\bar{p}(1-p)}{n}} = \pm 1.645\sqrt{\frac{0.833(0.167)}{900}} = \pm 0.02045151$$

So the new margin error

$$e = \frac{e}{2} = \frac{0.02045151}{2} = 0.010225755$$

The sample size

$$n = \frac{z_{\alpha/2}^2 \; p(1-p)}{e^2} = \frac{(1.645)^2 (0.833)(0.167)}{(0.010225755)^2} = 3600$$

So, the additional units = 3600 – 900 = 2700.

Example: Q 7.58 Page 294

n = 1000 passengers, x = 345 passengers have more than one bag.

a. Conctruct 95% C.I for the proportion for the passengers with more one bag $\overline{p}=\dfrac{345}{1000}=0.345$, $1-\alpha=0.95 \Rightarrow \alpha=0.05 \Rightarrow Z_{\frac{0.05}{2}}=1.96$

$0.345 \pm 1.96(\sqrt{[(0.345)(1-0.345]/1000})$ = [0.3155 , 0.3745].

With 95% confidence we believe that the proportion of customers that carry more than one bag on the airline is between 0.3155 and 0.3745.

b. Conctruct 95% C.I for the number of the passengers with more one bag

[(0.3155) – (568) , (0.3745) + (568)] = [179.20 , 212.716].

c. n = 1000 , 690 males and 310 femals, 280 males have more one bag. Conctruct a 95% C.I for the proportion of the male passengers with more than one bag

$$n_{male}=690,\ x_{male}=280 \Rightarrow \overline{p}=\frac{x_{male}}{n_{male}}=\frac{280}{690}=0.405797$$

$$1-\alpha=0.95 \Rightarrow \alpha=0.05 \Rightarrow Z_{\frac{0.05}{2}}=1.96$$

$0.4058 \pm 1.96(\sqrt{[(0.4058)(1-0.4058]/690})$ = [0.3692 , 0.4424].

Based on 95% confidence, the proportion of males that would be affected is between 0.3692 and 0.4424. So less than half would be affected. However, this might imply that men may be more affected than the population overall.

d. 15% say YES, C.I is 95% error = 0.02, find the sample size

$$1-\alpha=0.95 \Rightarrow \alpha=0.05 \Rightarrow Z_{\frac{0.05}{2}}=1.96$$

$$p=0.15 \quad and \quad e=\pm 0.02$$

$$n=\frac{z_{\frac{\alpha}{2}}^{2}\,p(1-p)}{e^2} = \frac{1.96^2\left(0.15\right)\left(1-0.15\right)}{\left(0.02\right)^2}$$ = 1,224.51 = 1,225.