Business Statistics
0676-201

College of Business Administration
Quantitative Methods Department

Chapter 1:

Estimation and Confidence Intervals

Goals

- Define a point estimate.
- Define level of confidence.
- Construct a confidence interval for the population mean when the population standard deviation is known.
- Construct a confidence interval for a population mean when the population standard deviation is unknow
- Construct a confidence interval for a population proportion.
- Determine the sample size for attribute and variable sampling

1. Concepts of Estimation

Point and Interval Estimators

Point Estimator

A point Estimator is the statistic, computed from sample information, which is used to estimate the value of an unknown parameter using a single value or point.

Remark

The sample mean, $\bar{x}$, is not the only point estimate of a population parameter. For example, $\hat{p}$, a sample proportion, is a point estimate of p, the population proportion; and s, the sample standard deviation, is a point estimate of σ, the population standard deviation.

Example

Recent financial initiatives in Saudi Arabia emphasize the importance of savings and investment for improving overall financial health and well-being, particularly in securing long-term financial stability. The director of finance at Saudi Aramco, a leading energy company, wants to estimate the average amount of savings employees have accumulated over the past year. A sample of 70 employees reveals that the mean savings amount last year was SAR 25,000. What is the point estimate of the unknown population mean?

Solution:

The sample mean = SAR 25,000

The point estimate of the unknown population mean is the sample mean (SAR 25,000).

Interval Estimator or Confidence Interval Estimator

An **interval estimator is a** range of values constructed from sample data so that the population parameter is likely to occur within that range at a specified probability. The specified probability is called the **level of confidence**

In general the level of confidence is denoted by **1-α**. The most used levels are are 90%,95%, 99%...

The level of significance denoted by α is the probability that the population parameter is not in that interval estimator, for example 10%,%5,1%

Example 1

In assessing the financial risks associated with a new construction project in Saudi Arabia, we estimate that the average cost overrun is SAR 150,000. The range of this estimate might be from SAR 140,000 to SAR 160,000. We can describe how confident we are that the actual cost overrun will fall within this interval. For instance, we might say that we are 90% confident that the cost overrun for the project will be between SAR 140,000 and SAR 160,000.

Remark: Factors Affecting Confidence Interval Estimates

The factors that determine the width of a confidence interval are:

1. The sample size, *n*.

2. The variability in the population, usually σ estimated by *s*.

3. The desired level of confidence.

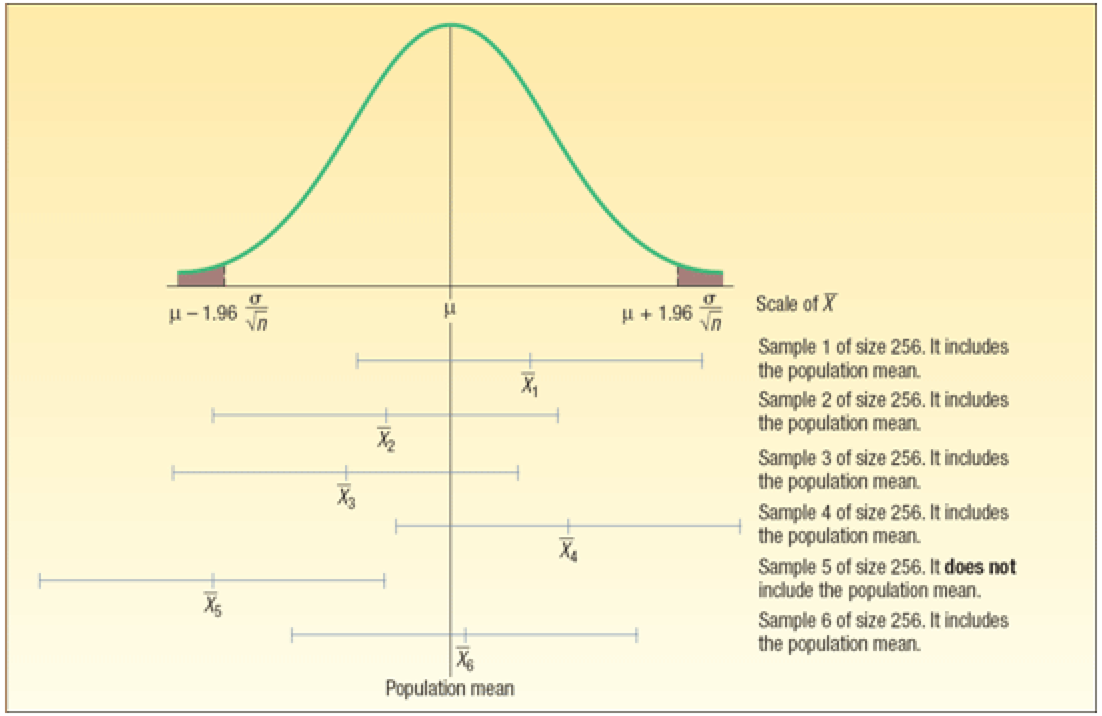
Concepts of Estimation

Point and Interval Estimators

Interval Estimates - Interpretation

For a 95% confidence interval about 95% of the similarly constructed intervals will contain the parameter being estimated.

Example: 95% of the sample means for a specified sample size will lie within 1.96 standard deviations of the hypothesized population



2. Estimating the Population Mean

We represent the Confidence Interval of the estimator as

$$\mu^* = \bar{X} \pm Me,$$

Me is called the margin of error. Here the e is the standard error. It is depending on the size of the sample, and the standard deviation is known or unknown and M is depending on the distribution. More precisely, we have the following summarizing table

2. Estimating the Population Mean

Sample size	*Population Variance*	*Distribution used*	*Standard Value*	*Standard error*
Large $n \geq 30$	***known***	***Normal***	$Z_{\alpha/2}$	$\frac{\sigma}{\sqrt{n}}$
	unknown	***Normal***	$Z_{\alpha/2}$	$\frac{s}{\sqrt{n}}$
Small $n < 30$	***known***	***Normal***	$Z_{\alpha/2}$	$\frac{\sigma^2}{n}\left(\frac{N-1}{n-1}\right)$
	unknown	***Student (t)***	$t_{n-1,\alpha/2}$	$\frac{s}{\sqrt{n}}$

$Z_{\alpha/2}$ is the standard value for which the area at the right is equal to $\alpha/2$ for the Standard Normal Distribution

$t_{n-1,\alpha/2}$ is the standard value for which the area at the right is equal $\alpha/2$ to for the t Distribution of degree of freedom n-1

2. Estimating the Population Mean

Confidence interval for the population mean $\boldsymbol{\mu}$

Sample size	*Population Variance*	*Confidence Interval*
Large $n \geq 30$	***known***	$\boldsymbol{\mu} = \bar{\boldsymbol{X}} \pm \boldsymbol{Z}_{\alpha/2} \frac{\sigma}{\sqrt{\boldsymbol{n}}}$
	unknown	$\boldsymbol{\mu} = \bar{\boldsymbol{X}} \pm \boldsymbol{Z}_{\alpha/2} \frac{\boldsymbol{s}}{\sqrt{\boldsymbol{n}}}$
Small $n < 30$	***known***	$\boldsymbol{\mu} = \bar{\boldsymbol{X}} \pm \boldsymbol{Z}_{\alpha/2} \frac{\sigma^2}{n}\left(\frac{N-1}{n-1}\right)$
	unknown	$\boldsymbol{\mu} = \bar{\boldsymbol{X}} \pm \boldsymbol{t}_{\boldsymbol{n-1},\alpha/2} \frac{\boldsymbol{s}}{\sqrt{\boldsymbol{n}}}$

2. Estimating the Population Mean

Four Commonly Used Confidence Levels and $z_{\alpha/2}$

$1-\alpha$	α	$\alpha/2$	$z_{\alpha/2}$
0.90	0.1	0.05	1.645
0.95	0.05	0.025	1.96
0.98	0.02	0.01	2.33
0.99	0.01	0.005	2.575

2. Estimating the Population Mean

Example 2 :

An economic analyst wishes to investigate the average household expenditure on utilities in a certain region of Saudi Arabia. A sample of 10 households over a year revealed a sample mean monthly expenditure of SAR 1,200 with a standard deviation of SAR 300. Construct a 95 percent confidence interval for the population mean. Would it be reasonable for the analyst to conclude that the population mean monthly expenditure on utilities is SAR 1,150?

Solution: (Example 2)

Given in the problem:

n=10, $\overline{X} = 1200$, s=300

The level of confidence is $1 - \alpha = 95\% = 0.95$

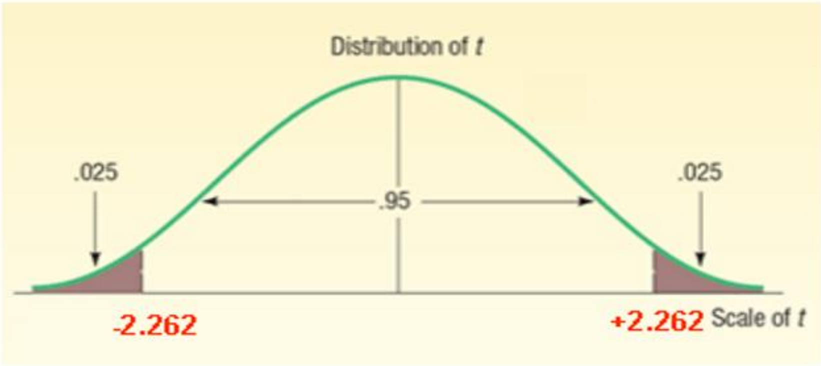
Since the variance is unknown, then we use the t- distribution, and we have

$$\boldsymbol{\mu} = \overline{\boldsymbol{X}} \pm \boldsymbol{t}_{\boldsymbol{n-1},\alpha/2} \frac{\boldsymbol{s}}{\sqrt{\boldsymbol{n}}} = 1200 \pm \boldsymbol{t}_{\boldsymbol{9,0,025}} \frac{300}{\sqrt{\boldsymbol{10}}}$$

$$\boldsymbol{\mu} = 1200 \pm 214.592 = (985.408,1414.592)$$

We conclude that the economic analyst can be reasonably sure (95% confident) that that the population mean monthly expenditure on utilities is between 985.408 and 1414.592 SAR

	Confidence Intervals				
	80%	90%	95%	98%	99%
	Level of Significance for One-Tailed Test				
df	0.100	0.050	0.025	0.010	0.005
	Level of Significance for Two-Tailed Test				
	0.20	0.10	0.05	0.02	0.01
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	$\boxed{2.262}$	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169



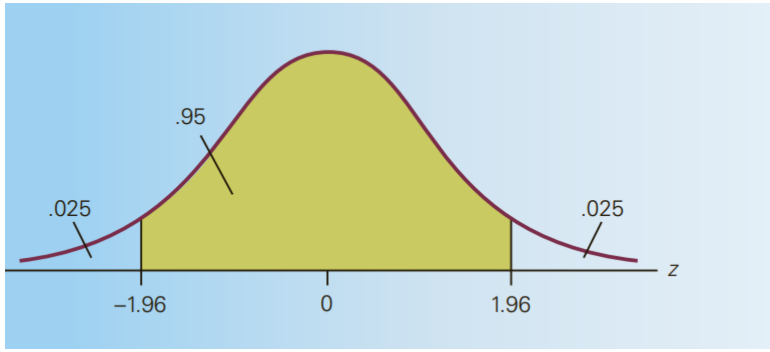
2. Estimating the Population Mean

Example 3:

Al-Shawarma Express is a franchise fast-food restaurant chain located in the Eastern Province of Saudi Arabia, specializing in shawarma wraps, falafel sandwiches, and grilled chicken plates. Fresh juices and French fries are also available. The Marketing Department of Al-Shawarma Express LLC reports that the distribution of daily sales for their restaurants follows the normal distribution and that the population standard deviation is SAR11,250. A sample of 40 franchises showed the mean daily sales to be SAR 75,000. Develop a 95% confidence interval for the population mean of daily sales. Interpret the confidence interval.

2. Estimating the Population Mean

Solution: (Example 3)



Given in the problem:

n=40, $\overline{X} = 75000$, σ= 11250

The level of confidence is $1 - \alpha = 95\% = 0.95$

Since the variance is known, then we use the Normal distribution, and we have

$$\boldsymbol{\mu = \overline{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}} = 75000 \pm \boldsymbol{Z_{0.025} \frac{11250}{\sqrt{40}}}$$

$$\boldsymbol{Z_{0.025} = 1.96, so}$$

$$\boldsymbol{\mu} = 75{,}000 \pm 3486.41 = (71513.59, 78486.41)$$

We can be 95% confident that the true population mean of daily sales for Al-Shawarma Express franchises in the Eastern Province of Saudi Arabia falls between SAR 71,513.59 and SAR 78,486.41.

2. Estimating the Population Mean

Example 4: (Excel workbook)

The manager of the Inlet Square Mall, near Ft. Myers, Florida, wants to estimate the mean amount spent per shopping visit by customers. A sample of 20 customers reveals the following amounts spent.

$ 48.16	$ 42.22	$ 46.82	$ 51.45	$ 23.78	$ 41.86	$ 54.86
37.92	52.64	48.59	50.82	46.94	61.83	61.69
49.17	61.46	51.35	52.68	58.84	43.88	

Determine a 95% confidence interval. Interpret the result. Would it be reasonable to conclude that the population mean is $50? What about $60?

2. Estimating the Population Mean

Solution: (Example 4)

Instructions:

1. Open the sheet Example4 from the Excel Workbook Chapter1, calculate the sample mean and the sample standard deviation
2. Use the command CONFIDENCE.T to find the margin of error then deduce the lower and upper limit of the confidence interval estimator

t-estimate of a Mean			
Sample Mean	49.348	Confidence interval estimate	
Sample Standard deviation	9.012111	Margin of Error	=CONFIDENCE.T(1-J11,J9,J10)
Sample size	20	Lower Confidence limit	**CONFIDENCE.T**(alpha, standard_dev, size)
Confidence level	95%	Upper Confidence limit	53.5658

2. Estimating the Population Mean

Solution: (Example 4)

The manager of Inlet Square wondered whether the population mean could have been $50 or $60.

t-estimate of a Mean			
Sample Mean	49.348	Confidence interval estimate	
Sample Standard deviation	9.012111	Margin of Error	4.217798
Sample size	20	Lower Confidence limit	45.1302
Confidence level	95%	Upper Confidence limit	53.5658

The confidence interval is (45.13,53.57)

The value of $50 is within the confidence interval. It is reasonable that the population mean could be $50.

The value of $60 is not in the confidence interval. Hence, we conclude that the population mean is unlikely to be $60.

2. Estimating the Population Mean

Example 5: (Excel workbook)

Because of different sales ability, experience, and devotion, the incomes of real estate agents vary considerably. Suppose that in a large city the annual income is normally distributed with a standard deviation of $15,000. A random sample of 60 real estate agents was asked to report their annual income (in $1,000). The responses are saved in sheet Example 5. Determine the 99% confidence interval estimate of the mean annual income of all real estate agents in the city.

2. Estimating the Population Mean

Solution: (Example 5)

Instructions:

1. Open the sheet Example5 from the Excel Workbook Chapter1, calculate the sample mean
2. Use the command CONFIDENCE.NORM to find the margin of error then deduce the lower and upper limit of the confidence interval estimator

Z-estimate of a Mean			
Sample Mean	76.15	Confidence interval estimate	
Population Standard deviation	15	Margin of Error	=CONFIDENCE.NORM(1-E7,E5,E6)
Sample size	60	Lower Confidence limit	71.16193
Confidence level	99%	Upper Confidence limit	81.13807

CONFIDENCE.NORM(alpha, standard_dev, size)

2. Estimating the Population Mean

Solution: (Example 5)

Z-estimate of a Mean			
Sample Mean	76.15	Confidence interval estimate	
Population Standard deviation	15	Margin of Error	4.988072
Sample size	60	Lower Confidence limit	71.16193
Confidence level	99%	Upper Confidence limit	81.13807

The confidence interval of the mean annual income of all real estate agents in the city is (71.16,81.14)

2. Estimating the Population Mean

Exercise 1:

An insurance company in Saudi Arabia conducted a study to determine the average claim amount for car accidents in Riyadh. A random sample of 50 claim amounts was collected and the mean was computed to be 26.75 minutes. An analysis of these claim amounts reveals that they are normally distributed with a standard deviation of SAR 8,000.

Estimate with 95% confidence the mean claim amount for all car accidents in Riyadh.

What do your results tell you about the financial risk associated with car accident claims?

3. Estimating the Population proportion

PROPORTION

The fraction, ratio, or percent indicating the part of the sample or the population having a particular trait of interest

Example: a recent survey at a company in Saudi Arabia indicated that 92 out of 100 employees were satisfied with the company's new flexible working hours policy during Ramadan. The sample proportion is 92/100, or 0.92, or 92%.

3. Estimating the Population proportion

Point estimator of Population proportion

If we let $\hat{p}$ represent the sample proportion, x the number of "successes," and n the number of items sampled, we can determine a sample proportion as follows

$$\hat{p} = \frac{x}{n}$$

3. Estimating the Population proportion

Confidence Interval for Population proportion

If we denote p the Population proportion and $\hat{p}$ the sample proportion, and n the number of items sampled, then the Confidence interval is given by

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

Remark: The values np and n(1 − p) should both be greater than or equal to 5. So that we can employ the standard normal distribution, that is, z, to complete a confidence interval.

3. Estimating the Population proportion

Confidence Interval for Population proportion

Example 6:

A market research consultant was hired to estimate the proportion of households in Riyadh that associate the brand name of a popular oud-scented air freshener with its distinctive bottle design. The consultant randomly selected 1,400 households in Riyadh. From the sample, 420 were able to identify the brand by name based only on the shape and design of the bottle. Develop a 99% confidence interval for the population proportion.

3. Estimating the Population proportion

Confidence Interval for Population proportion

Solution: (Example 6)

n=1400, x=420,

Confidence level=99%, so α =1% and $z_{\frac{\alpha}{2}}$ =2.575

$$\hat{p} = \frac{x}{n} = \frac{420}{1400} = 0.3$$

$$p = \hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.3 \pm 0.012 = (0.288, 0.312)$$

3. Estimating the Population proportion

Confidence Interval for Population proportion

Exercise 2:

Al-Madina Printing Co. purchases ceramic mugs and imprints them with calligraphy and designs for Ramadan, Eid celebrations, national day, and other special occasions. Ahmed Al-Harbi, the owner, received a large shipment this morning from a new supplier in China. To ensure the quality of the shipment, he selected a random sample of 300 mugs and inspected them for defects. He found 15 to be defective.

a. What is the estimated proportion of defective in the population?
b. Develop a 95% confidence interval for the proportion defective.
c. Ahmed has an agreement with his supplier that if 10% or more of the mugs are defective, he can return the order. Should he return this lot? Explain your decision.

4. Determining the Sample Size

Determining the Sample Size to Estimate the Population Mean

The determination of the sample size n to Estimate the Population Mean is based on three variables:

1. The margin of error the researcher will tolerate.
2. The level of confidence desired, for example, 95%.
3. The variation or dispersion of the population being studied.

4. Determining the Sample Size

Determining the Sample Size to Estimate the Population Mean

If the variance is known,

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

If the variance is unknown, then

$$n = \frac{(Z_{\alpha/2})^2 s^2}{E^2}$$

where:

n is the size of the sample.

$Z_{\alpha/2}$ is the standard normal z-value corresponding to the desired level of confidence.

E is the maximum allowable error.

4. Determining the Sample Size

Determining the Sample Size to Estimate the Population Mean

Example 7 :

A student in public administration wants to estimate the mean monthly allowance of Shura Council members in Saudi Arabia. He can tolerate a margin of error of 375 SAR in estimating the mean. He would also prefer to report the interval estimate with a 95% level of confidence. The student found a report by the Ministry of Human Resources and Social Development that reported a standard deviation of 3,750 SAR. What is the required sample size?

Determining the Sample Size to Estimate the Population Mean

Solution: (Example 7)

n : the size of the sample.

Confidence level=95%, so $\alpha = 5\%$ and $z_{\frac{\alpha}{2}}$ =1.96

E =SAR 375, σ=SAR 3,750.

$$n = \frac{(Z_{\alpha/2})^2\sigma^2}{E^2} = \left(\frac{1.96 \times 3750}{375}\right)^2 = 384.16$$

Since we can't have a fractional sample size, we round up to the nearest whole number. Therefore, the required sample size is 385 Shura Council members.

4. Determining the Sample Size

Determining the Sample Size to Estimate the Population Mean

Exercise 3:

An insurance company in Saudi Arabia wants to estimate the average claim amount for car accidents in a specific region. Claim amounts range from SAR 1,000 to SAR 100,000. The estimate of the population mean claim amount should be within plus or minus SAR 500 of the population mean. Based on historical data, the population standard deviation is SAR 2,500. Using a 95% level of confidence, how many claim records need to be selected?

4. Determining the Sample Size

Determining the Sample Size to Estimate the Population Proportion

The determination of the sample size n to Estimate the Population Proportion is based on three variables:

1. The margin of error the researcher will tolerate.
2. The level of confidence desired, for example, 95%.
3. The proportion of the population being studied.

4. Determining the Sample Size

Determining the Sample Size to Estimate the Population Proportion

$$n = \left(\frac{Z_{\alpha/2}}{E}\right)^2 p(1-p)$$

where:

n is the size of the sample.

$Z_{\alpha/2}$ is the standard normal z-value corresponding to the desired level of confidence.

p is the population proportion.

E is the maximum allowable error.

Remark: If a reliable value cannot be determined with a pilot study or found in a comparable study, then a value of .50 can be used for p.

4. Determining the Sample Size

Determining the Sample Size to Estimate the Population Proportion

Example 8 :

An economist wants to estimate the proportion of businesses in a specific industry that are currently experiencing a decline in sales. The economist wants to estimate the population proportion with a margin of error of 0.10, prefers a level of confidence of 90%, and has no estimate for the population proportion. What is the required sample size?

Determining the Sample Size to Estimate the Population Proportion

Solution: (Example 8)

n : the size of the sample.

Confidence level=90%, so $\alpha = 10\%$ and $z_{\frac{\alpha}{2}}$ =1.645

E =0.1, p=0.5.

$$n = \left(\frac{Z_{\alpha/2}}{E}\right)^2 p(1-p) = (\frac{1.645}{0.1})^2 \times 0.5 \times (1-0.5) = 67.65$$

Since we can't have a fractional sample size, we round up to the nearest whole number. Therefore, the required sample size is 68 businesses.

4. Determining the Sample Size

Determining the Sample Size to Estimate the Population Proportion

Exercise 4:

Previous studies indicate that 30% of Saudi Arabian consumers are satisfied with the quality of locally produced goods. The Saudi Standards, Metrology, and Quality Control Organization (SASO) wants to update this percentage. How many consumers should be randomly selected to estimate the population proportion with a 90% confidence level and a 1% margin of error?

End of Chapter 1

Thank you