Chapter 6

Estimating the value of a Parameter

Using Confidence Intervals

CLO6	**Estimate the value of a parameter using confidence intervals.**

1. Introduction

- In this chapter we will focus on the following points:

  ✓ Compute a point estimate of the population mean.

  ✓ Construct and interpret a confidence interval for a population mean μ, assuming that the standard deviation σ of the population is known.

  ✓ Explain the role of margin of error in constructing a confidence interval.

  ✓ Determine the sample size necessary for estimating the population mean with a specified margin of error.

2. Confidence intervals for a population mean μ when σ is known:

- **Confidence interval:** confidence interval for an unknown parameter consists of an interval of numbers.

- **Level of confidence:** represents the expected proportion of intervals that will contain the parameter if a large number of different samples is obtained. It is denoted By:

$$(1-\alpha).100\% \qquad (1)$$

- For example, a 95% level of confidence $(\alpha = 0.05)$ implies that if 100 different confidence intervals are constructed, each based on different sample from the same population, then we will expect 95 of the intervals to include the parameter and 5 to not include the parameter.

- Confidence interval estimates for the population mean μ are of the form:

Point estimate $\pm$ margin error

- **The margin of error** of a confidence interval estimate of a parameter is a measure of how accurate the point estimate is and depends on 3 factors:
    - **Level of confidence**: increases as the margin of error increases.
    - **Sample size**: As sample size increases, the margin of error decreases.
    - **Standard deviation of the population**: The more spread the population, the wider will be the level of confidence.

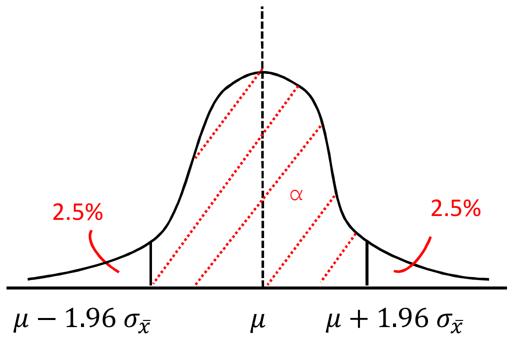
2.1 Construction of a confidence interval for a population mean

- ✓ The distribution of the sample mean will be normal if the population is normal, or approximately normal if the population is not normal, but the sample size is large $(\boldsymbol{n} \geq \mathbf{30})$.
- ✓ The mean of the distribution of sample mean has mean $\boldsymbol{\mu}_{\overline{X}} = \boldsymbol{\mu}$ and standard deviation $\boldsymbol{\sigma}_{\overline{X}} = \frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{n}}}$.
- ✓ The sample mean is a point estimate of the population mean, so we expect that it is close the mean of the population. But the question is how much it is close?
- ✓ Because the sample mean is normally distributed, 95% of all sample means lie within 1.96 standard deviations:

$$\boldsymbol{\mu - 1.96\,\sigma_{\bar{X}} < \bar{X} < \mu + 1.96\,\sigma_{\bar{X}}}$$

$$\boldsymbol{\bar{X} - 1.96\,\sigma_{\bar{X}} < \mu < \bar{X} + 1.96\sigma_{\bar{X}}}$$

$$\boldsymbol{P\left(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\,\frac{\sigma}{\sqrt{n}}\right) = 0.95}$$



- This means that 95% of all sample means will results in confidence interval estimates that contain the population mean.

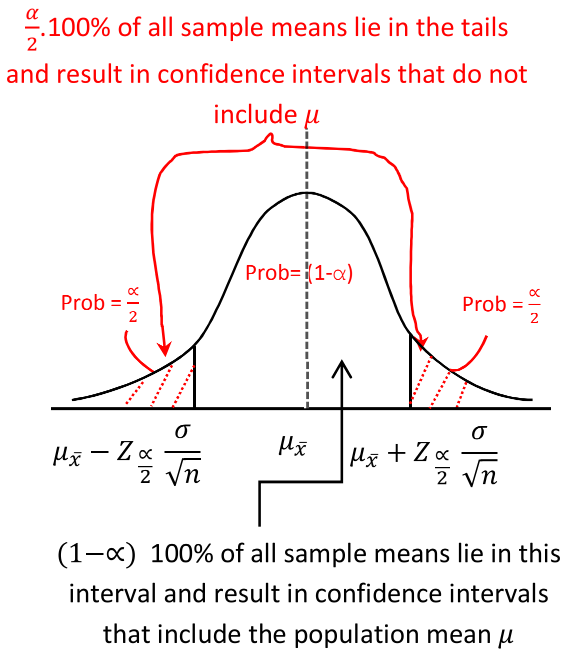
- It is common to write the 95% confidence interval as:

$$\boldsymbol{\bar{X} \pm 1.96\sigma_{\bar{X}}}$$

Point estimate ± margin of error.

- **Important:** Any sample mean $\boldsymbol{\bar{X}} \in \left[\boldsymbol{\mu - \frac{\sigma}{\sqrt{n}}, \mu + \frac{\sigma}{\sqrt{n}}}\right]$ will result in confidence interval that contains μ, and any $\boldsymbol{\bar{X} > \mu + \frac{\sigma}{\sqrt{n}}}$ $\boldsymbol{or}$ $\boldsymbol{\bar{X} < \mu - \frac{\sigma}{\sqrt{n}}}$ will result in confidence interval that does not contain μ.

- **Constructing a (1-α)100% confidence interval for μ,α known:**



- ✓ Suppose a simple random sample of size in is taken from a population with unknown mean μ and known standard deviation σ. Then (1-α)100% confidence interval for μ is given by:

  Lower bound: $\overline{X} - Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}$

  Upper bound: $\overline{X} + Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}$

  Where $Z_{\alpha/2}$ is the critical Z-value.

- ✓ Note that n must be large (n≥ 30) or the population is normally distributed.

Level of confidence	Area in each tail, $\frac{\propto}{2}$	Critical value, $Z_{\propto/2}$
90	0.05	1.645
95	0.025	1.96
99	0.005	2.575

Example 1:

A road has a speed limit of 45km/h. the residents near the road are concerned that the speed of cars is excessive. Based on the simple random sample of the speed of 12 cars, find:

(a) The estimate of the population mean speed of cars.

(b) Construct 90% confidence interval about the population mean; assume the standard deviation speed of cars is 8km/h.

57.4	56.1	70.3	65.6
44.2	58.6	66.1	57.3
62.2	60.4	64.5	52.7

(c) Interpret the results

Solution:

$\bar{X} = \frac{57.4+56.1+\cdots+52.7}{12} = 59.62$ km/h.

(b) we assume the distribution of the speed of cars to follow normal distribution, since n=12 is less than 30.

Since we are constructing 90% confidence interval, ∝=0.10. Therefore:

$Z_{\propto/2} = Z_{0.05} = 1.645$ (from table)

Lower bound: $\bar{X} - Z_{\propto/2} . \frac{\sigma}{\sqrt{n}} = 59.62 - 1.645 \frac{8}{\sqrt{12}} = 55.82$

Upper bound: $\bar{X} + Z_{\propto/2} . \frac{\sigma}{\sqrt{n}} = 63.42$

$$55.82 < \mu < 63.42$$

We are 90% confident that the mean speed of all cars is between 55.82 and 63.42. Therefore, the resident have the right to be concerned since μ is 90% well above the speed limit of 45.

- The Margin of Error: E in a (1-∝) 100% confidence interval with σ is known is given by:

$$E = Z_{\propto/2} . \frac{\sigma}{\sqrt{n}}$$

Note that E is valid for a sample drawn from a normal distribution or n≥30.

- **Margin of Error:**

The margin of error E is an (1-α)100% confidence interval in which in which σ is known is given by:

$$E = Z_{\frac{\alpha}{2}} . \frac{\sigma}{\sqrt{n}}$$

- **Important:**

The margin of error expression is valid only for a sample drawn from a normal population or the sample size must be greater than or equal to 30.

Example 2:

In example 1

(a) Determine the effect on margin of error when increasing the level of confidence from 90% to 99%.

(b) Determine the effect of increasing the sample size n from 12 to 48 on the margin of error. Keep level of confidence 90%.

Solution:

(a) with 99% level of confidence, ∝=0.01:

$\boldsymbol{Z_{0.005} = 2.575}$ (from table)

$$\boldsymbol{E = Z_{\propto/2}\frac{\sigma}{\sqrt{n}} = 2.575\frac{8}{\sqrt{12}} = 5.95}$$

Note that E has increased from 3.8 to 5.95. If we want to be more confident that an interval contains μ, we need to increase the width of the interval.

(b) $E = Z_{0.05}\frac{\sigma}{\sqrt{n}} = 1.645\frac{8}{\sqrt{48}} = 1.90$

The margin of error has decreased from 3.8 to 1.9 as n increased from 12 to 48.

- **Determining the sample size *n*:**

The sample size required to estimate the population mean μ with (1-α)100% level of confidence and specified margin of error E is given by:

$$n = \left(\frac{Z_{\propto/2}\,.\,\sigma}{E}\right)^2$$

where n is rounded up to the nearest integer number.

Example 3:

In example 9.1, how large a sample is required to estimate the mean speed with a margin of error E equal to 2km/h with 90% confidence?

Solution

$Z_{\propto/2} = Z_{0.05} = 1.645, \sigma = 8\ , E = 2$:

$$n = \left(\frac{Z_{\propto/2}\,.\,\sigma}{E}\right)^2 = 43.29$$

Therefore, n = 44.

---

3. Confidence intervals for a population mean μ when σ is unknown

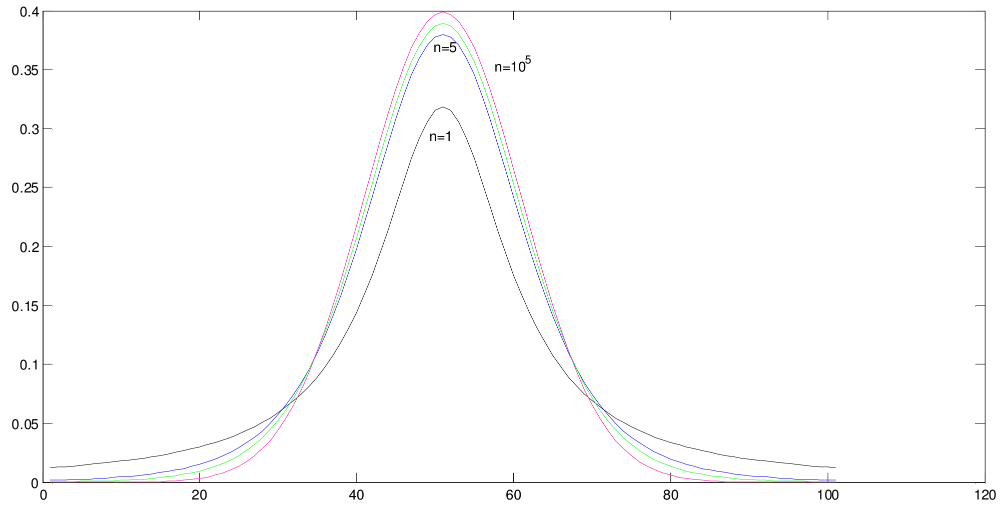
3.1 Theorem:

Suppose a simple random sample of size n is taken from a population. If the population follows a normal distributions, then the distribution of:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

Follows student's t-distribution with *n*-1 degree of freedom, where $\bar{x}$ is the sample mean and *s* is the sample standard deviation.

- **Properties of t- distribution:**
    - ✓ Function of the sample size n or the degrees of freedom.
    - ✓ Centered at 0 and symmetric about 0.
    - ✓ As n increases, t- distribution gets closes to the normal.



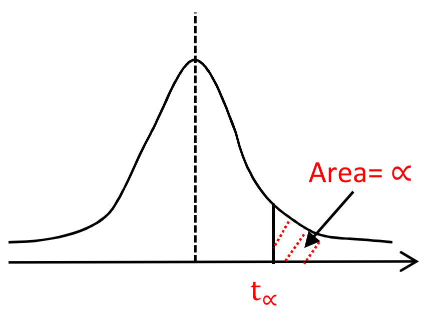
Example 4

Find the t-value such that the area under the t-distribution to the right of t-value is 0.1, assuming 15 degrees of freedom. That is find $t_{0.1}$ with 15 degrees of freedom.

Solution:

The area under the t-distribution to the right of this $t_{0.1}$ equal 0.1. From t-table:

$$t_{0.1} = 1{,}341$$



3.2 Constructing (1-∝) 100% confidence interval about μ, σ unknown:

- Suppose a simple random sample of size n is taken from a population with μ and σ are unknown.
- A (1-σ)100% confidence interval for μ is given by:

  Lower bound: $\bar{X} - t_{\propto/2}\frac{s}{\sqrt{n}}$

  Upper bound: $\bar{X} + t_{\propto/2}\frac{s}{\sqrt{n}}$

  Where $t_{\propto/2}$ is computed with n-1 degrees of freedom.

Example 5:

The data represents a simple random sample of student's exam. Construct a 95% confidence interval for the mean.

64	33.4	45.8	56	51.5	29.2	63.7

We have: $\boldsymbol{t_{0.025}}$ with $\boldsymbol{6\ df} = \boldsymbol{2.447}$ (from table)

Solution:

$\bar{X} = \frac{64+33.4+\cdots\ldots\ldots+63.7}{7} = 49.09$

$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{6}[(64-\bar{x})^2 + (33.4-\bar{x})^2 + \cdots \ldots.(65-\bar{x})^2]$

S=$\sqrt{S^2} = 13{,}8$

Now since 95% confidence interval, ∝=0.05, degrees of freedom $= \mathbf{7-1=6}$:

$\boldsymbol{t_{0.025}}$ with $\boldsymbol{6\ df} = \boldsymbol{2.447}$ (from table)

Lower bound: $\boldsymbol{\bar{X} - t_{\propto/2}\frac{s}{\sqrt{n}} = 49.09 - 2.447\frac{13.8}{\sqrt{7}} = 36.33}$

Upper bound: $\boldsymbol{\bar{X} + t_{\propto/2}\frac{s}{\sqrt{n}} = 61.85}$

So we are 95% confident that the mean exam mark is between 36.33 and 61.85.

t-Distribution

Area in Right Tail

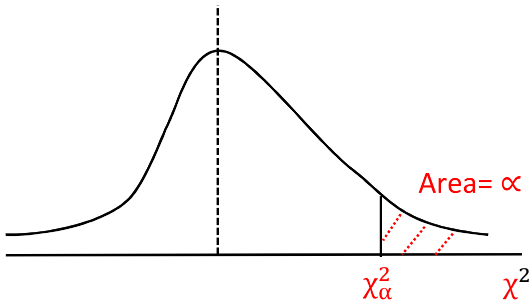
df	0.25	0.20	0.15	0.10	0.05	0.025	0.02	0.01	0.005
1	1.000	1.376	1.963	3.078	6.314	12.706	15.894	31.821	63.657
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797

4. Confidence Intervals for a Population Standard Deviation

- There are some situations in which we are interested to find the confidence intervals for the population standard deviation.
- If a simple random sample of size *n* obtained from a normally distributed population with mean μ and standard deviation σ, then:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

has a **chi–square distribution** with n-1 degree of freedom .



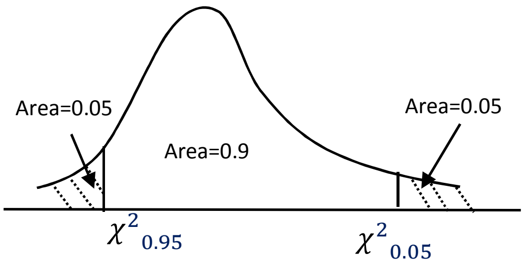
–

- Characteristics of the chi–square distribution :
  - It is not symmetric.
  - Depends on the sample size or degree freedom.
  - The value of χ^2 is always nonnegative.

Example 6:

1. Find the critical values that separate the middle 90% of the chi- square distribution assuming 15 degrees of freedom ($\frac{\alpha}{2}$ and $1-\frac{\alpha}{2}$).
2. Similarly find the critical values for 95%. (use table below)

Solution:



Chi-Square (χ^2) Distribution

Degrees of Freedom	Area to the Right of Critical Value							
	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01
1	—	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345
4	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277
5	0.554	0.831	1.145	1.610	9.236	11.071	12.833	15.086
6	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812
7	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475
8	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090
9	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666
10	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209
11	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725
12	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217
13	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688
14	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141
15	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578
16	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000
17	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409
18	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805
19	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191
20	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566
21	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932
22	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289
23	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638
24	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980

4.1 Constructing (1-α)100% confidence interval about σ^2:

- If a simple random sample of size n is taken from a normal population with mean μ and standard deviation σ, then a (1-α)100% confidence interval about σ^2 is given by:

Lower bound: $\frac{(n-1)s^2}{\chi^2{}_{\alpha/2}}$

Upper bound: $\frac{(n-1)s^2}{\chi^2{}_{1-\alpha/2}}$

Example 7:

Below are the prices of 12 randomly selected 3-year-old Corolla. Construct 90% confidence interval for the population variance and standard deviation of the price:

41.844	41.500	39.995	36.995	40.990	37.995
41.995	38.900	42.995	36.995	43,995	35.950

Assume the population has a normal distribution.

Solution:

$$s^2 = \frac{1}{n-1}\sum_{i-1}^{12}(x_i - \bar{x})^2 = (2,615.9)^2$$

$s = 2,615.9$

Now, for 90% confidence, ∝=0.1. Using the x^2 table with 11df:

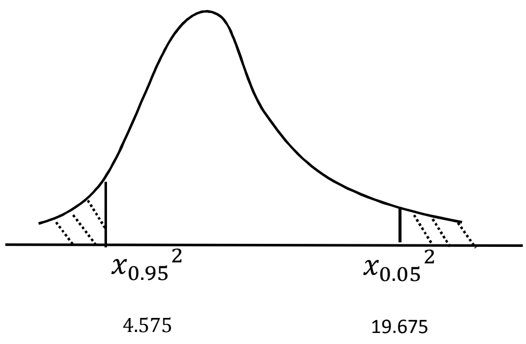
$x^2{}_{0.05} = 19.675$

$x^2{}_{0.95} = 4.575$

σ^2Lower bound:$= \frac{(n-1)s^2}{x^2{}_{0.05}} = \frac{11(2615.9)^2}{19.675} =$ $3,823,705.52$

σ^2Upper bound:$= \frac{(n-1)s^2}{x^2{}_{0.95}} = \frac{11(2615.9)^2}{4.575} =$ $16,444,023.19$

$$3,823,705.52 < \sigma^2 < 16,444,023.19$$



$\mathbf{1955 < \boldsymbol{\sigma} < 4055}$ **with 90% confidence.**