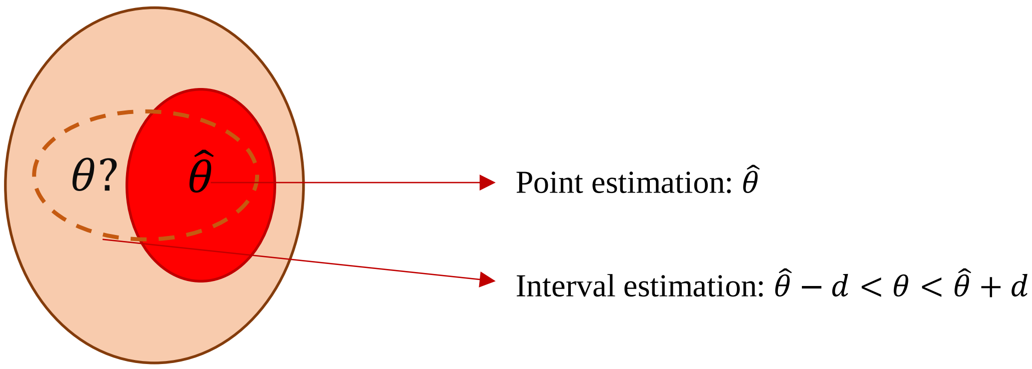
PARAMETER ESTIMATION

INTRODUCTION

Parameter estimation is the process of estimating the value of a parameter based on information obtained from a random sample. In inferential statistics, there are two types of parameter estimates, namely point estimation and interval estimation. Point estimation means obtaining a statistical value for estimating the parameter. Suppose that the mean of a random sample $(\bar{x})$ is used to estimate the population mean (μ). The interval estimate represents the range of values around the point estimate at which the parameter is estimated. The description of point estimation and interval estimation is as follows.

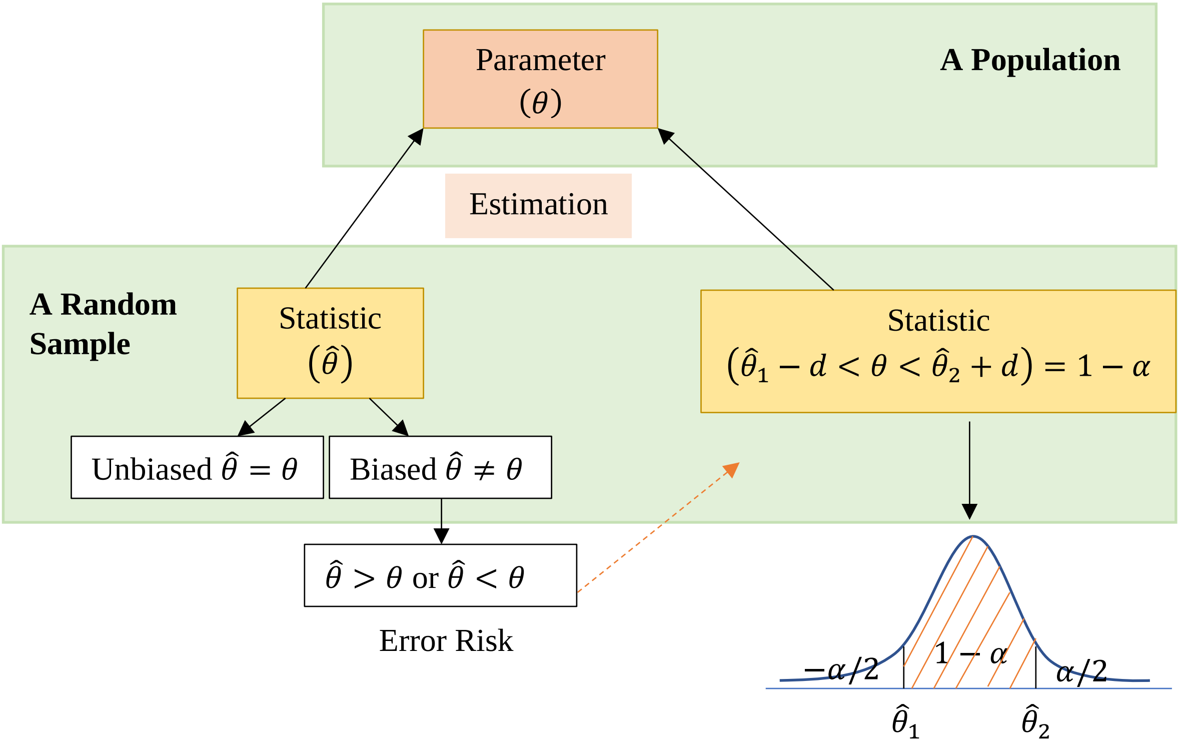


Suppose there are survey results as follows:

a. the mean of learning time for Tel U students in the even semester 2019/2020 is 14 hours in 1 week
b. the mean of Tel U student expenses for one month's study needs in the 2019/2020 even semester is between Rp. 1,000,000, - to Rp. 2,000,000, -

The two examples above are statistics from a sample with a population is all Tel U students. Example “a” is an example of point estimation, where 14 hours is the mean learning time of the sample used to describe the learning time conditions for all Tel U students. Example “b” is an example of interval estimation, where the average expenditure of all Tel U students is around Rp. 1,000,000, - to Rp. 2,000,000, -.

THE CONCEPT OF PARAMETER ESTIMATION



Note: d is the margin of error

1. Point Estimation

Point estimation has indirectly been described in the previous chapter, namely the sampling distribution. For example, as the result of $\mu_{\bar{X}} = \mu$ (in the example $\mu_{\bar{X}} = \mu = 21$). If $\hat{\theta} = \theta$, then the estimation results obtained are unbiased so that the estimator $\hat{\theta}$ can represent the population accurately. If $\hat{\theta} \neq \theta$, then the estimator $\hat{\theta}$ cannot represent the population accurately (biased). Biased in parameter estimation is the error risk in inferential statistics in terms of parameter estimation. To reduce the error risk, the estimate interval can be used.

The point estimation of the mean can be formulated as follows:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

As for the variance and the standard deviation:

Variance:

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

The variance formula above is the same as the variance formula below:

$$s^2 = \frac{n\sum_{i=1}^{n}{x_i}^2 - (\sum_{i=1}^{n}x_i)^2}{n(n-1)}$$

Proof:

Remember the quadratic equation formula: $(a-b)^2 = a^2 - 2ab + b^2$

So:

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

$$s^2 = \frac{\sum_{i=1}^{n}[(x_i)^2 - 2x_i\bar{x} + (\bar{x})^2]}{n-1}$$

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

$$s^2 = \frac{\sum_{i=1}^{n}(x_i)^2 - 2\bar{x}.n\bar{x} + n\bar{x}^2}{n-1}$$

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

$$s^2 = \frac{\sum_{i=1}^{n}(x_i)^2 - n\frac{1}{n^2}(\sum_{i=1}^{n}x_i)^2}{n-1}$$

$$s^2 = \frac{\sum_{i=1}^{n}(x_i)^2 - \frac{1}{n}(\sum_{i=1}^{n}x_i)^2}{n-1}$$

$$s^2 = \frac{\frac{n\sum_{i=1}^{n}(x_i)^2}{n} - \frac{1}{n}(\sum_{i=1}^{n}x_i)^2}{n-1} = \frac{n\sum_{i=1}^{n}(x_i)^2 - (\sum_{i=1}^{n}x_i)^2}{n(n-1)}$$

- $\sum_{i=1}^{n}x_i$ at $2\bar{x}\sum_{i=1}^{n}x_i$ can be converted into $n\bar{x}$ (see mean formula)
- $\sum_{i=1}^{n}\bar{x}^2 = n\bar{x}^2$ ($\bar{x}^2$ is a constant, if you add up n times, then the result is the same as $n\bar{x}^2$)
- $\bar{x}^2 = \left(\sum_{i=1}^{n}\frac{x_i}{n}\right)^2 = \frac{1}{n^2}(\sum_{i=1}^{n}x_i)^2$

The standard deviation:

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

or

$$s = \sqrt{\frac{n\sum_{i=1}^{n}(x_i)^2 - (\sum_{i=1}^{n}x_i)^2}{n(n-1)}}$$

Example:

As many as 50 students in class A, there are 5 students whose entrance examination results are in the top 3. The distribution of values is 100, 95, 98, 100, and 97. From these 5 values, calculate the mean and standard deviation!

Answer:

The Mean:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$= \frac{100 + 95 + 98 + 100 + 97}{5} = \frac{490}{5} = 98$$

Calculate the standard deviation:

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

$$= \frac{(100-98)^2 + (95-98)^2 + (98-98)^2 + (100-98)^2 + (97-98)^2}{5-1}$$

$$= \frac{(2)^2 + (-3)^2 + (0)^2 + (2)^2 + (-1)^2}{4} = \frac{18}{4} = 4.5$$

$$s = \sqrt{s^2} = \sqrt{4.5} = 2.12$$

You can also use the second formula:

$$\sum_{i=1}^{n} (x_i)^2 = (100)^2 + (95)^2 + (98)^2 + (100)^2 + (97)^2 = 48038$$

$$\left(\sum_{i=1}^{n} x_i\right)^2 = (100 + 95 + 98 + 100 + 97)^2 = 490^2 = 240100$$

$$s = \sqrt{\frac{n \sum_{i=1}^{n} (x_i)^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)}} = \sqrt{\frac{5(48038) - 240100}{5(5-1)}} = \sqrt{\frac{90}{20}} = 2.12$$

2. Interval Estimation

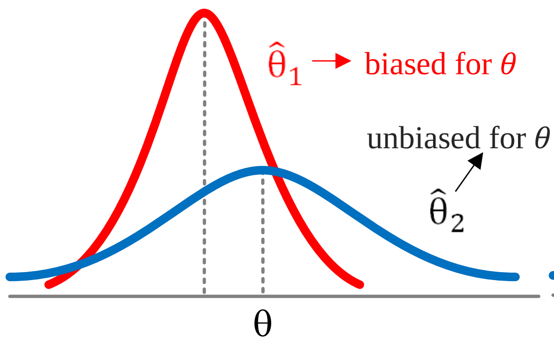
If the parameter estimation is stated in the form of an interval $\left(\hat{\theta}_1 < \theta < \hat{\theta}_2\right)$ or it is called a confidence interval (CI), then the amount of distance between the intervals is used as the accuracy of the estimation results. $(1 - \alpha)$ is the confidence value of the area/interval of the parameter estimation results, usually called the confidence level and the error tolerance (level of significance) of the CI is α. This α describes the level of error in estimation.

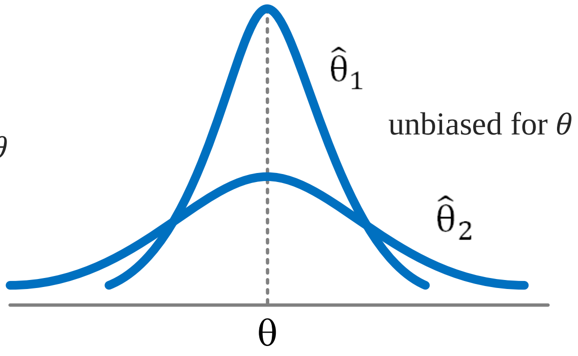
ESTIMATOR PROPERTIES

The statistic used to estimate the parameter is called an estimator so that $\hat{\theta}$ is the estimator for θ. A good estimator has the following properties.

a. Unbiased estimation

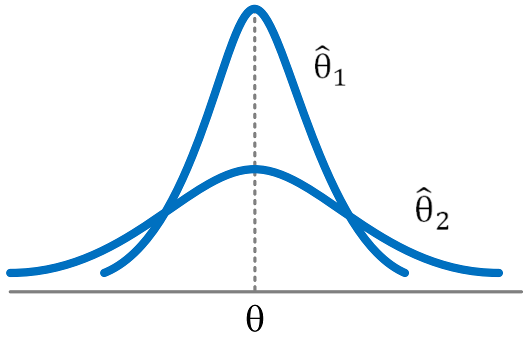
The condition for the unbiased estimator is when the statistics of a sample are equal to the parameters of a population $\hat{\theta} = \theta$





b. Relatively Efficient Estimator

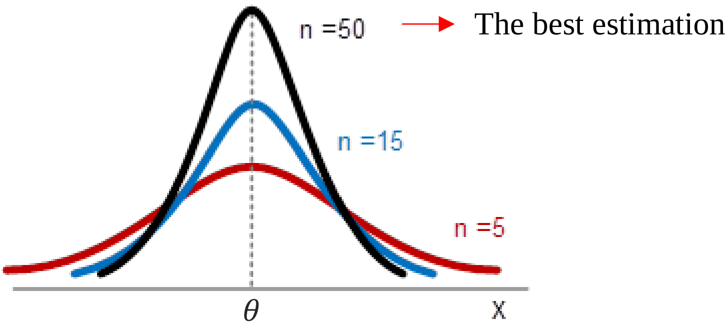
Relatively Efficient Estimator is the variance of the sample used to estimate very small parameters (high precision)



Variance of $\hat{\theta}_1$ less than $\hat{\theta}_2$ So the precision of $\hat{\theta}_1$ greater than $\hat{\theta}_2$

c. Consistent Estimator

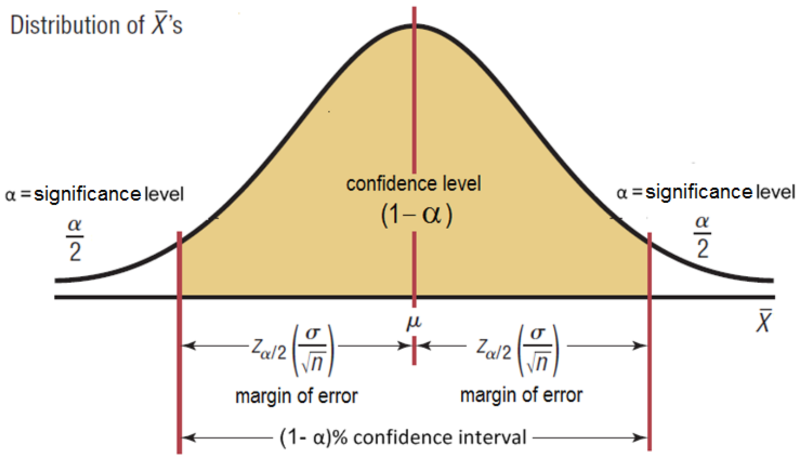
The statistic used as an estimator has a value that is close to the parameter value in line with the increase in the sample size used.



TERMS IN ESTIMATION

- **The interval estimation** is the interval or range of values from a statistic used to estimate a parameter
- **Confidence Interval** is a possible interval estimate containing parameters derived from sample selection.
- **The confidence level** is a level of confidence which is equivalent to $(1-\alpha)$ or in percentage form, it can be written as $100(1-\alpha)\%$. This level of confidence that is often used is 90%, 95%, and 99%
- **Error tolerance or Level of significance** is a value of the error level of an estimate, its size is α. Because the confidence interval is in the form of an interval, then the position of $(1-\alpha)$ is in the middle, so α is divided into two sides, namely the left side and the right side so that the size becomes $\alpha/2$.

Description of the relationship between $1-\alpha$ and α



ESTIMATION OF THE MEAN

Interval estimation of the mean is divided into two categories, namely the estimation with a large sample and a small sample. Besides, there is also the interval estimate for the difference between the two means. The following is the description.

1. Interval Estimation of the Mean in a Large Sample

a. The Variance of the Population is Known

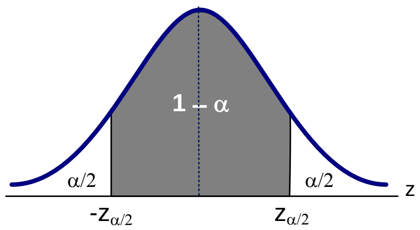
It is known that the Z statistical test (Z score) as follows:

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

From the formula above, the population mean (μ) is as follows:

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \quad \Longrightarrow \quad Z.{}^{\sigma}/_{\sqrt{n}} = \bar{x} - \mu \quad \Longrightarrow \boldsymbol{\mu = \bar{x} - Z.{}^{\sigma}/_{\sqrt{n}}}$$

If μ above is estimated using interval estimation, it can be described as follows.



$1 - \alpha$ is the area where the mean of a population is located. $Z_{\alpha/2}$ atau $-Z_{\alpha/2}$ is the interval limit. The confidence interval (CI) of the mean (μ) with the confidence level (CL) of $1 - \alpha$ can be formulated as:

$$\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$Z_{\alpha/2}$ or $-Z_{\alpha/2}$ is a statistical test that describes the value of the normal distribution resulting in an area of $\alpha/2$ on the right and left sides. By using a standard normal table, the value of $Z_{\alpha/2}$ can be found which is determined by the confidence level of $\alpha/2$.

The Margin of Error and the Sample Size

The formula of the margin of error

$$d = Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

The Sample size can be found:

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

Summary of $Z_{\alpha/2}$ values for a specific $(1-\alpha)$ or $(1-\alpha)100\%$ confidence level:

- If $(1-\alpha)100\% = 90\%$, then $\alpha = 10\%$ so $Z_{0.05} = 1.65$
- If $(1-\alpha)100\% = 95\%$, then $\alpha = 5\%$ so $Z_{0.025} = 1.96$
- If $(1-\alpha)100\% = 99\%$, then $\alpha = 1\%$ so $Z_{0.005} = 2.56$

b. The Variance of the Population is Unknown

If the variance of the population is unknown and the sample used is very large, then σ can be replaced by the standard deviation (s) of the sample so the confidence interval of this condition is:

$$\bar{x} - Z_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{x} + Z_{\alpha/2}\frac{s}{\sqrt{n}}$$

The margin of error:

$$d = Z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

The sample size:

$$n = \left(\frac{Z_{\alpha/2}.s}{d}\right)^2$$

2. Interval Estimation of the Mean in a Small Sample and the Variance of the Population is Unknown

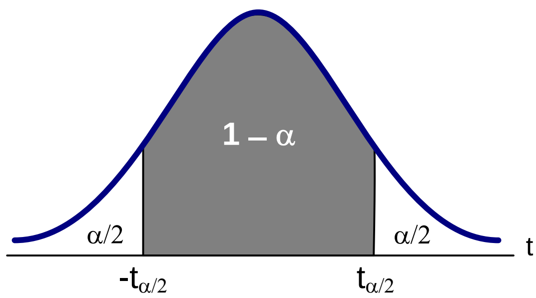
A small sample is when the sample size is less than 30. It is known that the t statistical test (t score) of the mean is:

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

From the formula above, the population mean (μ) is as follows:

$$t = \frac{\bar{x} - \mu}{{}^{s}/_{\sqrt{n}}} \quad \Rightarrow \quad t.{}^{s}/_{\sqrt{n}} = \bar{x} - \mu \quad \Rightarrow \boldsymbol{\mu = \bar{x} - t.{}^{s}/_{\sqrt{n}}}$$

If μ above is estimated using interval estimation, it can be described as follows.



$1-\alpha$ is the area where the mean of a population is located. $t_{\alpha/2}$ or $-t_{\alpha/2}$ is the interval limit. Confidence interval (CI) of the mean (μ) with the confidence level (CL) of $1-\alpha$ can be written as:

$$\bar{x} - t_{\alpha/2,v} \times \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2,v} \times \frac{s}{\sqrt{n}}$$

$t_{\alpha/2,v}$ is the t statistical test value obtained from the t distribution table with degrees of freedom of $v = n-1$ resulting in $\alpha/2$ on the right and left $(1-\alpha)$.

The margin of error:

$$d = t_{\alpha/2,v}\frac{s}{\sqrt{n}}$$

The sample size:

$$n = \left(\frac{t_{\alpha/2,v} \cdot s}{d}\right)^2$$

3. Interval Estimation of the Difference between the Two Means

This estimation is an estimate of the difference in means $(\mu_1 - \mu_2)$ derived from two populations. There are three possibilities for the estimation of the difference in means, namely:

a. If the variances (σ_1^2 and σ_2^2) of the populations are known.
b. If the variances (σ_1^2 and σ_2^2) of the populations are unknown and $\sigma_1^2 = \sigma_2^2$
c. If the variances (σ_1^2 and σ_2^2) of the populations are unknown and $\sigma_1^2 \neq \sigma_2^2$

a. Interval estimation of the variances (σ_1^2 and σ_2^2) of the populations are known

It is known that:

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \implies (\mu_1 - \mu_2) = (\bar{x}_1 - \bar{x}_2) - Z.\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

If $\overline{x}_1$ and $\overline{x}_2$ are the mean of random samples of size n_1 and n_1 respectively taken from the population with σ_1^2 and σ_2^2 are known, then the confidence interval (CI) with the confidence level of (1 – α)100% for $\mu_1 - \mu_1$ is:

$$(\bar{x}_1 - \bar{x}_2) - Z_{\frac{\alpha}{2}} \times \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_1 < (\bar{x}_1 - \bar{x}_2) + Z_{\frac{\alpha}{2}} \times \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

The margin of error:

$$d = Z_{\frac{\alpha}{2}} \times \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

b. Interval estimation of the variances ($\boldsymbol{\sigma_1^2}$ and $\boldsymbol{\sigma_2^2}$) of the populations are unknown and $\boldsymbol{\sigma_1^2 = \sigma_2^2}$

It is known that the pooled standard deviation of this condition is:

$$S_p = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

If $\overline{x}_1$ and $\overline{x}_2$ are the mean of random samples of size n_1 and n_1 respectively taken from the population with σ_1^2 and σ_2^2 are unknown and $\sigma_1^2 = \sigma_2^2$, then the confidence interval (CI) with the confidence level (CL) of (1 – α)100% for $\mu_1 - \mu_1$ is:

$$(\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2},v} \times S_p < \mu_1 - \mu_1 < (\bar{x}_1 - \bar{x}_2) + t_{\frac{\alpha}{2},v} \times S_p$$

$$v = n_1 + n_2 - 2$$

The margin of error:

$$d = t_{\frac{\alpha}{2},v} \times S_p = t_{\frac{\alpha}{2},v} \times \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

c. Interval estimation of the variances ($\boldsymbol{\sigma_1^2}$ and $\boldsymbol{\sigma_2^2}$) of the populations are unknown and $\boldsymbol{\sigma_1^2 \neq \sigma_2^2}$

It is known that the t statistical test for difference of the mean:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \Rightarrow \boldsymbol{(\mu_1 - \mu_2) = (\bar{X}_1 - \bar{X}_2) - t \times \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

If $\overline{x}_1$ & s_1^2 and $\overline{x}_2$ & s_2^2 are the means and variances of two random samples with sample sizes of n_1 and n_1 drawn from two populations with the condition $\sigma_1^2 \neq \sigma_2^2$, then the confidence interval (CI) with the confidence level of (1 – α)100% for $\mu_1 - \mu_1$:

$$(\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2},v} \times \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_1 < (\bar{x}_1 - \bar{x}_2) + t_{\frac{\alpha}{2},v} \times \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left[\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}\right]}$$

The margin of error:

$$d = t_{\frac{\alpha}{2},v} \times \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

ESTIMATION OF THE PROPORTION

1. Interval Estimation of the Population from a Single Sample

Suppose there is a condition A as x in a sample size of n, then the proportion of A is:

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

and $\hat{q} = 1 - \hat{p}$ (other conditions).

Example:

In a class of 50 students, there are 20 students from Bandung. The proportion of students who come from Bandung is:

$$\hat{p} = \frac{20}{50} = 0.4$$

and the proportion of students who come from outside Bandung:

$$\hat{q} = 1 - \hat{p} \Longrightarrow 1 - 0.4 = 0.6$$

If we want to know the range of the proportion of students coming from Bandung in a population, it can be estimated using interval estimation at a certain confidence level and the formula is divided into two, namely for a large and small sample.

a. A Large Sample Size

$$\hat{p} - Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}} < p < \hat{p} + Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}$$

The margin of error:

$$d = Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}$$

The sample size:

$$n = \left(\frac{Z_{\alpha/2}}{d}\right)^2 \hat{p} \cdot \hat{q}$$

b. A Small Sample Size

$$\hat{p} - t_{\alpha/2,v} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}} < p < \hat{p} + t_{\alpha/2,v} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}$$

$$v = n - 1$$

The margin of error:

$$d = t_{\alpha/2,v} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}$$

The sample size:

$$n = \left(\frac{t_{\alpha/2,v}}{d}\right)^2 \hat{p} \cdot \hat{q}$$

If the sample used is very small, the proportion of the population tends to 0 or 1, then the method of finding this confidence interval is unreliable and should not be used.

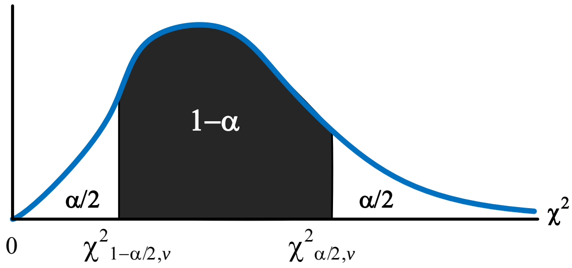
2. Interval Estimation of the Difference between the Two Proportions

If $\hat{p}_1$ is the proportion of a random sample with a sample size of n_1 and $\hat{p}_2$ is the proportion of a random sample with a sample size of n_2, then confidence interval for the difference in proportions at the confidence level of $(1-\alpha)100\%$ is:

$$(\hat{p}_1 - \hat{p}_2) - Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

ESTIMATION OF THE VARIANCE

Interval estimation of the variance is used to estimate the variance or standard deviation of a normally distributed population with unknown μ and σ^2. Suppose there is a χ^2 curve as shown below.



$1-\alpha$ is an area whose value is between 0 and 1 and can be found with:

$$1-\alpha = P\left(\chi^2_{1-\alpha/2,v} < \chi^2 < \chi^2_{\alpha/2,v}\right)$$

$$v = n-1$$

with

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

Based on the formulas above, it can be seen that the formula for the confidence interval for the variance is as follows:

$$\chi^2_{1-\alpha/2,v} < \frac{(n-1)s^2}{\sigma^2} < \chi^2_{\alpha/2,v}$$

the formula above changes to:

$$\boldsymbol{\frac{(n-1)s^2}{\chi^2_{\alpha/2,v}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,v}}}$$

EXERCISES

1. Interval Estimation of the Mean (A Large Sample Size)

Suppose a student wants to examine the learning time for one week from students in his program study. He gets a sample of 30 students, with a mean of 14 hours, and the variance is 9 hours. Based on this statement, then calculate:

a. 95% confidence interval (CI) for the mean of the study time of all students in the study program for 1 week!
b. If the desired margin of error (d) is 1 hour, then specify the sample to be taken!

The answer:

a. $\bar{x} = 14 \qquad \sigma^2 = 9 \Longrightarrow \sigma = 3 \qquad n = 30$

$(1 - \alpha) = 95\%$ so $\alpha = 5\% = 0.05$ and $\alpha/2 = 0.025$

Then:

$-Z_{\frac{\alpha}{2}} = -Z_{0.025} = -1.96$ (for left side)

$Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$ (for left side)

95% CI of the mean is:

$$\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$\Longrightarrow 14 - 1.96\frac{3}{\sqrt{30}} < \mu < 14 + 1.96\frac{3}{\sqrt{30}} = 14 - 1.07 < \mu < 14 + 1.07$$

$$\Longrightarrow 12.93 < \mu < 15.07$$

conclusion:

With a confidence level of 95%, an estimate for the mean of learning time for 1 week is between 12.93 hours to 15.07 hours or the equivalent of 12 hours 55 minutes to 15 hours 4 minutes.

b. $d = 1 \qquad Z_{\alpha/2} = 1.96 \qquad \sigma = 2$

so

$$n = \left(\frac{Z_{\alpha/2}.\sigma}{d}\right)^2 \Longrightarrow n = \left(\frac{1.96 \times 3}{1}\right)^2 = 34.57 \approx 35$$

So to get a margin of error of 1, the sample that must be obtained is 35 students.

2. Interval Estimation of the Mean (A Small Sample Size)

A telecommunications company is testing battery life on the latest mobile phones. The trial was carried out on 20 cellphones which were taken randomly from the production results in 1 day. The test results show that the mean of battery life in normal use of 20 cellphones is 14 hours and the variance is 4 hours. Determine a 90% confidence interval for the average battery life of all products for 1 day.

The answer:

$n = 20 \qquad \bar{x} = 14 \qquad s^2 = 4 \Rightarrow s = 2$

$100(1-\alpha)\% = 90\% \;\Rightarrow\; \alpha = 10\% = 0.1$

$v = n - 1 = 19$

$\alpha = 10\% = 0.1 \Rightarrow \alpha/2 = 0.05$

$-t_{\alpha/2,v} = -t_{0.05,19} = -1.729$ (for left side)

$t_{\alpha/2,v} = t_{0.05,19} = 1.729$ (for right side)

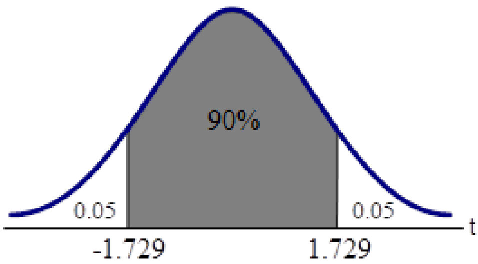
90% confidence interval of the mean:

$$\bar{x} - t_{0.05,19}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{0.05,19}\frac{s}{\sqrt{n}}$$

$$14 - 1.729\frac{2}{\sqrt{20}} < \mu < 14 + 1.729\frac{2}{\sqrt{20}}$$

$$14 - 0.773 < \mu < 14 + 0.773$$

$$13.227 < \mu < 14.773$$



Conclusion:
Based on the CI calculation, it can be concluded that with a confidence level of 90%, the mean of battery life when the cellphone is used in normal conditions is between 13,227 hours to 14,773 hours or 13 hours 14 minutes to 14 hours 28 minutes.

3. Estimation of the Proportion

The results of a survey on the use of gadgets in children aged 5-6 years showed that out of 100 children randomly taken in a city, only 40 children used gadgets to learn. By using a 90% confidence level, determine the confidence interval of all children in the city using gadgets to learn!

The Answer:

$n = 100 \qquad x = 40 \qquad \hat{p} = \frac{40}{100} = 0.4 \Rightarrow \hat{q} = 0.6$

$(1-\alpha)100\% = 90\% \;\Rightarrow \alpha = 10\%$

$\alpha = 10\% = 0.1 \Rightarrow \alpha/2 = 0.05$

$-Z_{0.05} = -1.65$ (for left side)

$Z_{0.05} = 1.65$ (for right side)

90% CI of the proportion:

$$\hat{p} - Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}} < p < \hat{p} + Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}$$

$$\Rightarrow 0.4 - 1.65 \cdot \sqrt{\frac{0.4 \cdot 0.6}{100}} < p < 0.4 + 1.65 \cdot \sqrt{\frac{0.4 \cdot 0.6}{100}}$$

$$\Rightarrow 0.4 - 0.08 < p < 0.4 + 0.08$$

$$\Rightarrow 0.32 < p < 0.48$$

You can write: $32\% < p < 48\%$

Conclusion:

With a confidence level of 90%, it was found that the proportion of children aged 5-6 years in a city using gadgets for learning was between 32% and 48% or between 32 to 48 children in that city using gadgets for learning.

4. Estimation of the Variance

It is known that the Matematika Diskrit test score of 15 students are:

100	90	80	85	90
75	85	80	80	85
75	90	100	90	95

Of the 15 scores, we want to know with a confidence level of 90%, what is the confidence interval for the variance and the standard deviation of all students' Matematika Diskrit test scores?

$n = 15 \qquad (1 - \alpha) = 90\%$

$v = 15 - 1 = 14$

- calculate the variance of a sample using one of the formulas:

$$\sum_{i=1}^{n} (x_i)^2 = (100)^2 + (90)^2 + (80)^2 + (85)^2 + (90)^2 + \cdots + (95)^2 = 113550$$

$$\left(\sum_{i=1}^{n} x_i\right)^2 = (100 + 90 + 80 + 85 + 90 + \cdots + 95)^2 = 1690000$$

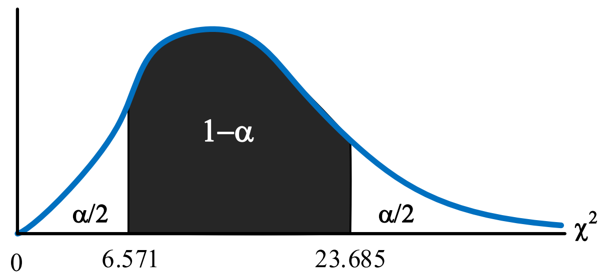
$$s^2 = \frac{n\sum_{i=1}^{n}(x_i)^2 - (\sum_{i=1}^{n} x_i)^2}{n(n-1)} = \frac{15(113550) - 1690000}{15(14)} = \frac{13250}{210} = 63.09$$

- Determine the location of $1-\alpha$ on the chi-square table

  $1-\alpha = 90 \Longrightarrow \alpha = 10\% \Longrightarrow \alpha/2 = 5\% = 0.05$

  $\chi^2_{1-\alpha/2,v} = \chi^2_{0.95,14} = 6.571$ (for left side)

  $\chi^2_{\alpha/2,v} = \chi^2_{0.05,14} = 23.685$ (for right side)



- Determine the confidence interval for the variance with a 90% confidence level

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,v}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,v}} \Longrightarrow \frac{14(63.09)}{23.685} < \sigma^2 < \frac{14(63.09)}{6.571}$$

$$\Longrightarrow 37.29 < \sigma^2 < 134.42$$

- Determine the confidence interval for the standard deviation with a 90% confidence level

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2,v}}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,v}}} = 37.29 < \sigma^2 < 134.42 = 6.11 < \sigma < 11.59$$

Conclusion:

Based on the calculation of the confidence interval above, the estimation of the variance in the test scores of all students taking Matematika Diskrit courses is between 37.29 and 134.42 and the standard deviation is between 6.11 and 11.59.