Universiti Malaysia PAHANG
Engineering • Technology • Creativity

CHAPTER 2

SAMPLING DISTRIBUTIONS AND CONFIDENCE INTERVAL (CI)

Universiti Malaysia Pahang
26600 Pekan, Pahang Darul Makmur
Tel: +609 424 5000 • Fax: +609 424 5555

Expected Outcomes

- ✓ Able to find the confidence interval for the population mean when population variance is known or unknown.
- ✓ Able to find the confidence interval for the difference between two population means of independent samples when population variances are known or unknown.
- ✓ Able to find the confidence interval for the difference between two population means of dependent samples when the population variance of the differences is known or unknown.
- ✓ Able to find the confidence interval for population proportion and the confidence interval for the difference between two population proportions.
- ✓ Able to find the confidence interval for population variance and standard deviation; and the confidence interval for the ratio of two population variances.

CONTENT

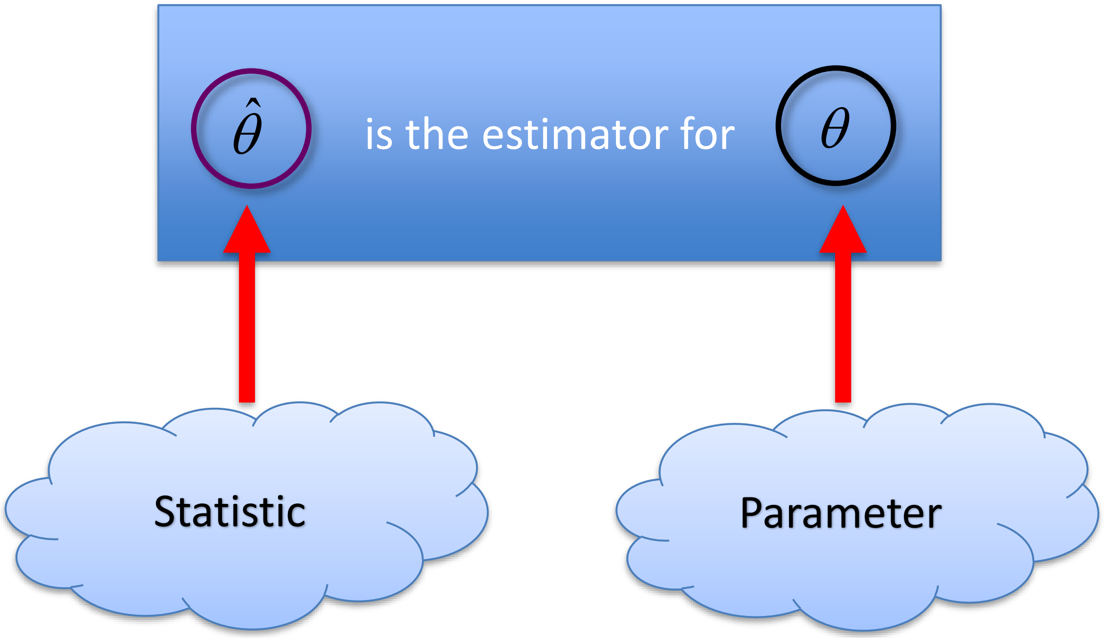
2.1 Estimate, Estimation, Estimator
2.1.1 Definition
2.1.2 Properties of Good Point Estimator
2.1.3 Confidence Interval
2.2 Confidence Interval for Population Mean
2.3 Confidence Interval for Difference Between Two Population Means
2.3.1 Independent Samples
2.3.2 Dependent Samples
2.4 Confidence Interval for Population Proportion
2.5 Confidence Interval for Difference Between Two Population Proportions
2.6 Confidence Interval for Population Variance and Population Standard Deviation
2.7 Confidence Interval for Ratio of Two Population Variances

2.1 ESTIMATE, ESTIMATION, ESTIMATOR

- ✓ Define and identify the statistical term of estimate, estimation and estimator.

2.1.1 Definition

- The probability functions that represent statistical models such as Poisson, Binomial and Normal distributions may include one or more parameter(s).
- The parameters that have been widely used are **mean**, **variance** and **proportion**.
- Any functions of a random sample whose objective is to approximate a parameter is called an **estimator**.



- Estimation is the entire process of using an estimator $(\hat{\theta})$ to estimate the parameter, θ.
- Two types of estimation

1. **Point Estimation -** a process that produces a specific numerical value calculated from a sample data representing the unknown population parameter.

Parameter		Point Estimator
μ	→	$\bar{x}$
σ^2	→	s^2
π	→	p

2. **Interval Estimation** - a process that produces range of values calculated from the sample data, forming an interval within which the parameter is estimated to lies $a < \theta < b$ where $a, b \in (-\infty, \infty)$.

2.1.2 Properties of Good Point Estimator

- There are four main properties of a **good estimator**.

    1. Unbiased
    2. Consistency
    3. Efficiency
    4. Sufficient

- These properties should be **satisfied** for choosing the best estimator of a parameter. In this course, we **assume** that the best estimator for a parameter has satisfied all the properties.

2.1.3 Confidence Interval

- A confidence interval gives an estimated range of values which is likely to include an unknown population parameter, θ with a specified probability within that interval.
- The interval is usually written as (a, b) or $a < \theta < b$.
- **The end points** of the interval are the confidence limits and the specified probability is called the confidence level.
- Commonly used are 95%, 90% and 99%.
- **Confidence limits** represent the **lower and upper boundaries** which define the **range or width** of a confidence interval.
- By looking at the width of a confidence interval, we can get a good sense of the **estimator precision**.
- The smaller the interval width, the better the confidence interval describes the parameter indicating that the interval estimate is more precise and accurate.

For a confidence interval $a < \theta < b$, where a is lower boundary and b is upper boundary:

Interval width = $b - a$

- The **confidence level** is the probability value, $(1-\alpha)$ that is associated with a confidence interval.

- In general, $P(a < \theta < b) = 1 - \alpha$, where α value represents the **level of significance** (often expressed as a percentage).

- The confidence level is also referred to the reliability of the interval if we think the width of the interval as identifying its precision or accuracy.

- **Example**, at significance level of $\alpha = 0.05 = 5\%,$ the confidence level is $100\% \times (1 - 0.05) = 95\%.$

- Interpretation of confidence interval
  - The 95% confidence interval means that we are 95% confident that the true parameter lies within the interval.
  - It means the probability that interval includes θ is 95% or $P(a < \theta < b) = 0.95.$

In general, a $(1-\alpha)100\%$ **Confidence Interval for** θ **is given by:**

$$P\left(\hat{\theta}-\underbrace{\left(\text{distribution for } \hat{\theta}\right)\left(s.d \text{ for } \hat{\theta}\right)}_{\text{Estimation Error, } E} < \theta < \hat{\theta}+\underbrace{\left(\text{distribution for } \hat{\theta}\right)\left(s.d \text{ for } \hat{\theta}\right)}_{\text{Estimation Error, } E}\right)=1-\alpha$$

Lower bound — Upper bound

where

$s.d$: a standard deviation of the estimator $\hat{\theta}$

θ : a parameter

$\hat{\theta}$: a statistic

$P\left(\hat{\theta}-E<\theta<\hat{\theta}+E\right)=1-\alpha$

Interval width $=2E$.

Note that, the estimation error *E* is also called as margin of error

Hence, a $(1-\alpha)$ 100% **Confidence Interval for** θ

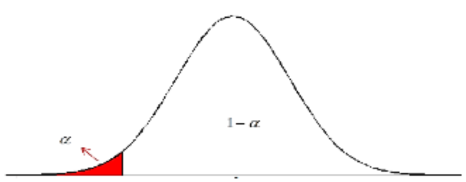
$$=\hat{\theta}\pm\left(\text{distribution for } \hat{\theta}\right)\left(s.d \text{ for } \hat{\theta}\right).$$

$$=\hat{\theta}\pm E$$

One-Sided Confidence Interval

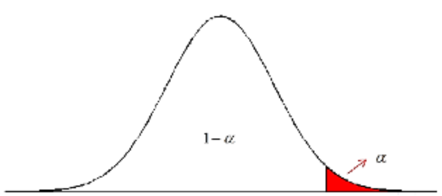
- In some cases, a researcher is interested to construct one-sided confidence interval, namely, one-sided lower bound or one-sided upper bound for θ.
- A one-sided confidence bound for θ results from replacing $\alpha/2$ by α, and $\pm$ sign by either + or – sign in the confidence interval formula for θ.

A One-Sided Lower Bound and A One-Sided Upper Bound for θ



$$\theta > \hat{\theta} - \left(distribution\ for\ \hat{\theta}\right)\left(s.d\ for\ \hat{\theta}\right)$$

one-sided lower bound



$$\theta < \hat{\theta} + \left(distribution\ for\ \hat{\theta}\right)\left(s.d\ for\ \hat{\theta}\right)$$

one-sided upper bound

where

$s.d$: a standard deviation of the estimator $\hat{\theta}$

θ : a parameter

$\hat{\theta}$: a statistic

2.2 CONFIDENCE INTERVAL FOR POPULATION MEAN

- ✓ Estimate the confidence interval for the population mean when population variance/standard deviation is known or unknown.
- ✓ Estimate the sample size by using the concept of confidence interval for population mean.

How to construct CI

- We assume that all data are comes **from normal distribution .** A random variable with normal distribution, X
  - Sample mean, $\bar{X}$ is the best estimator for population mean, μ .
  - $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$
- Based on the CLT,

$$Z = \frac{X - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

the confidence interval for the population mean, μ can be constructed as follows (next slide).

$$1-\alpha = P\left(-Z_{\alpha/2} < Z < Z_{\alpha/2}\right)$$

$$= P\left(-Z_{\alpha/2} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < Z_{\alpha/2}\right)$$

$$= P\left(-\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < -\mu < -\bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

$$= P\left(\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

- Thus, a $(1-\alpha)100\%$ confidence interval for population mean, μ is given by

$$\left(\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \quad \text{or} \quad \left(\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

if the population variance and standard deviation is known.

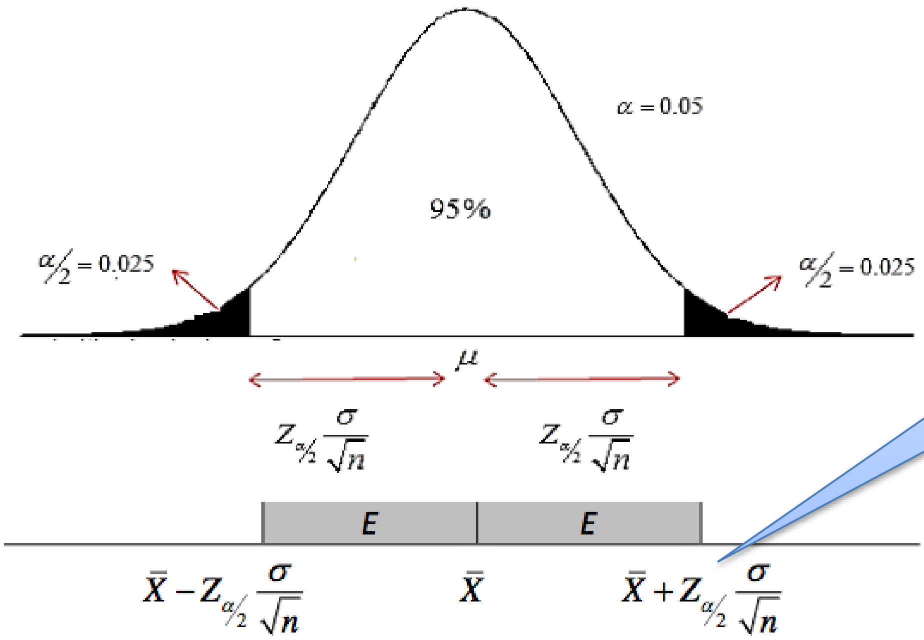


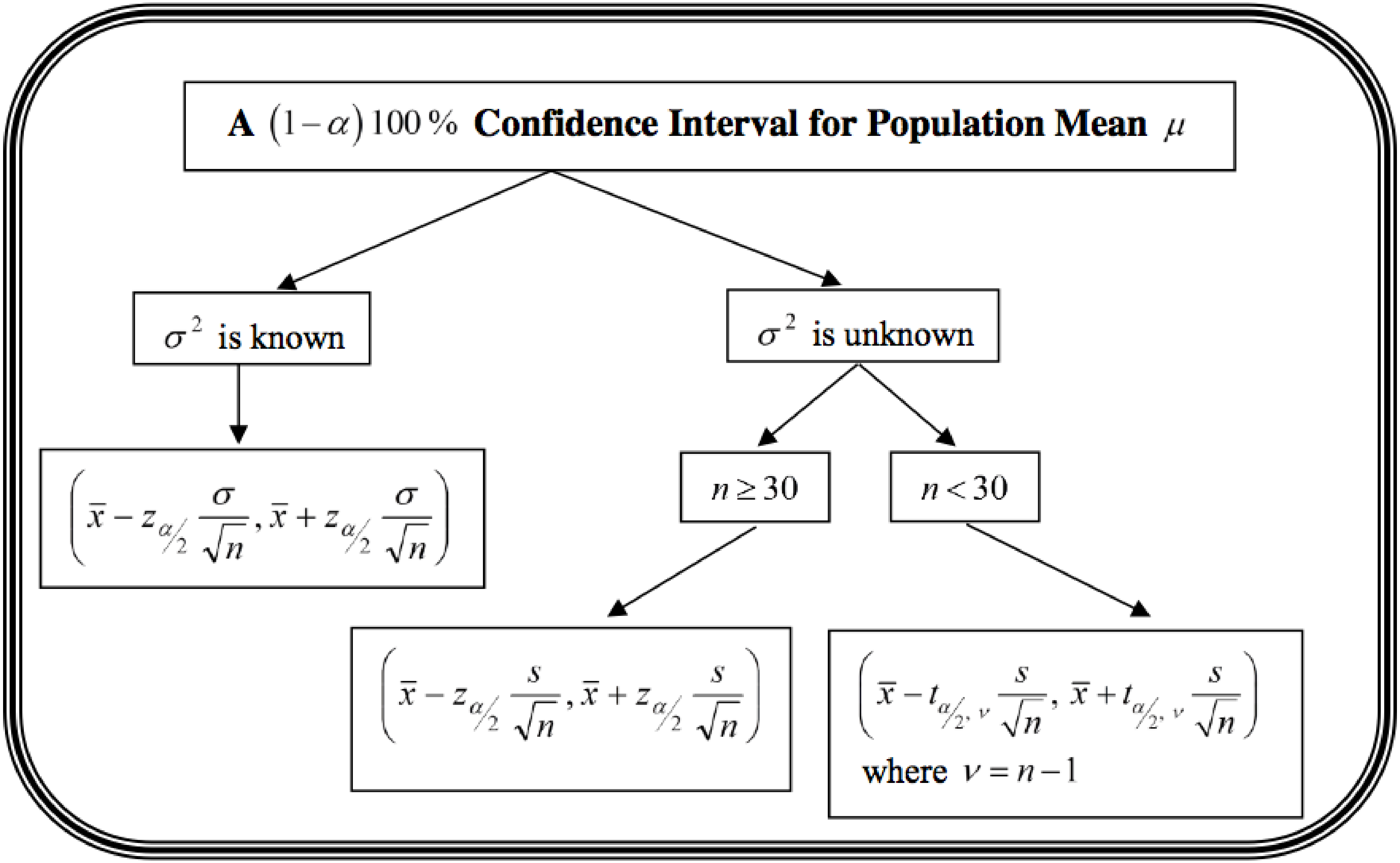
Figure 2.5: Distribution of $\bar{X}$

The term is a critical value $Z\alpha/_2$ obtained from the statistical tables

- $P\left(\bar{x} - E < \mu < \bar{x} + E\right) = 1 - \alpha$
- In general, for a specific value, let say $\alpha = 0.05$, 95% of the sample means will fall within this error value on either side of the population mean, μ

Consider $E = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$,

- if $n\uparrow$ $E\downarrow$ (the width of the confidence interval is small)
- if $\sigma\uparrow$ $E\uparrow$ (the width of the confidence interval is large)
- if $\alpha\uparrow$ $E\downarrow$ (the width of the confidence interval is small)


A $(1-\alpha)100\%$ Confidence Interval for Population Mean μ
σ^2 is known
$\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$
σ^2 is unknown
$n \geq 30$
$n < 30$
$\left(\bar{x} - z_{\alpha/2}\frac{s}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{s}{\sqrt{n}}\right)$
$\left(\bar{x} - t_{\alpha/2,\nu}\frac{s}{\sqrt{n}}, \bar{x} + t_{\alpha/2,\nu}\frac{s}{\sqrt{n}}\right)$
where $\nu = n-1$

Example 2.1

A factory produces steel sheets whose weights are known to be normally distributed with a standard deviation of 2.4 kg. A random sample of 26 sheets had a mean weight of 3.14 kg. Find a 99% confidence interval for the population mean weight and give an interpretation of the parameter estimate.

Solution:

X : The weight of steel sheets

$n = 26, \quad \bar{x} = 3.14 \text{ kg}, \quad \sigma = 2.4 \text{ kg} \quad \rightarrow \quad \sigma^2$ is known

$z_{0.01/2} = z_{0.005} = 2.5758$

A 99% confidence interval for population mean, μ

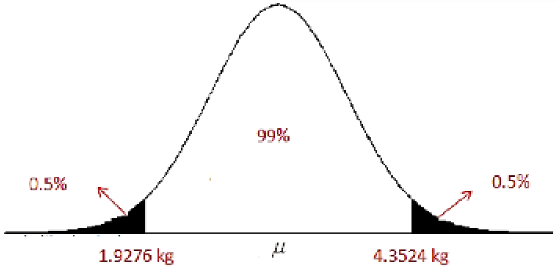
$$= \left(\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right)$$

$$= \left(3.14 - 2.5758 \left[\frac{2.4}{\sqrt{26}} \right], 3.14 + 2.5758 \left[\frac{2.4}{\sqrt{26}} \right] \right)$$

$$= (3.14 - 1.2124, 3.14 + 1.2124)$$

$$= (1.9276,\ 4.3524) \text{ kg}$$

Noted that
- the estimation error is and E=1.2124 kg.
- the width of the confidence interval is 4.3524 - 1.9276 = 2.4248 kg.



$$P(1.9276 < \mu < 4.3524) = 0.99$$

Interpretation: We are 99% confident that the population mean weight of steel sheets lies within 1.9276 and 4.3524 kg.

EXAMPLE 2.2:

Based on Example 2.1, investigate the effect of sample size, standard deviation and confidence level values on the confidence limit if we

a. Change the sample size to 100.

a. Change the population standard deviation to 1 kg.

a. Change the confidence level to 95%.

Example 2.3

A random sample of 25 bulbs was selected and the specified brightness were evaluated for each bulb by measuring the amount of electrical current required. The sample mean of the electrical current required for the bulbs is 280.3 micro amps and the sample standard deviation is 10.3 micro amps. Construct a 90% confidence interval for the population mean and interpret your answers.

Solution:

X : The electrical current required for bulbs

$n = 25$, $\bar{x} = 280.3$ micro amps, $s = 10.3$ micro amps $\rightarrow$ σ^2 is unknown and $n < 30$.

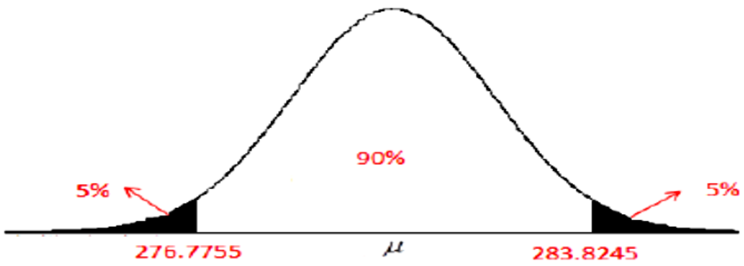
$t_{0.1/2,25-1} = t_{0.05,24} = 1.7109$

A 90% confidence interval for the population mean, μ

$$= \left(\bar{x} - t_{\alpha/2,\, n-1} \frac{s}{\sqrt{n}},\ \bar{x} + t_{\alpha/2,\, n-1} \frac{s}{\sqrt{n}} \right)$$

$$= \left(280.3 - 1.7109 \frac{10.3}{\sqrt{25}},\ 280.3 + 1.7109 \frac{10.3}{\sqrt{25}} \right)$$

$$= (276.7755,\ 283.8245) \text{ micro amps}$$



$P(276.7755 < \mu < 283.8245) = 0.99$

Interpretation: We are 90% confident that the population or true mean of the electrical current required for bulb lies within 276.7755 and 283.8245 micro amps.

Determination of Sample Size

Sample size, $n = \left(\dfrac{z_{\alpha/2}\sigma}{E} \right)^2$

where

$z_{\alpha/2}$: critical value

σ : population standard deviation

E : estimation error

Example 2.4

A telephone company wants to estimate the mean number of minutes people in a city spend talking long distance with 90% confident. From past records, an estimate of the standard deviation is 12 minutes. What is the **minimum sample size** required so that the estimation error of the mean lies within ±5?

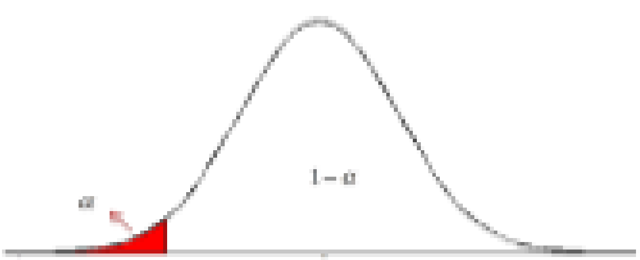
Solution:

$$\sigma = 12 \text{ minutes}, \; E = 5, \; z_{\alpha/2} = z_{0.05} = 1.6449$$

$$n \geq \left(\frac{z_{\alpha/2}\sigma}{E} \right)^2 \geq \left(\frac{1.6449(12)}{5} \right)^2 = 15.5848 \approx 16 \text{ people}$$

One-Sided Lower Bound and One-Sided Upper Bound for μ

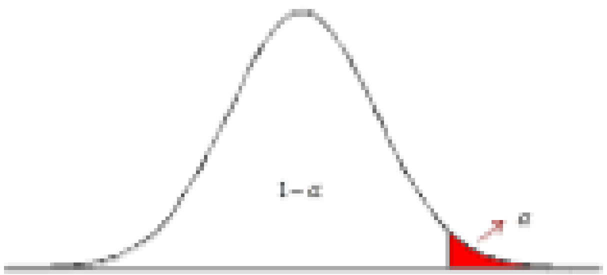
One-Sided Lower Bound and One-Sided Upper Bound for μ



$$\mu > \bar{x} - z_{\alpha}\frac{\sigma}{\sqrt{n}}$$

one-sided lower bound

$\mu \in (a, \infty)$



$$\mu < \bar{x} + z_{\alpha}\frac{\sigma}{\sqrt{n}}$$

one-sided upper bound

$\mu \in (-\infty, b)$

where $a, b \in \mathbb{R}$

Note: Refer ME.7 for example.

EXAMPLE 2.5:

A packet of baking powder supposed to have a mean weight of 200 g. The distribution of weight is normal and the population standard deviation is 7 g. A random sample of nine packets of baking powder had the following weights.

218, 207, 219, 200, 205, 221, 206, 205, 211

Construct a one sided upper-bound of a 98% confidence interval for the population mean weight of a packet of baking powder. Interpret the result.

Solution:

X: Weight of a packet of baking powder

$n = 9, \quad \mu = 200g, \quad \sigma = 7g, \quad \bar{x} = 210.2222g, \quad z_{0.02} = 2.0537$

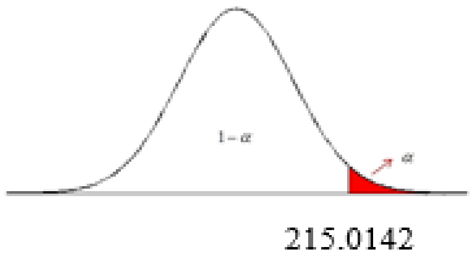
A one-sided upper bound of a 98% confidence interval for the population mean, μ

$$= \left(\bar{x} \pm z_{\alpha} \frac{\sigma}{\sqrt{n}} \right)$$

$$= \left(0, 210.2222 + \frac{2.0537(7)}{\sqrt{9}} \right)$$

$$= (0, 210.2222 + 4.7920)$$

$$= (0, 215.0142)\ g$$



Interpretation: We are 98% confident that the population mean weight of a packet of baking powder that contain one-sided upper bound is between 0g to 215.0142g.

EXERCISE 2.1

1. *The mass of vitamin E in a capsule manufactured by a certain drug company is normally distributed with standard deviation 0.042 mg. A random sample of 5 capsules was analysed and the mean mass of vitamin E was found to be 5.12 mg. Find a 95% confidence interval for the population mean mass of vitamin E per capsule. Give an interpretation of the parameter estimate.*

2. *A random number of 100 pieces of wood is cut using a machine. The sample mean of length is 1.06 cm and the sample standard deviation is 0.08 cm.*

   a) Find a 90% confidence interval for mean length of all the woods cut by the machine and interpret your answer.

   b) What is the width of this confidence interval?

3. *A study on number of drills that can be made by microdrill machine is conducted. A random sample of 15 microdrill machines was drawn and the data is given as follows.*

   13 11 10 15 12 8 16 13 11 11 14 10 15 12 9

   Find a 95% confidence interval for the mean number of drills that can be made by the microdrill machine. Give comment on the parameter estimate.

4. *The result of a stress test, X is known to be normally distributed random variable with mean μ and standard deviation 1.3. It is required to have a 95% confidence interval for μ with total width less than 2. Find the least number of tests that should be carried out to achieve this.*

2.3 CONFIDENCE INTERVAL FOR DIFFERENCE BETWEEN TWO POPULATION MEANS

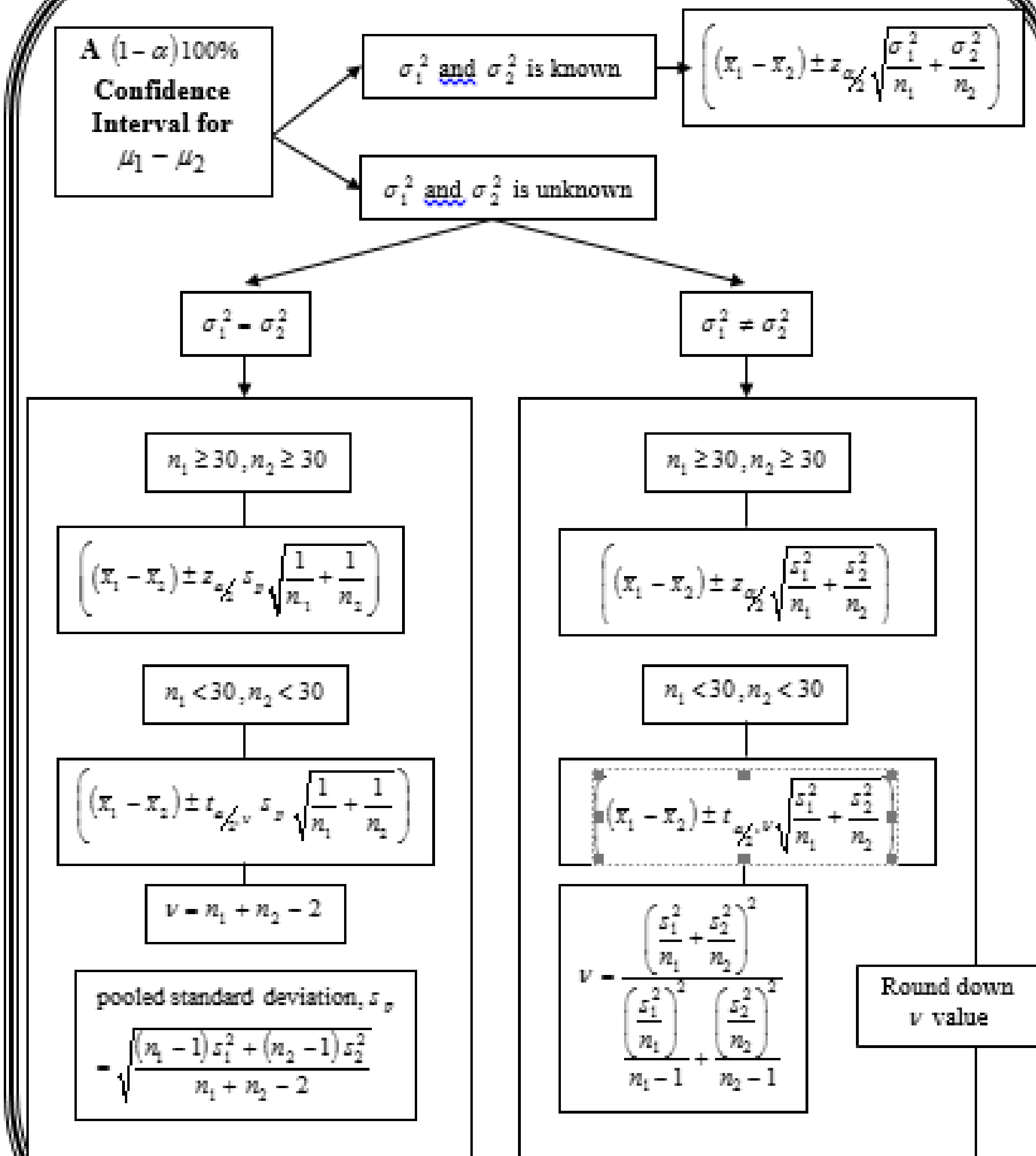
- ✓ Estimate the confidence interval for the difference between two population means of independent samples when the population variances are known or unknown.
- ✓ Estimate the difference between two population means of dependent samples when the population variance of the differences is known or unknown.

2.3.1 INDEPENDENT SAMPLES

- **Independent samples** are measurements made on two different sets of items which the samples are taken independently.

- For **example**, in a factory that producing a certain chemical, it is thought that a new process for producing this chemical is cheaper than the currently process used. So, we have two independent samples which come from the new process and current process for producing the certain chemical in the factory.

- This sub-topic considers estimation procedures for the mean difference of **two independent samples.**



Noted that $-\infty < \mu_1 - \mu_2 < \infty$, $\mu_1 - \mu_2$ from two different populations (independent cases).

Example 2.6

Two chemical companies, Company A and Company B supply a raw material where the most important element in this material is the concentration. The standard deviations of concentration produced by both companies are known as 5.81 and 4.70, respectively. The mean of concentration in a random sample of 10 batches produced by Company A is 90.25 grams per litre, while for Company B, the mean of concentration in a random sample 15 batches is 87.54 grams per litre. Consider the random variable concentration is normally distributed, construct a 96% confidence interval for the difference of the population means of the concentration produced by both companies. Give an interpretation for the parameter estimate.

Solution:

Universiti Malaysia PAHANG

X_A : The concentration of raw material produced by Company A

X_B : The concentration of raw material produced by Company B

Company	**A**	**B**
Sample size, n	10	15
Population Standard deviation, σ	5.81	4.70
Sample mean, $\bar{x}$	90.25	87.54

$\rightarrow$ σ^2 is known.

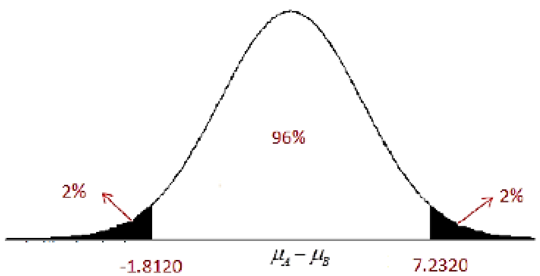
$$z_{\alpha/2} = z_{0.02} = 2.0537$$

A 96% confidence interval for difference between two population means, $\mu_A - \mu_B$

$$= \left((\bar{x}_A - \bar{x}_B) \pm z_{\alpha/2} \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}} \right)$$

$$= \left((90.25 - 87.54) \pm 2.0537 \sqrt{\frac{5.81^2}{10} + \frac{4.70^2}{15}} \right)$$

$$= (-1.8120,\ 7.2320) \text{ grams per liter}$$



$$P(-1.8120 < \mu_A - \mu_B < 7.2320) = 0.96$$

$\mu_A < \mu_B$ $\qquad$ $\mu_A > \mu_B$

Interpretation: We are 96% confident that the difference of the population mean for the concentration produced by both companies is between 1.8120 grams per litre more by Company B to 7.2320 grams per litre more by Company A, compared to each other, respectively.

Example 2.7

A biotechnology company produces a therapeutic drug which concentration is normally distributed. Two new methods of producing this drug have been proposed, and the method that produces more drugs will be chosen by management. The company chose a sample of size 15 from each method and obtained the following data in grams per litre.

Method	Mean	Standard deviation
X	16.22	0.05
Y	17.41	1.00

By assuming the population variances for both methods are equal, construct a 99% confidence interval on the difference between means of the concentration by method Y and method X. Interpret the answer resulted.

Solution:

X_X : The concentration of therapeutic drug by using Method X
X_Y : The concentration of therapeutic drug by using Method Y

Method	Mean	Standard deviation
X	$\bar{x}_X = 16.22$	$s_X = 0.05$
Y	$\bar{x}_Y = 17.41$	$s_Y = 1.00$

$\rightarrow$ σ^2 is unknown and $n_1 < 30,\ n_2 < 30$.

$$\text{pooled standard deviation, } s_p = \sqrt{\frac{(n_Y - 1)\, s_Y^2 + (n_X - 1)\, s_X^2}{n_Y + n_X - 2}}$$

$$= \sqrt{\frac{(15-1)\ 1.00^2 + (15-1)\ 0.05^2}{15+15-2}}$$

$$= 0.7080$$

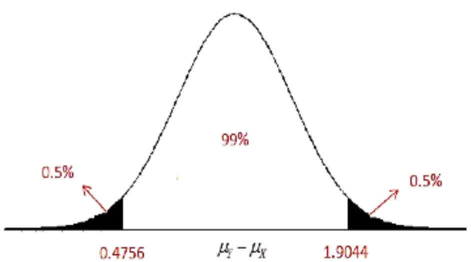
$$t_{\alpha/2,\ n_Y + n_X - 2} = t_{0.01/2,\ 15+15-2} = 2.7633$$

A 99% confidence interval for difference between two population means, $\mu_Y - \mu_X$

$$= \left((\bar{x}_Y - \bar{x}_X) \pm t_{\alpha/2,\ n_Y + n_X - 2}\, s_p \sqrt{\frac{1}{n_Y} + \frac{1}{n_X}} \right)$$

$$= \left((17.41 - 16.22) \pm (2.7633)(0.7080) \sqrt{\frac{1}{15} + \frac{1}{15}} \right)$$

$$= (0.4756,\ 1.9044)\ \text{g/l}$$

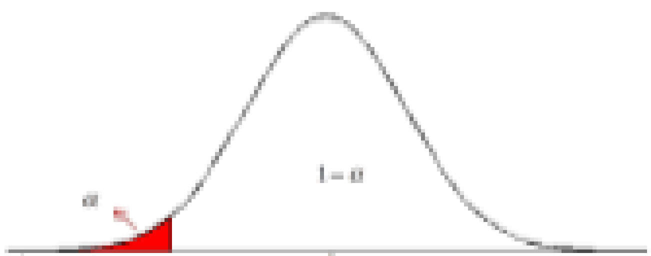


$$P(0.4756 < \mu_Y - \mu_X < 1.9044) = 0.99$$

Interpretation: We are 99% confident that the difference between the population means of the concentration produced by using method Y and method X lies within 0.4756 and 1.9044 grams per litre.

One-Sided Lower Bound and One-Sided Upper Bound for $\mu_1 - \mu_2$

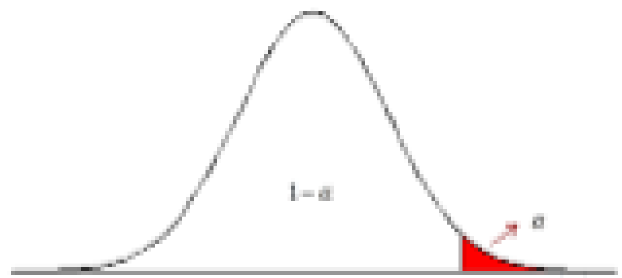
One-Sided Lower Bound and One-Sided Upper Bound for $\mu_1 - \mu_2$



$$\mu_1 - \mu_2 > (\bar{x}_1 - \bar{x}_2) - z_\alpha \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

one-sided lower bound

$\mu_1 - \mu_2 \in (a, \infty)$



$$\mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_\alpha \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

one-sided upper bound

$\mu_1 - \mu_2 \in (-\infty, b)$

where $a, b \in \mathbb{R}$

2. The burning rates of two different solid-fuel propellants used in aircrew escape systems are being studied. It is known that both propellants are normally distributed and have the same population variances. Two random samples of 26 specimens for each propellant are tested. The sample mean and standard deviation of burning rates for both propellants are summarised as in table below.

Type of Propellant	Mean (cm/s)	Standard deviation (cm/s)
Propellant 1	18.312	3.110
Propellant 2	24.412	2.708

Construct a 90% confidence interval on the difference in population mean for the propellant 2 and propellant 1. Interpret the answer resulted.

4. Two groups of students are given a problem solving test, and the test marks are compared. The data are as follows:

Mathematics Majors	***Computer Science majors***
$n_1 = 15$	$n_2 = 14$
$\bar{x}_1 = 83.6$	$\bar{x}_2 = 79.2$
$s_1 = 3.3$	$s_2 = 2.8$

Find the 98% confidence interval for the difference population mean of test marks between the two groups of students. Assume the variance population test marks are not the same for both groups. Give the comment on the parameter estimate.

1.The burning rates of two different solid-fuel propellants used in aircrew escape systems are being studied. It is known that both propellants are normally distributed and have the same population variances. Two random samples of 26 specimens for each propellant are tested. The sample mean and standard deviation of burning rates for both propellants are summarised as in table below.

Construct a 90% confidence interval on the difference in population mean for the propellant 2
and propellant 1. Interpret the answer resulted

EXERCISE 2.3.1

Universiti Malaysia PAHANG

3. *The mean of sleep time for 50 IPTS students are 7 hours with standard deviation of 1 hour. The mean of sleep time for 60 IPTA students is 6 hours with standard deviation of 0.7 hour. Assume that the sleep time for the IPTS and IPTA students are normally distributed, find the 99% confidence interval for the difference population mean of sleep time between the IPTS and IPTA students. Interpret the answer resulted.*

   a) Assume the population variance are same

   b) Assume the population variance are different

4. Two groups of students are given a problem solving test, and the test marks are compared. The data are as follows:

Mathematics Majors	***Computer Science majors***
$n_1 = 15$	$n_2 = 14$
$\bar{x}_1 = 83.6$	$\bar{x}_2 = 79.2$
$s_1 = 3.3$	$s_2 = 2.8$

Find the 98% confidence interval for the difference population mean of test marks between the two groups of students. Assume the variance population test marks are not the same for both groups. Give the comment on the parameter estimate.

2.3.2 DEPENDENT SAMPLES

- This sub-topic considers estimation procedures for the mean difference of **two dependent samples i.e. paired data,** μ_D

- In paired data, each data point in one sample is matched to a unique data point in the second sample (matched couplings).

- An example of a paired data is a **pre-test/post-test** study design in which a random variable is measured **before and after** an intervention/treatment.

A $(1-\alpha)\,100\,\%$ Confidence Interval for μ_D

- σ_D^2 is known:

  $$\left(\bar{x}_D - z_{\alpha/2}\frac{\sigma_D}{\sqrt{n}},\ \bar{x}_D + z_{\alpha/2}\frac{\sigma_D}{\sqrt{n}}\right)$$

- σ_D^2 is unknown:
  - $n \geq 30$:

    $$\left(\bar{x}_D - z_{\alpha/2}\frac{s_D}{\sqrt{n}},\ \bar{x}_D + z_{\alpha/2}\frac{s_D}{\sqrt{n}}\right)$$

  - $n < 30$:

    $$\left(\bar{x}_D - t_{\alpha/2,\,\nu}\frac{s_D}{\sqrt{n}},\ \bar{x}_D + t_{\alpha/2,\,\nu}\frac{s_D}{\sqrt{n}}\right)$$

    where $\nu = n - 1$

where

$D = X_1 - X_2$ is the difference between two data sets

μ_D : a mean difference of the population, $-\infty < \mu_D < \infty$

$\bar{x}_D$: a sample mean of the differences

σ_D : a population standard deviation of the differences

s_D : a sample standard deviation of the differences

Example 2.8

A tyre manufacturer wishes to compare the made of tread wear of tyres between the new and conventional materials. One tyre of each type is placed on each front wheel of each of tenth front wheel drive vehicles. The choice as to which type of tyre goes on the right wheel and which goes on the left is made with the flip of a coin. Each car is driven for 40 000 km, then the tyres are removed, and the depth of the tread (in cm) on each is measured. The results were as follows.

Car	1	2	3	4	5	6	7	8	9	10
New material	4.35	5.00	4.21	5.03	5.71	4.61	4.70	6.03	3.80	4.70
Old material	4.19	4.62	4.04	4.72	5.52	4.26	4.27	6.24	3.46	4.50
Difference, D	0.16	0.38	0.17	0.31	0.19	0.35	0.43	-0.21	0.34	0.20

Calculate a 90% confidence interval for the population mean difference in tread wear between new and old materials in a way that takes advantage of the reduces variability produced by the paired design. Hence give the comment on your answer.

Solution:

X_N : The depth of the tread wear of tyres made by using new materials
X_O : The depth of the tread wear of tyres made by using old materials

$n = 10, \quad \bar{x}_D = 0.2320, \quad s_D = 0.1826 \ \rightarrow \sigma_D^2$ is unknown and $n < 30$.

$t_{0.10/2,\ 10-1} = t_{0.05,\ 9} = 1.8331$

A 90% confidence interval for μ_D

$$= \left(\bar{x} - t_{\alpha/2,\ n-1} \frac{s_D}{\sqrt{n}},\ \bar{x}_D + t_{\alpha/2,\ n-1} \frac{s_D}{\sqrt{n}} \right)$$

$$= \left(0.2320 - 1.8331 \left(\frac{0.1826}{\sqrt{10}} \right),\ 0.2320 + 1.8331 \left(\frac{0.1826}{\sqrt{10}} \right) \right)$$

$$= (0.1262,\ 0.3378)\text{cm}$$

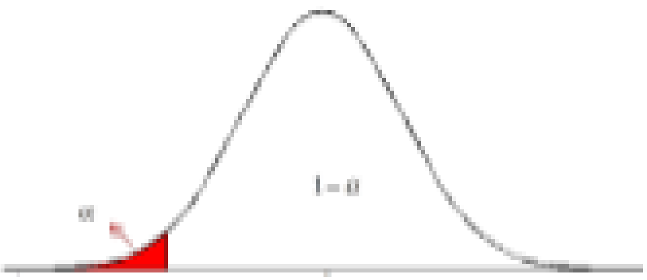
$$P(1.1262 < \mu_D < 0.3378) = 0.90$$

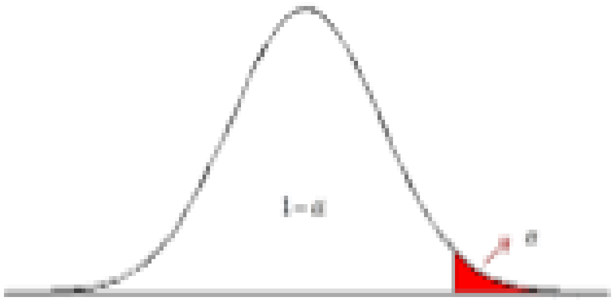
Interpretation: We are 90% confident that the population mean difference in tread wear between new and old materials lies within 0.1262 and 0.3378 cm.

One-Sided Lower Bound and One-Sided Upper Bound for μ_D

Universiti Malaysia PAHANG
Engineering • Technology • Creativity

One-Sided Lower Bound and One-Sided Upper Bound for μ_D





$$\mu_D > \bar{x}_D - z_\alpha \frac{\sigma_D}{\sqrt{n}}$$

one-sided lower bound

$$\mu_D \in (a, \infty)$$

$$\mu_D < \bar{x}_D + z_\alpha \frac{\sigma_D}{\sqrt{n}}$$

one-sided upper bound

$$\mu_D \in (-\infty, b)$$

where $a, b \in \mathbb{R}$

EXERCISE 2.3.2

Universiti Malaysia PAHANG
Engineering • Technology • Creativity

1. *Adding computerised medical images to a database promises to provide great resources for physicians. However, there are other methods of obtaining such information, so the issue of efficiency of access needs to be investigated. An experiment was conducted in which eight computer-proficient medical professionals were timed (in seconds) both while retrieving an image from a library of slides and while retrieving the same image from a computer database with a Web front end and the data are given as follows.*

Subject	1	2	3	4	5	6	7	8
Slide	30	35	40	25	20	30	35	62
Digital	25	16	15	15	10	20	17	16

Assume that the time retrieving the image is normally distributed, find the 95% confidence interval for the true mean difference between slide retrieval time and digital retrieval time. Interpret your result.

2. *A sample of 38 diesel lorries was run for both hot and cold engines. A study is conducted to estimate the difference in fuel economy. The mean difference of fuel mileage between hot and cold engines is 0.250 litre per km and the sample variance of the differences is 0.013 litre per km. Assume that the fuel mileage is normally distributed, find the 98% confidence interval for the population mean difference of fuel mileage between hot and cold engines. Give a comment on the parameter estimates.*

2.4 CONFIDENCE INTERVAL FOR POPULATION PROPORTION

- ✓ Estimate the confidence interval for a population proportion.
- ✓ Estimate the sample size by using the concept of confidence interval for population proportion.

A $(1-\alpha)$ 100% Confidence Interval for Population Proportion π

$$\left(p - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}},\ p + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right)$$

Number of successes in a given sample size

Sample proportion

$$p = \frac{x}{n} \quad \text{and} \quad q = 1 - p$$

Sample size

where $np \geq 5$ and $nq \geq 5$

Noted that $0 \leq \pi \leq 1$ and $p + q = 100\%$

Universiti Malaysia PAHANG
Engineering • Technology • Creativity

$$\textbf{Sample size, } n = \left(\frac{z_{\alpha/2}}{E} \right)^2 p(1-p)$$

where

$z_{\alpha/2}$: critical value

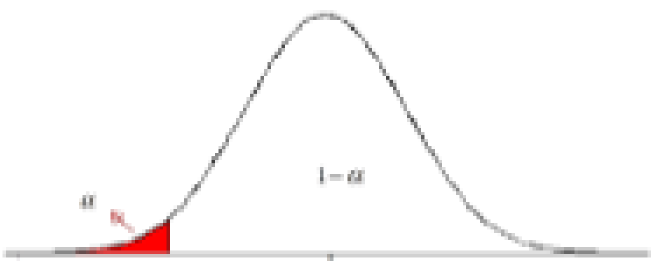
p : sample proportion

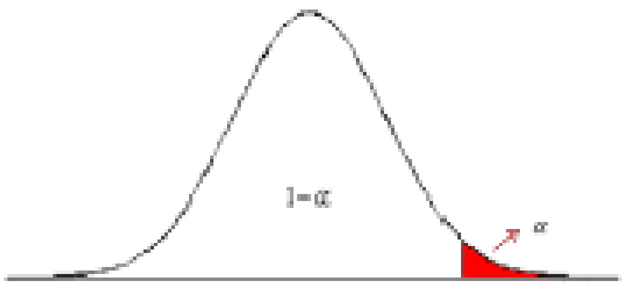
E : estimation error

Note that if the sample proportion is not given, the value of sample proportion is considered equal, that is $p = q = 0.5$.

One-Sided Lower Bound and One-Sided Upper Bound for π

One-Sided Lower Bound and One-Sided Upper Bound for π





$$\pi \geq p - z_{\alpha}\sqrt{\frac{p(1-p)}{n}}$$

one-sided lower bound

$$\pi \in (a, 1]$$

$$\pi \leq p + z_{\alpha}\sqrt{\frac{p(1-p)}{n}}$$

one-sided upper bound

$$\pi \in [0, b)$$

where $a, b \in [0,1]$

Example 2.9

The fraction of defective integrated circuits produced in a photolithography process is being studied. A random sample of 200 circuits is tested, revealing 13 defectives.

a) Calculate a 95% confidence interval on the fraction of defective circuits produced by this process and give a comment on the resulted confidence interval.

b) How large the sample would be if we wish to be at least 95% confident that the error in estimating p is less than 0.02?

c) Construct a one sided 95% lower bound confidence interval on the fraction of defective circuits produced by this process. Give an interpretation of the parameter estimate.

Solution:

a) X : The number of defective integrated circuits produced in a photolithography process

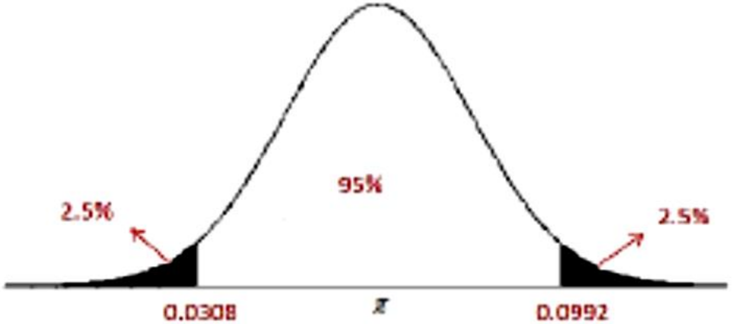
$$p=\frac{x}{n}=\frac{13}{200}=0.0650, \quad z_{\alpha/2}=z_{0.05/2}=1.9600$$

A 95% confidence interval for population proportion, π

$$=\left(p-z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}},\ p+z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right)$$

$$=\left((0.0650)-(1.9600)\sqrt{\frac{0.0650(1-0.0650)}{200}},\ (0.0650)+(1.9600)\sqrt{\frac{0.0650(1-0.0650)}{200}}\right)$$

$$=(0.0308,\ 0.0992)$$



$$P(0.0308<\pi<0.0992)=0.95$$

Interpretation: We are 95% confident that the fraction of the defective integrated circuits produced in a photolithography process is between 3.08% and 9.92%.

b) The estimates sample size,

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

$$\geq \left(\frac{1.9600}{0.02} \right)^2 (0.0650)(1-0.0650)$$

$$\geq 583.6831$$

$$\therefore n \approx 584 \text{ circuits}$$

c) $z_{\alpha} = z_{0.05} = 1.6449$

A 95% lower bound confidence interval for population proportion, π

$$= \left(p - z_{\alpha}\sqrt{\frac{p(1-p)}{n}},\ 1 \right]$$

$$= \left((0.0650) - (1.6449)\sqrt{\frac{0.0650(1-0.0650)}{200}},\ 1 \right]$$

$$= (0.0363, 1]$$

Interpretation: We are 95% confident that the fraction of the defective integrated circuits produced in a photolithography process that contain one-sided lower bound is between 3.63% to 100%.

■

EXERCISE 2.4

1. *The cure rate for the standard treatment of a disease is 45%. Dr Amani has introduced her new treatment which she claims is much better. She tested the new treatment on 50 patients with the disease and cured 30 of them.*
   a) *Find the sampling distribution of the cure rate of a disease.*
   b) *Construct a 99% confidence interval on the cure rate of a disease for the new treatment. Give a comment on the resulted confidence interval.*
   c) *If no estimation of the sample proportion available, how large is the samples of patient would be if we are 90% confident that the estimation error is 5%?*

2. *In a random sample of 85 automobile engine crankshaft bearings, 12% have a surface finish that meets the roughness specifications.*
   a) *Construct a 98% confidence interval of bearings that do not meet the roughness specification. Give an interpretation of the parameter estimate.*
   b) *How large the sample size would be if we wish to be at least 95% confident that the error in estimating proportion of the bearings in the population that exceeds the roughness specification is 0.02?*

2.5 CONFIDENCE INTERVAL FOR DIFFERENCE BETWEEN TWO POPULATION PROPORTIONS

- ✓ Estimate the confidence interval for the difference between two population proportions.

A $(1-\alpha)100\%$ Confidence Interval for $\pi_1 - \pi_2$

$$\left((p_1 - p_2) \pm z_{\alpha/2} \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \right)$$

Noted that $0 \le \pi_1 - \pi_2 \le 1$

Example 2.10

Two different types of injection-moulding machines are used to form plastic parts. A plastic part is considered defective if it has excessive shrinkage or is discoloured. Two random samples, each of size 300, are selected, and 15 defective parts are found in the sample from machine 1 while 8 defective parts are found in the sample from machine 2. Construct a 98% confidence interval on the difference in the two population proportions of the defective parts. Give an interpretation for your answer.

Solution:

X_1 : The number of defective plastic parts produced by machine 1
X_2 : The number of defective plastic parts produced by machine 2

Machine 1	Machine 2
$n_1 = 300$	$n_2 = 300$
$x_1 = 15$	$x_2 = 8$
$p_1 = \frac{15}{300} = 0.0500$	$p_2 = \frac{8}{300} = 0.0267$

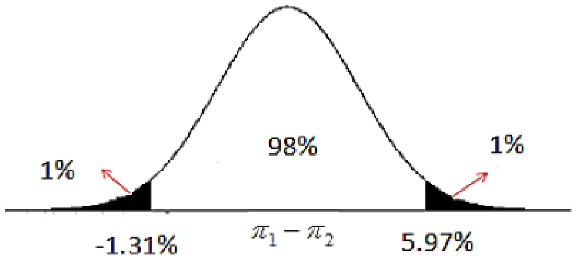
$$z_{\alpha/2} = z_{0.02/2} = 2.3263$$

A 98% confidence interval for difference between two population proportions, $\pi_1 - \pi_2$

$$= \left((p_1 - p_2) \pm z_{\alpha/2} \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \right)$$

$$= \left((0.0500 - 0.0267) \pm 2.3263 \sqrt{\frac{0.0500(0.9500)}{300} + \frac{0.0267(0.9733)}{300}} \right)$$

$$= (-0.0131,\ 0.0597)$$



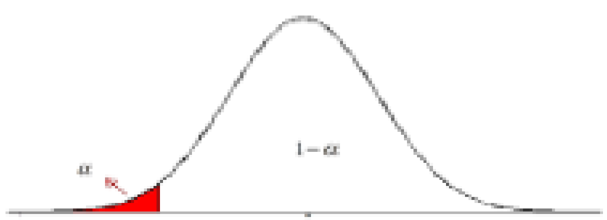
$$P(-1.31\% < \pi_1 - \pi_2 < 5.97\%) = 0.98$$

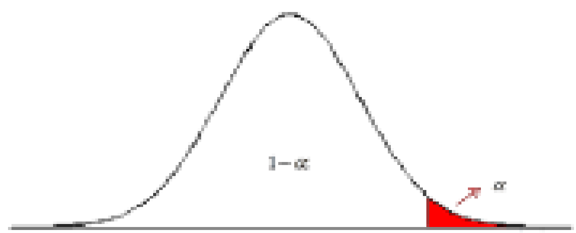
$\pi_1 < \pi_2$ $\quad$ $\pi_1 > \pi_2$

Interpretation: We are 98% confident that the difference in the two true proportions for the defective plastic parts produced by using machine 1 and machine 2 is between 1.31% more by machine 2 to 5.97% more by machine 1, compared to each other respectively.

One-Sided Lower Bound and One-Sided Upper Bound for $\pi_1 - \pi_2$

One-Sided Lower Bound and One-Sided Upper Bound for $\pi_1 - \pi_2$





$$\pi_1 - \pi_2 > (p_1 - p_2) - z_\alpha \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

one-sided lower bound

$\pi_1 - \pi_2 \in (a, 1]$

$$\pi_1 - \pi_2 < (p_1 - p_2) + z_\alpha \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

one-sided upper bound

$\pi_1 - \pi_2 \in [0, b)$

where $a, b \in [0,1]$

Note: Refer ME.8 for example.

EXERCISE 2.5

1. *An experiment was conducted to estimate the brace force for a compression web brace. In a sample of 380 short test columns and 394 long test columns, there were found that 304 and 360 of them have overestimated forces, respectively. Construct a 97% confidence interval for the difference of population proportions between the overestimated forces for long and short columns. Interpret your answer.*

2. *The specification for the pull strength of a wire that connects an integrated circuit to its frame is 10g or more. In a sample of 85 units made with gold wire, 68 units met the specification, while in a sample of 120 units made with aluminium wire, 105 units met specification. Find the 98% confidence interval for the difference between population proportion of circuit made with aluminium and gold wires that not met the specification. Give comment on the parameter estimate.*

3. *Two processes for manufacturing a certain microchip are being compared. A sample of 400 chips was selected from a less expensive process, and 16% were found to be defective. A sample of 100 chips was selected from a more expensive process, and 12% were found to be defective. Find a 90% confidence interval for the difference between the population proportions of defective chips produced by the two processes. Explain your answer.*

2.6 CONFIDENCE INTERVAL FOR POPULATION VARIANCE AND POPULATION STANDARD DEVIATION

✓ Estimate the confidence interval for a population variance and standard deviation.

A $(1-\alpha)100\%$ Confidence Interval for σ^2

$$\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2,\,\nu}},\ \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,\,\nu}} \right)$$

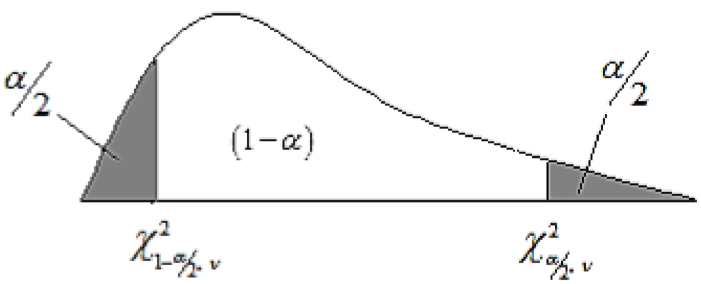
where $\nu = n-1$,

$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$ (Chi-square distribution),

and $\sigma^2 \geq 0$

A $(1-\alpha)100\%$ Confidence Interval for σ

$$\left(\sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2,\,\nu}}},\ \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,\,\nu}}} \right)$$



Example 2.11

A random sample of 10 rulers produced by a machine gives a set of data below (in cm).

100.131	**100.072**	**100.023**	**99.994**	**99.885**
100.146	**100.037**	**100.108**	**99.929**	**100.210**

Find the 95% confidence interval for the population variance and standard deviation of all the rulers produced by the machine. Give a comment on the 95% confidence interval for the population variance.

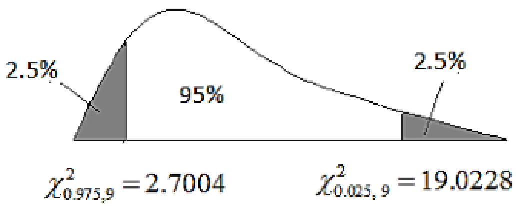
Solution:

X : The length of rulers produced by a machine

$n = 10, \quad s = 0.1006, \quad s^2 = 0.0101$

$\chi^2_{\alpha/2,\, n-1} = \chi^2_{0.05/2,\, 10-1} = 19.0228$

$\chi^2_{1-\alpha/2,\, n-1} = \chi^2_{1-0.05/2,\, 10-1} = 2.7004$



A 95% confidence interval for population variance, σ^2

$$= \left(\frac{(n-1)s^2}{\chi^2_{\alpha/2,\, n-1}},\ \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,\, n-1}} \right)$$

$$= \left(\frac{9\ (0.0101)}{19.0228},\ \frac{9\ (0.0101)}{2.7004} \right)$$

$$= (0.0048,\ 0.0337) \longrightarrow 0.0048 < \sigma^2 < 0.0337$$

Interpretation: We are 95% confident that the population variance of the length for the rulers produced by a machine lies within 0.0048 and 0.0337.

A 95% confidence interval for population standard deviation, σ

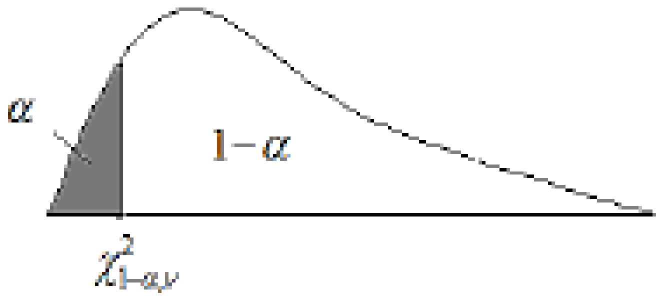
$$= \left(\sqrt{0.0048},\ \sqrt{0.0337} \right)$$

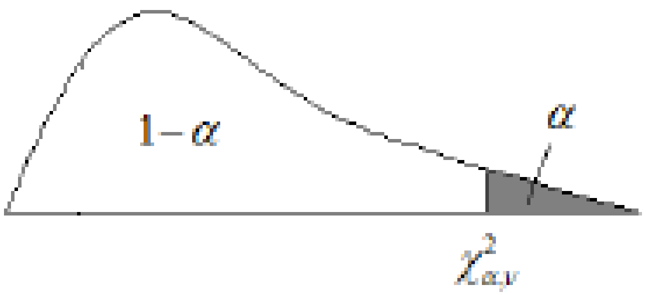
$$= (0.0693,\ 0.1836) \text{ cm} \longrightarrow 0.0693 < \sigma < 0.1836$$

Interpretation: We are 95% confident that the population standard deviation of the length for the rulers produced by a machine lies within 0.0693 cm and 0.1836 cm.

One-Sided Lower Bound and One-Sided Upper Bound for σ^2

One-Sided Lower Bound and One-Sided Upper Bound for σ^2





$$\sigma^2 > \frac{(n-1)s^2}{\chi^2_{\alpha,\nu}}$$

one-sided lower bound

$\sigma^2 \in (a, \infty)$

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

one-sided upper bound

$\sigma^2 \in [0, b)$

where $a, b \geq 0$

EXERCISE 2.6

Universiti Malaysia PAHANG
Engineering • Technology • Creativity

1. *A dairy processing company claims that the variance of the amount of fat in the whole milk processed by the company is 0.25. A random sample of 41 milk containers is selected and gives a variance of 0.27. Construct a 95% confidence interval on the population standard deviation of the amount of fat in the whole milk. Give a comment on the parameter estimate.*

2. *Ten independent measurements of the dissolution rate of a certain chemical are taken at a temperature of 10°C. The results (in °C) are as follows.*

   2.28 1.66 2.56 2.64 1.92 4.63 4.56 4.42 4.79 4.26

   Find the 98% confidence interval for the population variance of dissolution rate. Interpret your answer.

3. *Using data from question (2), find the 95% one-sided lower and upper bound for the population variance of the dissolution rate.*

2.7 CONFIDENCE INTERVAL FOR RATIO OF TWO POPULATION VARIANCES

- ✓ Estimate the confidence interval for the ratio of two population variances.

A $(1-\alpha)100\%$ Confidence Interval for $\left(\dfrac{\sigma_1^2}{\sigma_2^2}\right)$

$$\left(\frac{s_1^2}{s_2^2} f_{1-\alpha/2,\nu_2,\nu_1}, \frac{s_1^2}{s_2^2} f_{\alpha/2,\nu_2,\nu_1}\right)$$

or

$$\left(\frac{s_1^2}{s_2^2}\frac{1}{f_{\alpha/2,\nu_1=n_1-1,\ \nu_2=n_2-1}}, \frac{s_1^2}{s_2^2} f_{\alpha/2,\nu_2=n_2-1,\nu_1=n_1-1}\right)$$

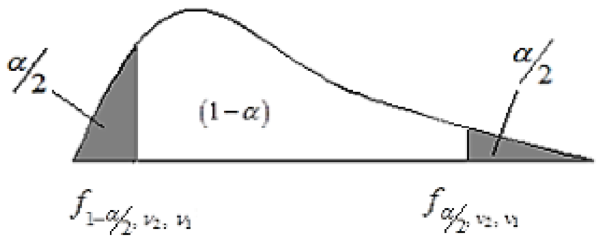
where

$$\nu_1 = n_1 - 1,\ \nu_2 = n_2 - 1$$

$$\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \sim F_{\nu_1,\nu_2} \text{ (} F \text{ distribution)}$$

$$f_{1-\alpha/2,\nu_1,\nu_2} = \frac{1}{f_{\alpha/2,\nu_2,\nu_1}}$$

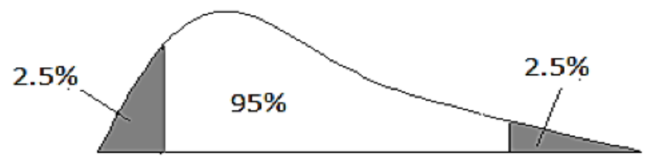
$$\sigma_1^2, \sigma_2^2 \geq 0$$



Note:

(i) Which degrees of freedom come first in the formula should be read as ν_1 in the statistical table.

(ii) The F-distribution of the lower and upper bounds of the confidence interval of two population variances is equal to $\dfrac{1}{f_{\alpha/2,\, \nu_1,\, \nu_2}}$ and $\dfrac{s_1^2}{s_2^2} f_{\alpha/2,\, \nu_2,\, \nu_1}$, respectively. This happened because of the changing inequalities when we derive the confidence interval for $\dfrac{\sigma_1^2}{\sigma_2^2}$ from the F sampling distributions to the confidence interval form.



$$f_{0.975,\,11,9} = \frac{1}{f_{0.025,\,9,\,11}} = \frac{1}{3.5879} = 0.2787 \qquad f_{0.025,\,11,\,9} = 3.9121$$

Figure 2.6: The Critical values of *F*-Distribution

Figure 2.6 illustrates how the critical value for both tailed of the F-distribution is identified using the statistical table. In statistical table, ν_1 is referred to the value of the first degrees of freedom, while ν_2 is referred to the value of the second degrees of freedom. Suppose the 95% confidence interval of $\dfrac{\sigma_1^2}{\sigma_2^2}$ is given by $\left(\dfrac{s_1^2}{s_2^2} f_{0.975,\,11,9},\ \dfrac{s_1^2}{s_2^2} f_{0.025,\,11,\,9} \right)$, then the critical values of the F-distribution is shown as in **Figure 2.6**.

Example 2.12

The machine in *Example 2.9* is serviced. A random sample of 12 rulers (in cm) produced by the machine after the service gives a set of data given as follows.

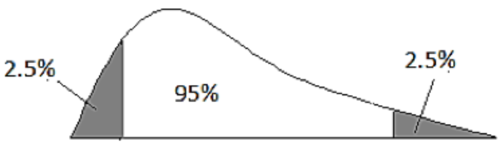
100.030	100.011	100.022	100.043	99.904	99.965
100.046	100.067	100.088	99.989	100.110	100.051

Find the 95% confidence interval for the ratio of population variances for all rulers produced by the machine before and after the service. Hence, interpret your answer.

Solution:

X_B : The length of rulers produced by a machine before the service
X_A : The length of rulers produced by a machine after the service

Before	After
$n_B = 10$	$n_A = 12$
$s_B = 0.1006$	$s_A = 0.0557$
$s_B^2 = 0.0101$	$s_A^2 = 0.0031$



$$f_{0.975,\,11,9} = \frac{1}{f_{0.025,\,9,\,11}} = \frac{1}{3.5879} = 0.2787$$

$$f_{0.025,\,11,\,9} = 3.9121$$

$$f_{\alpha/2,\,n_B-1,\,n_A-1} = f_{0.05/2,\,10-1,\,12-1} = f_{0.025,\,9,\,11} = 3.5879$$

$$f_{\alpha/2,\,n_A-1,\,n_B-1} = f_{0.05/2,\,12-1,\,10-1} = f_{0.025,\,11,\,9} = 3.9121$$

A 95% confidence interval on the ratio of two population variances, $\dfrac{\sigma_B^2}{\sigma_A^2}$

$$= \left(\frac{s_B^2}{s_A^2} f_{1-\alpha/2,\,n_A-1,\,n_B-1},\ \frac{s_B^2}{s_A^2} f_{\alpha/2,\,n_A-1,\,n_B-1} \right)$$

$$= \left(\frac{s_B^2}{s_A^2} \frac{1}{f_{\alpha/2,\,n_B-1,\,n_A-1}},\ \frac{s_B^2}{s_A^2} f_{\alpha/2,\,n_A-1,\,n_B-1} \right)$$

$$= \left(\frac{0.0101}{0.0031} \frac{1}{f_{0.025,\,9,\,11}},\ \frac{0.0101}{0.0031} f_{0.025,\,11,\,9} \right)$$

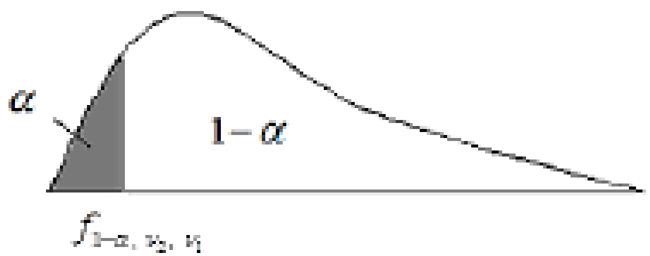
$$= \left(\frac{0.0101}{0.0031} \frac{1}{3.5879},\ \frac{0.0101}{0.0031}(3.9121) \right)$$

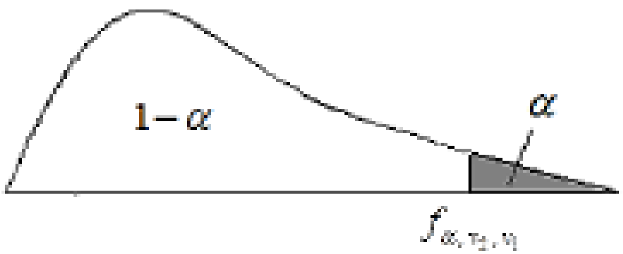
$$= (0.9081,\ 12.7459) \longrightarrow 0.9081 < \frac{\sigma_B^2}{\sigma_A^2} < 12.7459$$

Interpretation: We are 95% confident that the ratio of population variances for the length of rulers produced by a machine before and after service lies within 0.9081 and 12.7459.

One-Sided Lower Bound and One-Sided Upper Bound for σ_1^2 / σ_2^2

One-Sided Lower Bound and One-Sided Upper Bound for $\dfrac{\sigma_1^2}{\sigma_2^2}$





$$\frac{\sigma_1^2}{\sigma_2^2} > \frac{s_1^2}{s_2^2} f_{1-\alpha,\ \nu_2,\ \nu_1}$$

one-sided lower bound

$$\frac{\sigma_1^2}{\sigma_2^2} \in (a, \infty)$$

$$\frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} f_{\alpha,\ \nu_2,\ \nu_1}$$

one-sided upper bound

$$\frac{\sigma_1^2}{\sigma_2^2} \in (0, b)$$

where $a, b \geq 0$

EXERCISE 2.7

Universiti Malaysia PAHANG
Engineering • Technology • Creativity

1. *Two catalysts are analysed to determine how they affect the mean yield of a chemical process. Specifically, Catalysts A is currently in use, but Catalyst B is acceptable. The standard deviation of seven mean yields analysed by Catalyst A is 2.21 while eight mean yields analysed by Catalyst B has a standard deviation of 3.13. Construct a 90% confidence interval for the ratio of two population variances for all the mean yields produced by Catalyst A and Catalyst B. Give a comment on the parameter estimate.*
2. *Long-term exposure of textile workers to cotton dust released during processing can result in substantial health problems. A textile researcher conducts a study to investigate the methods to reduce risks while preserving important fabric properties. The accompanying data on roving cohesion strengths (kilo newton metre per kilogram (kN.m/kg)) for specimens produced at eight different twist multiples is drawn for two different levels of strength and the data are given as follows.*

Control	11.83	8.24	7.17	13.25	10.86	10.12	14.63	14.04
Heated	12.15	8.32	3.86	7.50	12.48	11.13	10.18	13.79

Construct a 96% confidence interval for the ratio of two population variances for all the roving cohesion produced by control and heated strength. Interpret your answer.

Do all ME and Tutorial questions.

"With determination, discipline and hard work all dreams become a reality." — **Lailah Gifty Akita**