Recall that:

We use point estimate to estimate the population parameter using a one value called: (Sample Statistic).

Now, we use an interval to estimate the population parameter. This interval is called: (Confidence interval).

Confidence intervals= point estimate $\mp$ margin of error

To construct our confidence intervals: starts by determine the level of significance (α).

Then, we say that: we trust (insure) that our confidence interval <u>includes</u> the <u>value of the true parameter</u> by the percent $(1-\alpha)100\%$.

If we assume ($\alpha = 0.05$), then we insure that our confidence interval <u>includes</u> the <u>value of the true parameter</u> by the percent $(1-0.05)100\%$ = 95%. Thus: here 95% is the confidence level.

And the value 0.95 is the confidence coefficient.

Note: Assuming the confidence interval (Lowe, Upper) = (L,U). The length of this interval increases as confidence level increases.

Note: The **Confidence Interval Range** is a **twice** of (**margin of error**).

* Confidence intervals for population mean (μ) when:

1.The sample size is small ($n < 30$) Or large ($n \geq 30$) and population variance is known. (use Z-tables)

$\overline{x} \mp$ margin of error : $$\overline{x} \mp Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$, $Z_{\frac{\alpha}{2}}$ value providing an area of α/2 in the upper tail of the Z distributio

2.The sample size is large ($n \geq 30$) and population variance is Unknown. (use Z-tables)

$\overline{x} \mp$ margin of error : $$\overline{x} \mp Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$,

3.The sample size is small ($n < 30$) and population variance is Unknown. (use t-tables)

$\overline{x} \mp$ margin of error : $$\overline{x} \mp t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$, $t_{\frac{\alpha}{2}}$ value providing an area of α/2 in the upper tail of the t distributio with (n − 1) degrees of freedom

* Confidence intervals for population proportion (p) :

$\overline{p} \mp$ margin of error : $$\overline{p} \mp Z_{\frac{\alpha}{2}} \sqrt{\frac{\overline{p}(1-\overline{p})}{n}}$$

Example:

Assuming a sample of 16 students, if the sample mean of student marks is 75.

Assuming the population variance is 64

Obtain a 95% confidence interval of the population mean (μ):

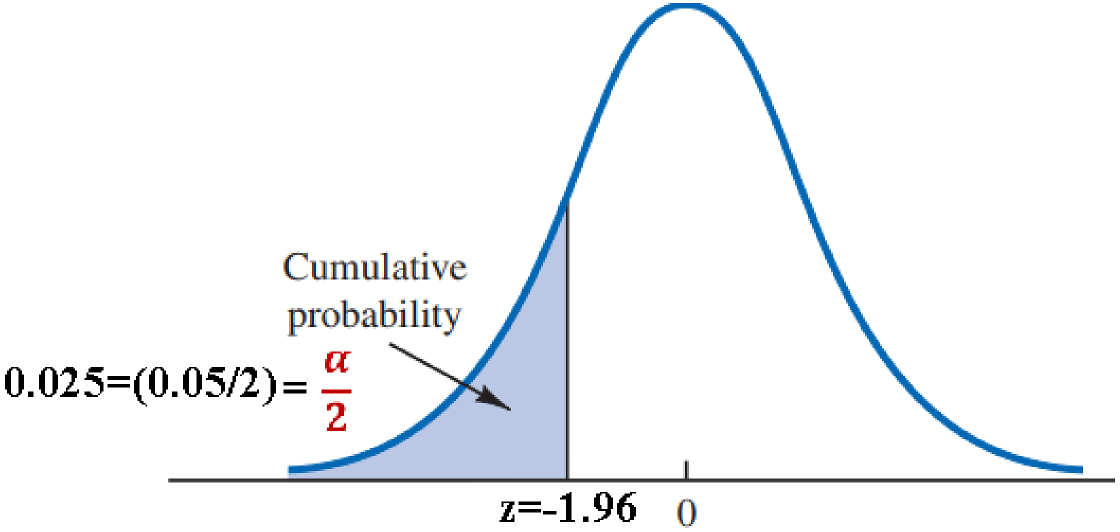
$$\bar{x} \mp Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$$75 \mp Z_{\frac{0.05}{2}} \frac{\sqrt{64}}{\sqrt{16}}$$

$$75 \mp Z_{0.025} \frac{8}{4} \quad , \qquad Z_{0.025} = 1.96$$

$$75 \mp (1.96)(2)$$

$$75 \mp 3.92 \qquad , \text{ margin of error} = 3.92$$



We trust by 95% that $\mu \in (71.08\ ,\ 78.92)$

95% confidence interval of μ is $(71.08\ ,\ 78.92)$ $\rightarrow$ $71.08 < \mu < 78.92$

Example:

Assuming a sample of 100 students, if the sample mean of student marks is 75.

Assuming the sample variance is 64

Obtain a 95% confidence interval of the population mean:

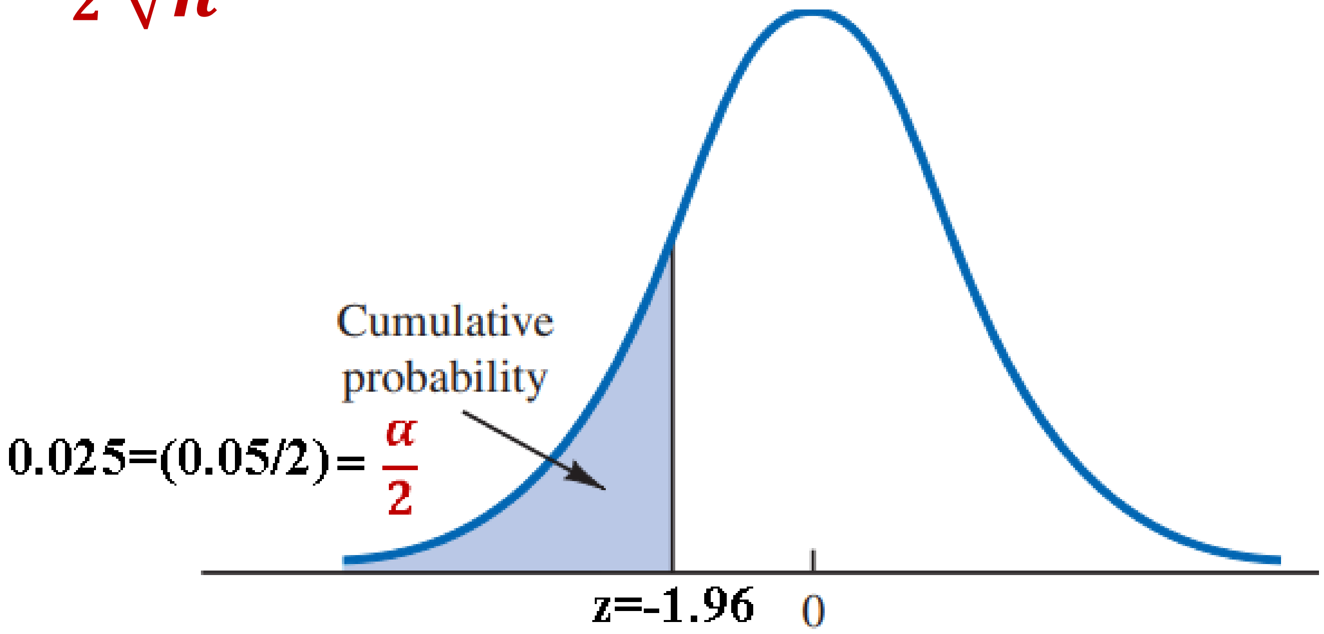
$$\bar{x} \mp Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

$$75 \mp Z_{\frac{0.05}{2}} \frac{\sqrt{64}}{\sqrt{100}}$$

$$75 \mp Z_{0.025} \frac{8}{10} \quad , \qquad Z_{0.025} = 1.96$$

$$75 \mp (1.96)(0.8)$$

$$75 \mp 1.568$$



We trust by 95% that $\mu \in (73.432 , 76.568)$

95% confidence interval of μ is $(73.432 , 76.568)$, margin of error = 1.568

Note:

Assuming $(1-\alpha)100\%$ confidence interval of the population **mean** $(\boldsymbol{\mu})$: $\bar{x} \mp Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$

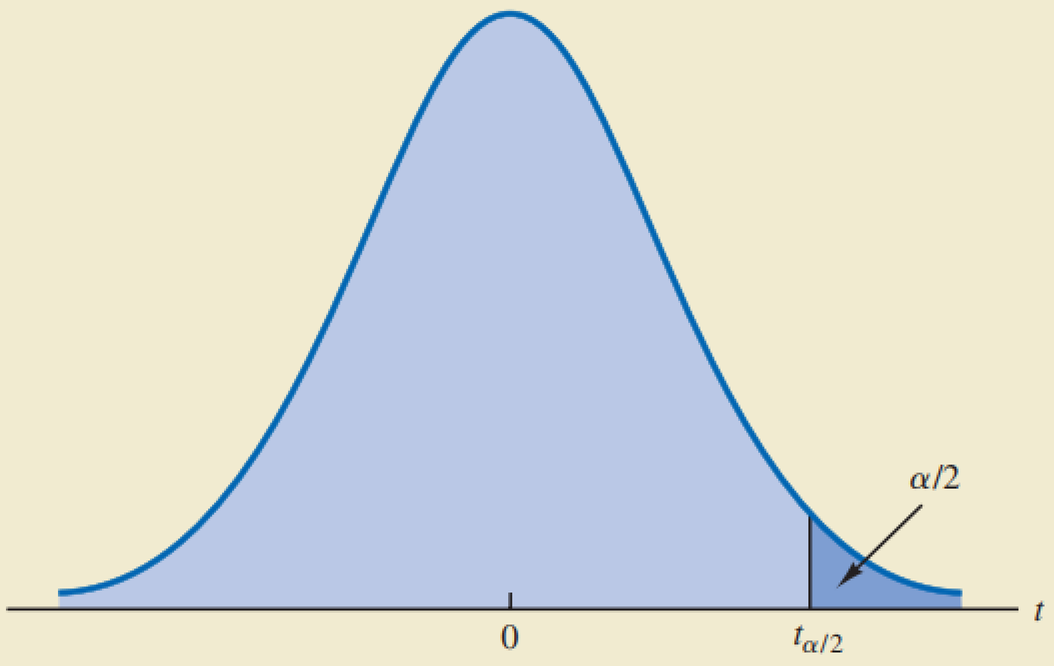
Then, the margin of error is $Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$,

margin of error increases as (n) decreases.

margin of error increases as (S) increases. And/or $Z_{\frac{\alpha}{2}}$ increases.

Where $Z_{\frac{\alpha}{2}}$ increases if confident level ($(1-\alpha)100\%$) increases i. e. when (α) decreases

FIGURE 8.5 t DISTRIBUTION WITH $\alpha/2$ AREA OR PROBABILITY IN THE UPPER TAIL



Example:

Assuming a sample of 16 students, if the sample mean of student marks is 75.

Assuming the sample variance is 64

Obtain a 95% confidence interval of the population mean:

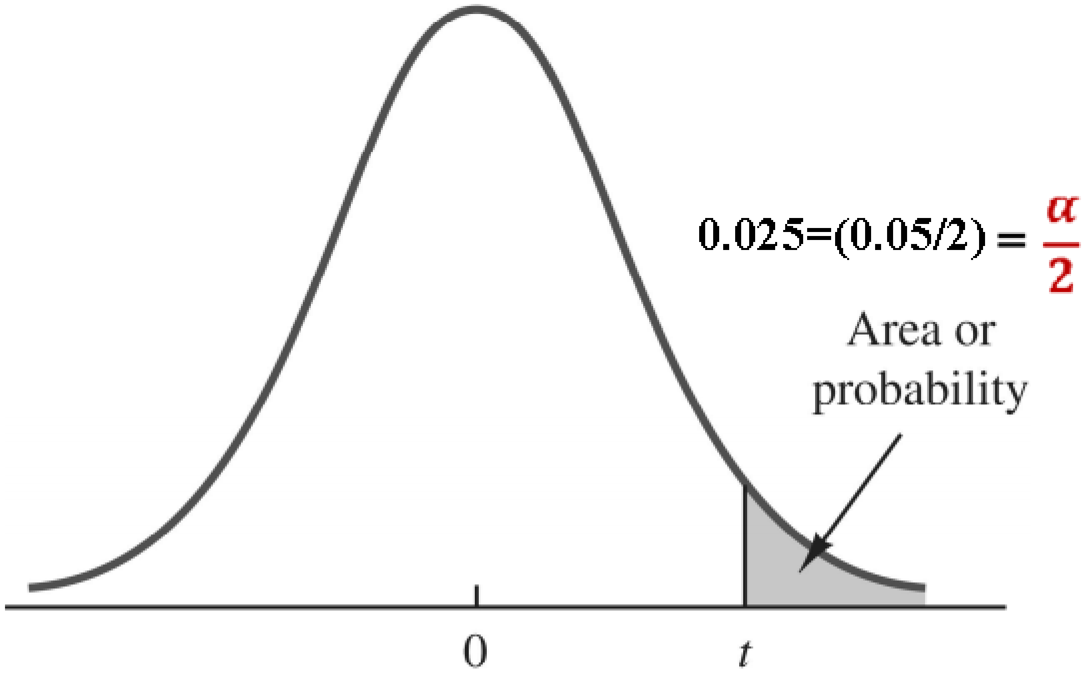
$$\bar{x} \mp t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

$$75 \mp t_{\frac{0.05}{2}} \frac{\sqrt{64}}{\sqrt{16}}$$

$$75 \mp t_{0.025} \frac{8}{4} \quad , \qquad t_{0.025\ ,15\ d.f.} = 2.131$$

$$75 \mp (2.131)(2)$$

$$75 \mp 4.262$$



We trust by 95% that $\mu \in (70.738\ ,\ 79.262)$

95% confidence interval of μ is $(70.738\ ,\ 79.262)$, margin of error = 4.262

Example:

Assuming a sample of 30 students, if the proportion of failed student is 0.6.

Obtain a 90% confidence interval of the population proportion:

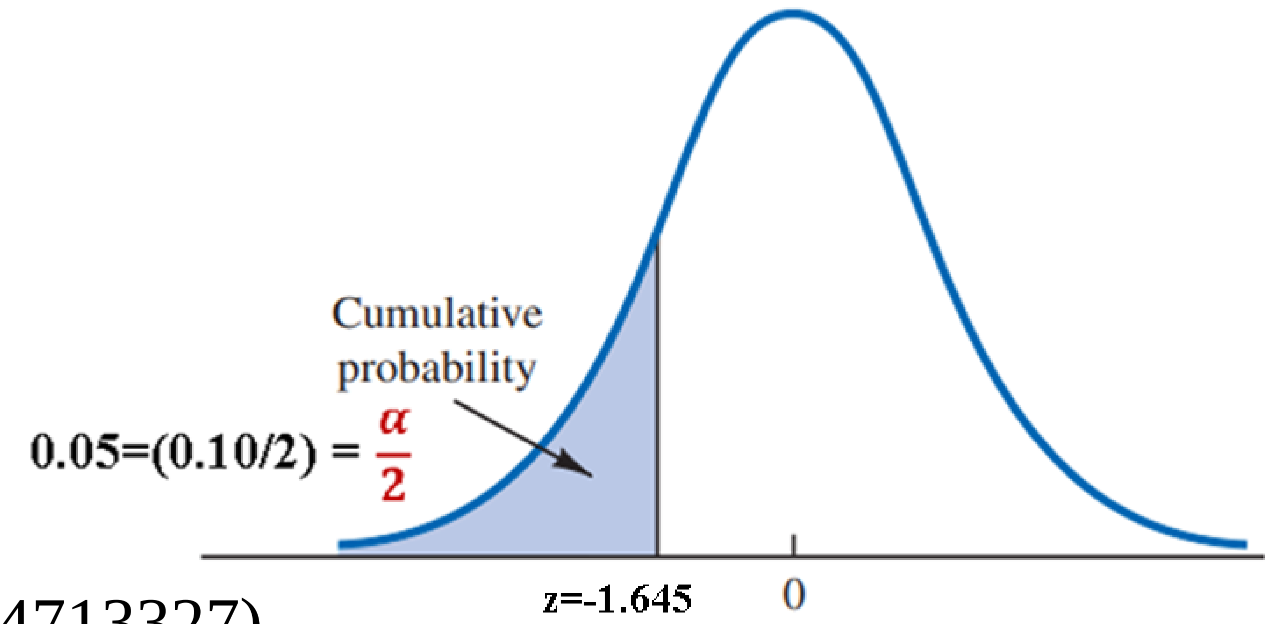
$$\bar{p} \mp Z_{\frac{\alpha}{2}} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

$$0.6 \mp Z_{\frac{0.10}{2}} \sqrt{\frac{0.6(1-0.6)}{30}}$$

$$0.6 \mp Z_{0.05}\sqrt{0.008} \quad , \quad Z_{0.05} = 1.645$$

$$0.6 \mp (1.645)\sqrt{0.008}$$

$$0.6 \mp 0.14713327$$



We trust by 90% that $p \in (0.4528667 , 0.74713327)$

90% confidence interval of p is $(0.4528667 , 0.74713327)$, margin of error = 0.14713327

Note:

Assuming $(1-\alpha)100\%$ confidence interval of the population proportion (P): $\bar{p} \mp Z_{\frac{\alpha}{2}}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$

Then, the margin of error is $Z_{\frac{\alpha}{2}}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$,

margin of error increases as (n) decreases.

margin of error increases as $(\bar{p})$ near 0.5. And/or $Z_{\frac{\alpha}{2}}$ increases.

Where $Z_{\frac{\alpha}{2}}$ increases if confident level $((1-\alpha)100\%)$ increases i. e. when (α) decreases

90% confidence interval of p is (0.4528667 , 0.74713327)

The **confidence interval Range =**

= Upper-Lower= 0.74713327 - 0.4528667 = 0.29426657 =(2* margin of error)

Note: The **Confidence Interval Range** is a **twice** of (**margin of error**).

The confidence interval Range = Upper - Lower = (2* margin of error)

Example:

Assuming 95% confidence interval (C.I.) of the population mean of 100 students

is [62,84], find:

a) **margin of error = (C.I. Range) / 2 = (84 – 62) / 2 = 11**

b) **Sample mean ($\overline{x} \mp$ margin of error)**

$\overline{x}$ = (Upper - marginal error) = (84 - 11) = 73 OR $\overline{x}$ = (Lower + marginal error) = (62 + 11) = 73

73 is the sample mean

c) **Population variance** $\overline{x} \mp Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$, the marginal error $= Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 11$

$$Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 1.96 \frac{\sigma}{\sqrt{100}} = 11 \ \ggg \sigma = 56.12 \gg variance = \sigma^2 = 3149.59$$

Note:

The most Z values for Confidence Intervals:

from Normal Table

Assuming 95% confidence interval (C.I.) >>>> $Z_{\frac{0.05}{2}} = 1.96$

Assuming 90% confidence interval (C.I.) >>>> $Z_{\frac{0.10}{2}} = 1.645$

Assuming 85% confidence interval (C.I.) >>>> $Z_{\frac{0.15}{2}} = 1.44$

Assuming 80% confidence interval (C.I.) >>>> $Z_{\frac{0.20}{2}} = 1.28$

Determining the Sample Size

A Case Based on Confidence intervals for population mean (μ):

point estimate $\mp$ margin of error

$$\bar{x} \mp \text{margin of error}$$

We can use the above formula to make an estimate value of the <u>sample size</u> (n) that provide a given determined [margin of error]:

we use the confidence interval formula: $\bar{x} \mp Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$.

Then, assuming the [margin of error] is $Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

Where: $Z_{\frac{\alpha}{2}}$ **is determined based on how you need a confidence level for your estimator.**

The value of the population standard deviation (σ) is determined using several ways:

The value of the population standard deviation (σ) is determined using several ways:

A common method is to replace the population standard deviation **(σ)** by its estimator from a sample(use **(S)**).

If the sample standard deviation **(S)** is unknown, we can use equation (8.3) provided we have a preliminary or planning value for **(S)** .

$$\text{margin of error} = \mathrm{E} = Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \rightarrow Then\ E = Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

In practice, one of the following procedures can be chosen to determine (*S*) :

1. Use the estimate of the population standard deviation computed from data of **previous studies** as the planning value for **(*S*).**
2. Use a pilot study to select a **preliminary sample**. The sample standard deviation from the preliminary sample can be used as the planning value for **(*S*)**.
3. Use **judgment** or a "**best guess**" for the value of **(*S*)**. For example, we might begin by estimating the largest and smallest data values in the population.

   The difference between the largest and smallest values provides an estimate of the **range** for the data. Finally, the [range divided by 4] is often suggested as a rough approximation of the standard deviation **(*S*)**. and thus an acceptable planning value for **(*S*)**.

Then the formula $E = Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$ will be used to estimate **(*n*)** **at least**: $n \geq \left(\frac{Z_{\alpha/2} \; S}{E} \right)^2$

Ex. if $\left(\frac{Z_{\alpha/2} \; S}{E} \right)^2$ **=57.2 → Then → We need at least a sample size of 58**

Note: $n \geq \left(\dfrac{Z_{\alpha/2} \quad S}{E} \right)^2$

Then, the sample size is increases if $Z_{\frac{\alpha}{2}}$ increases.

Where $Z_{\frac{\alpha}{2}}$ increases if confident level ($(1-\alpha)100\%$) increases i. e. when (α) decreases

Example:

Suppose that we wanted to estimate the true **average** number of salmon fish eggs lays with 95% confidence.

The margin of error we are willing to accept is (0.7).

Suppose we also know that the standard deviation (S) from a previous sample equals to (20).

What is sample size should we use?

$$n \geq \left(\frac{Z_{\alpha/2} \quad S}{E} \right)^2$$

$$n \geq \left(\frac{1.96 \; * \; 20}{0.7} \right)^2 = (\; 56 \;)^2 = 3136 \;\; \textbf{is the sample size}$$

Exercise: solve the above example based on 80% confidence.

Determining the Sample Size

A Case Based on Confidence intervals for population proportion (P):

point estimate $\mp$ margin of error

$\bar{\mathrm{p}} \mp$ margin of error

We can use the above formula to make an estimate value of the <u>sample size</u> (n) that provide a given determined [margin of error]:

we use the confidence interval formula: $\bar{p} \mp Z_{\frac{\alpha}{2}}\sqrt{\frac{P(1-P)}{n}}.$

Then, assuming the [margin of error] is $Z_{\frac{\alpha}{2}}\sqrt{\frac{P(1-P)}{n}}$

Where: $Z_{\frac{\alpha}{2}}$ **is determined based on how you need a confidence level for your estimator.**

The value of the population proportion (P) is estimated by a planning value (P^{*}) is determined using several ways:

In practice, one of the following procedures can be chosen to determine a planning value (P^*) for population proportion:

1. Use the sample proportion from a **previous sample** of the same or similar units.
2. Use a **pilot study** to select a **preliminary sample**. The sample proportion from this sample can be used as the planning value, **(P^*)**.
3. Use **judgment** or a "**best guess**" for the value of **(P^*)**.
4. If none of the preceding alternatives apply, use a planning value of **(P^* = 0.50)**.

$Then\ the\ formula\ E = Z_{\frac{\alpha}{2}}\sqrt{\frac{P^*(1-P^*)}{n}}$ will be used to estimate **(n)** **at least** :

$$n \geq \frac{\left(Z_{\alpha/2}\right)^2 \ (P^*)(1-P^*)}{(\ E\ )^2}$$

The largest value of $(P^*)(1 - P^*)$ **is at** **the planning value** $(P^* = 0.5)$

TABLE 8.5 SOME POSSIBLE VALUES FOR $p^*(1 - p^*)$

p^*	$p^*(1 - p^*)$	
.10	$(.10)(.90) = .09$	
.30	$(.30)(.70) = .21$	
.40	$(.40)(.60) = .24$	
.50	$(.50)(.50) = .25$	← Largest value for $p^*(1 - p^*)$
.60	$(.60)(.40) = .24$	
.70	$(.70)(.30) = .21$	
.90	$(.90)(.10) = .09$	

Example:

Assuming that a random sample of 60 students from PTUK university has been given a qualified exam, and 48 of these students had success in the exam by levels over than 70.

Construct a 95% confidence interval for the population proportion of students in PTUK who passed the qualifying exam with graded over than 70. $\overline{p} = \frac{48}{60} = 0.8$

$$\overline{p} \mp Z_{\frac{\alpha}{2}} \sqrt{\frac{\overline{p}(1-\overline{p})}{n}}$$

$$0.8 \mp Z_{\frac{0.05}{2}} \sqrt{\frac{0.8(1-0.8)}{60}}$$

$$0.8 \mp Z_{0.025}\sqrt{0.00266667} \quad , \quad Z_{0.025} = 1.96$$

$$0.8 \mp (1.96)(0.0516398)$$

$$0.8 \mp 0.101214$$

We trust by 95% that $p \in (0.698786035 , 0.901213965)$

Example:

Assuming that a random sample of 100 students from PTUK university has been given a qualified exam, and 78 of these students had success in the exam by levels over than 70.

Construct a 80% confidence interval for the population proportion of students in PTUK who passed the qualifying exam with graded over than 70. $\bar{p} = \frac{78}{100} = 0.78$

$$\bar{p} \mp Z_{\frac{\alpha}{2}} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

$0.78 \mp Z_{\frac{0.20}{2}} \sqrt{\frac{0.78(1-0.78)}{100}}$

$0.78 \mp Z_{0.10} \sqrt{0.001716}$, $\quad Z_{0.025} = 1.96$

$0.78 \mp (1.28)(0.041425)$

0.78 ∓ 0.05302

We trust by 80% that $p \in (0.72698, 0.83302)$

Example:

Assuming that you want to estimate the population proportion of defect devices. Assuming a pilot study of 38 devices that contains 5 defected devices. Find the sample size that will be used to estimate the proportion of defected devices such that we confidence it by 90% with margin of error 0.206.

Solve:

$$n \geq \frac{\left(Z_{\alpha/2}\right)^2 \ (P^*)(1 - P^*)}{(\ E\ )^2}$$

$$n \geq \frac{\left(Z_{0.10/2}\right)^2 \ \left(\frac{5}{38}\right)\left(1 - \frac{5}{38}\right)}{(\ 0.206\ )^2}$$

$$n \geq \frac{(1.645)^2 \ \left(\frac{5}{38}\right)\left(\frac{33}{38}\right)}{(\ 0.206\ )^2}$$

$n \geq 7.28$ **>>> at least we need 8**