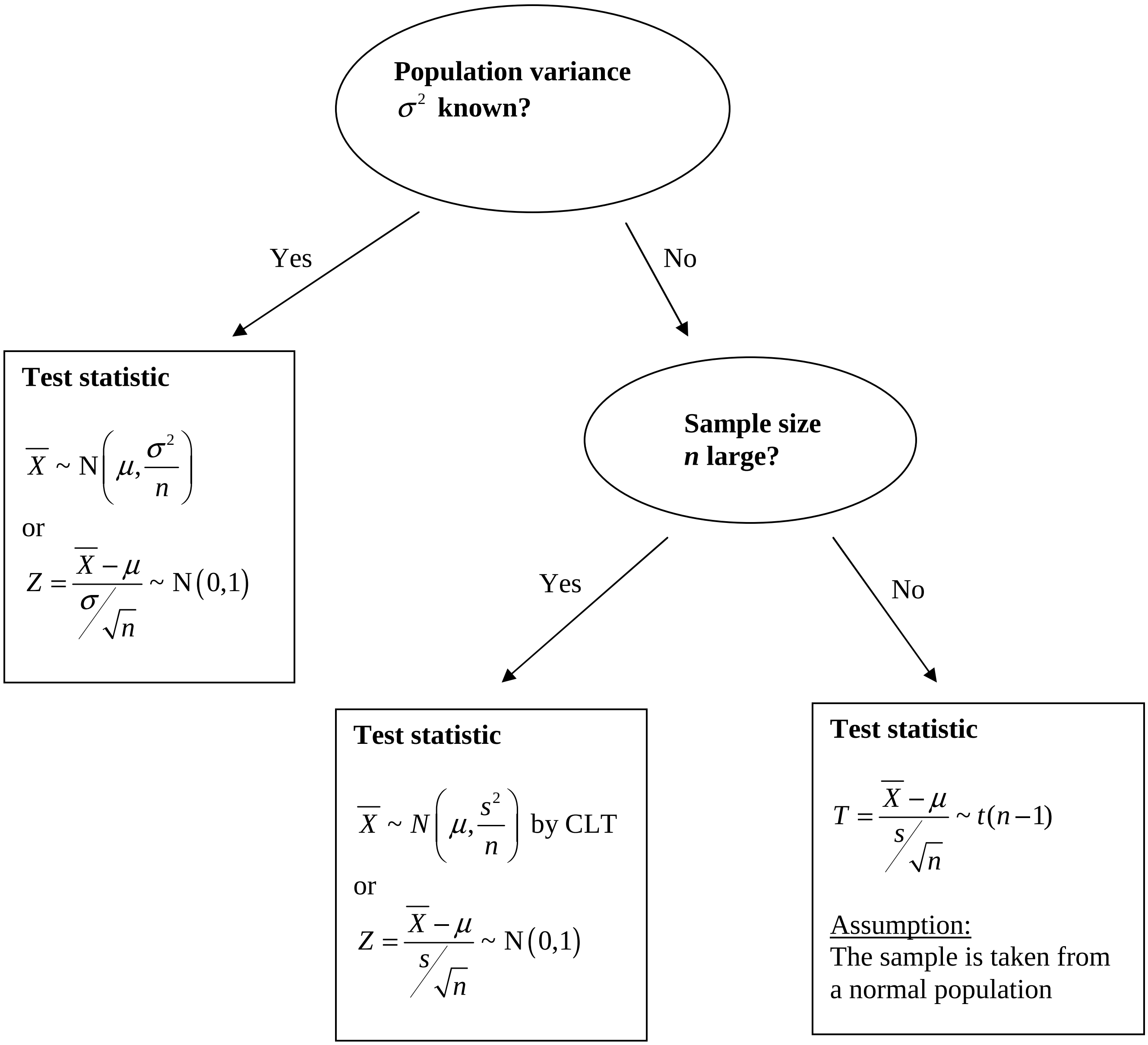
## A Summary on Hypothesis Testing (H2 Maths Paper 2)

**This set of well-written summary notes has been contributed by Mr Lau Haw Ping who lectures Mathematics at Anderson Junior College. Thank you very much!**

## (I) Hypothesis testing – flowchart on which test statistic to apply



**Remember:**

1) $s^2 = \frac{n}{n-1}$(sample variance)

2) For coding, $u = x - 200$, $\bar{x} = \bar{u} + 200$, $s_x^2 = s_u^2$ and sample var ($x$) = sample var ($u$).

## (II) Carrying out hypothesis testing involving unknowns

**<u>Case 1</u>: When *n* is unknown**

$H_0 : \mu = 5$

$H_1 : \mu < 5$

Given: $\bar{x} = 6, \quad \sigma = 1.2, \quad n = ?$ (unknown)

Given that $H_0$ is rejected at 5%, find the range of $n$.

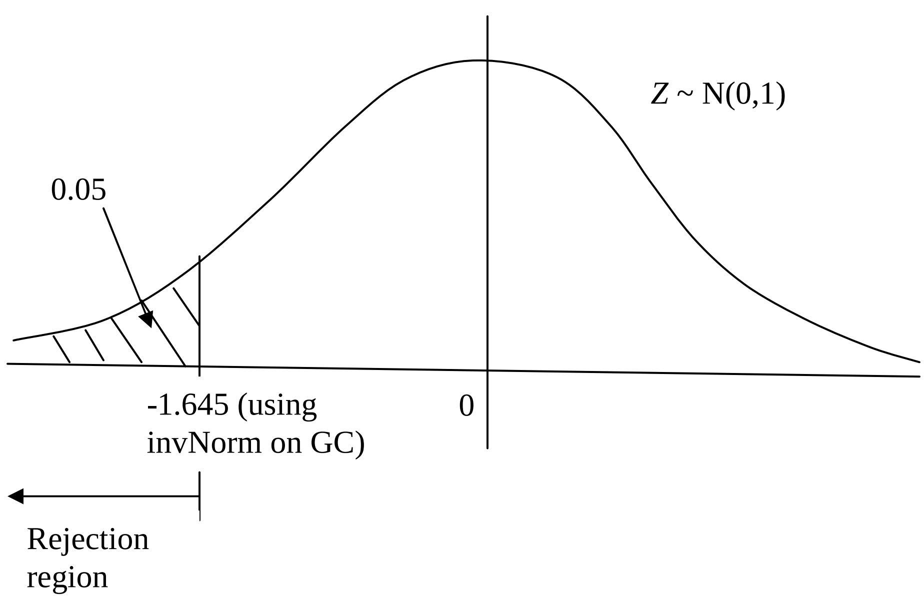
---

<u>Solution:</u>

Test statistic $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ i.e. $\bar{X} \sim N\left(5, \frac{(1.2)^2}{n}\right)$

The moment you have unknown in the parameters, you **MUST** seek the help of Standard Normal Distribution $Z \sim N(0,1)$.


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



To reject $\mathrm{H}_0$, $z < -1.645$

i.e. $\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}} < -1.645 \Rightarrow \dfrac{6-5}{1.2/\sqrt{n}} < -1.645 \Rightarrow$ ..............etc.

**<u>Case 2</u>: When $\mu_0$ is unknown (*z*-test)**

$\mathrm{H}_0 : \mu = \mu_0$
$\mathrm{H}_1 : \mu < \mu_0$

Given: $\bar{x} = 6$, $s = 1.2$ ($\sigma$ unknown), $n = 60$

Given that $\mathrm{H}_0$ is rejected at 5%, find the range of $\mu_0$.

---

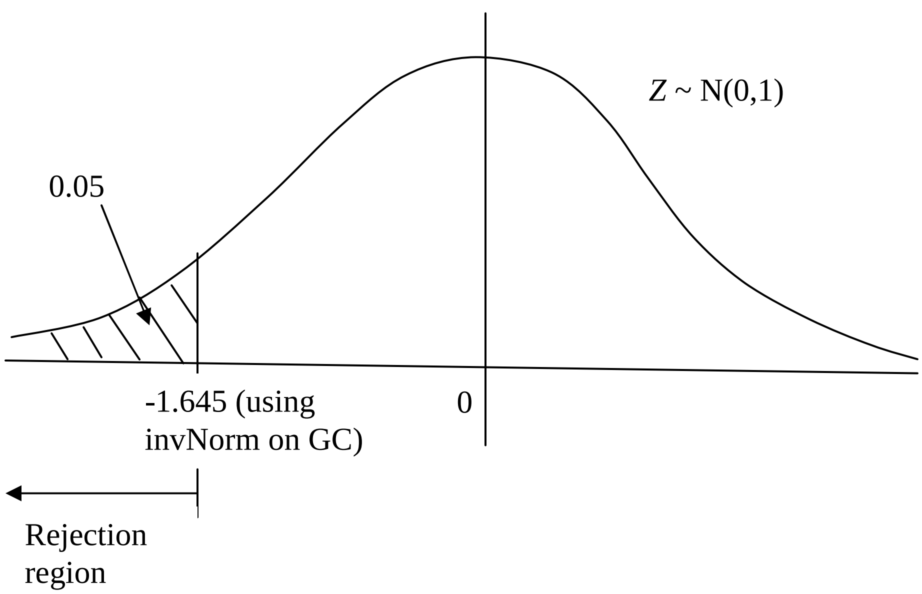
<u>Solution:</u>

Test statistic $\bar{X} \sim \mathrm{N}\left(\mu, \dfrac{s^2}{n}\right)$ i.e. $\bar{X} \sim \mathrm{N}\left(\mu_0, \dfrac{(1.2)^2}{60}\right)$

The moment you have unknown in the parameters, you **MUST** seek the help of Standard Normal Distribution $Z \sim \mathrm{N}(0,1)$.

HELP!!

$$Z = \frac{\bar{X}-\mu_0}{s/\sqrt{n}} \sim \mathrm{N}(0,1)$$



To reject $\mathrm{H}_0$, $z < -1.645$

i.e. $\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}} < -1.645 \Rightarrow \dfrac{6-\mu_0}{1.2/\sqrt{60}} < -1.645 \Rightarrow$ ……………etc.

**<u>Case 3</u>: When $\bar{X}$ is unknown**

$\mathrm{H}_0 : \mu = 5$
$\mathrm{H}_1 : \mu < 5$

Given: $\bar{x} = ?$ (unknown), $\sigma = 1.2$, $n = 60$

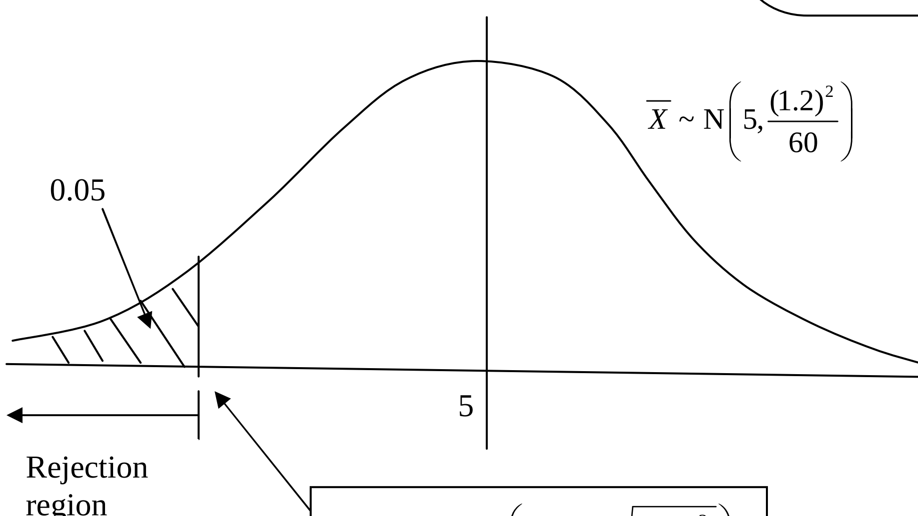
Given that $\mathrm{H}_0$ is rejected at 5%, find the range of $\bar{x}$.

---

<u>Solution:</u>

Test statistic $\bar{X} \sim \mathrm{N}\left(\mu, \dfrac{\sigma^2}{n}\right)$ i.e. $\bar{X} \sim \mathrm{N}\left(5, \dfrac{(1.2)^2}{60}\right)$

When parameters are known, you have the option of using $\bar{X} \sim \mathrm{N}\left(\mu, \dfrac{\sigma^2}{n}\right)$ (easier) or $Z \sim \mathrm{N}(0,1)$.

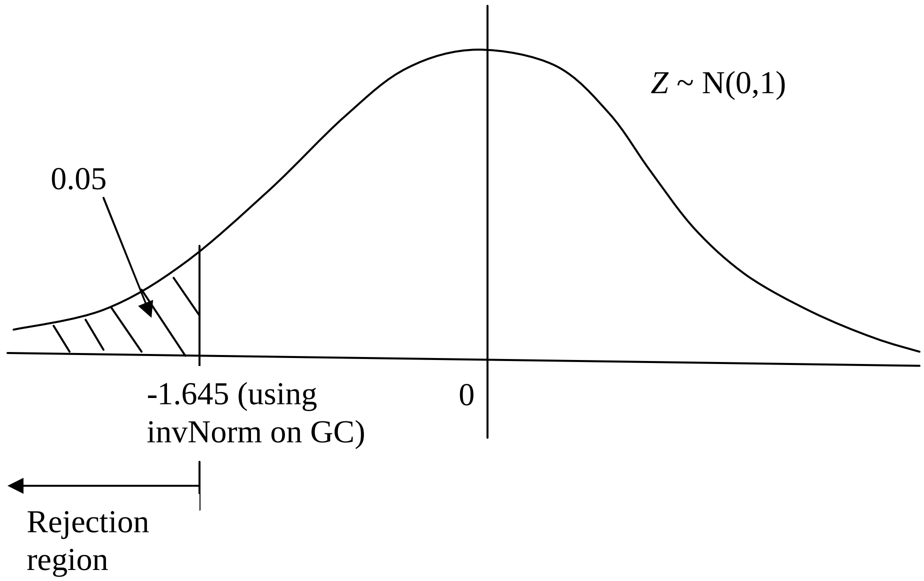



Use invNorm$\left(0.05, 5, \sqrt{\dfrac{(1.2)^2}{60}}\right)$ to find this value (i.e. 4.74518)

To reject $H_0$, $\bar{x} < 4.75$ (answer)

Alternatively,

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




To reject $H_0$, $z < -1.645$

ie $\frac{\overline{x} - \mu}{\sigma / \sqrt{n}} < -1.645 \Rightarrow \frac{\overline{x} - 5}{1.2 / \sqrt{60}} < -1.645 \Rightarrow$ ……………etc. (same answer)

**<u>Case 4</u>: When $\mu_0$ is unknown (*t*-test)**

$H_0 : \mu = \mu_0$
$H_1 : \mu < \mu_0$

Given: $\bar{x} = 6$, $s = 1.2$ ($\sigma$ unknown), $n = 12$
Given that $H_0$ is rejected at 5%, find the range of $\mu_0$.

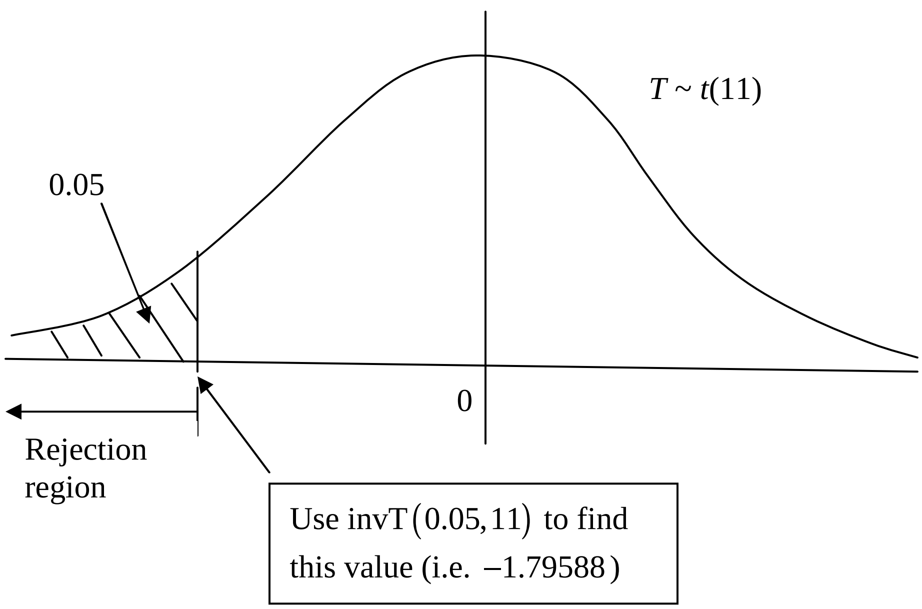
---

<u>Solution:</u>

Test statistic $T = \dfrac{\bar{X} - \mu}{s/\sqrt{n}} \sim t(n-1)$

> **MUST** use *t*-test when $\sigma$ is unknown and *n* is small.

$$T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim t(n-1)$$




To reject $H_0$, $t < -1.79588$

i.e. $\dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} < -1.79588 \Rightarrow \dfrac{6 - \mu_0}{1.2/\sqrt{12}} < -1.79588 \Rightarrow$ ……………etc.

**Note**: For the student whose GC does not have invT, you will need to learn how to read the invT value from the *t*-table in the formulae booklet. For our example, we start with the row with $\nu = 11$ and then followed by the column with $p = 0.95$. The value obtained is 1.796, but we need to include a negative sign because the *t*-value is on the left half of the *t*-distribution curve.

## (III) Exercises with worked solutions

**Q1** A random sample of 50 packets of the snack 'Yummi Crunch' is weighed and the mass *x* in grams is recorded. The results are summarised as follows:

$$\Sigma(x - 150) = -80, \ \Sigma(x - 150)^2 = 2500.$$

A test was carried out at the 5% significance level with the following hypotheses:

$H_0$ : the population mean mass of the snack is $\mu_0$

$H_1$ : the population mean mass of the snack is $< \mu_0$

Given that $H_0$ is rejected in favour of $H_1$, find the set of possible values of $\mu_0$. Assume that the distribution of the mass of the snack is normal. [6]

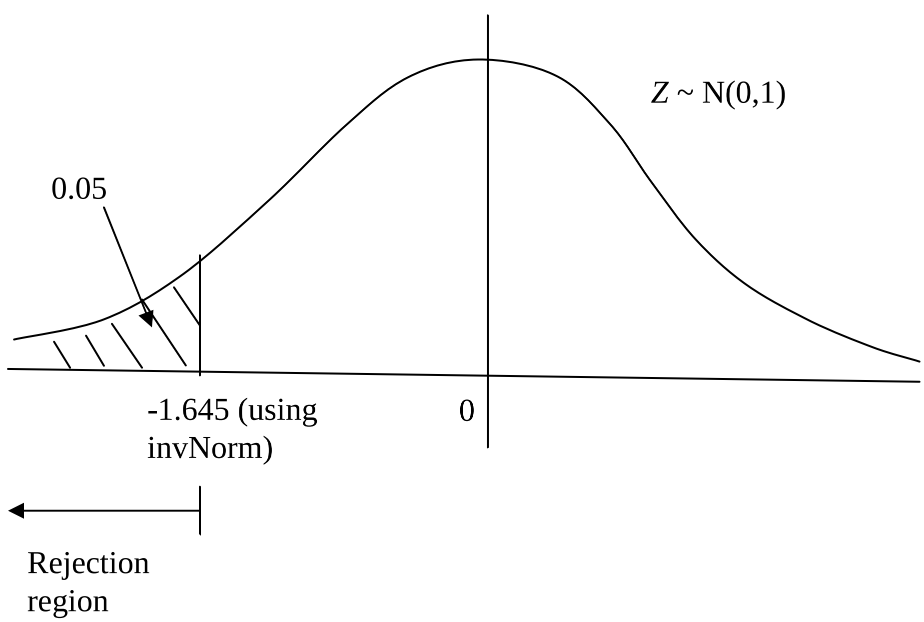
(Source: CJC'07 P2/9)

<u>Solution</u>:

This falls under **case 2**.

$H_0 : \mu = \mu_0$

$H_1 : \mu < \mu_0$




$$\bar{x} = \frac{-80}{50} + 150 = 148.4$$

$$s^2 = \frac{1}{49}\left(2500 - \frac{80^2}{50}\right) = 48.408$$

Under $H_0$, $Z = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim N(0,1)$

So $z = \dfrac{148.4 - \mu_0}{\sqrt{\dfrac{48.408}{50}}}$

Since $H_0$ is rejected, $\dfrac{148.4-\mu_0}{\sqrt{\dfrac{48.408}{50}}}$ < InvNorm(0.05) = –1.64485

$148.4 - \mu_0 < -1.61845$, so $\mu_0 > 150$ ❖

**Q2** In a particular workshop, it is claimed that the masses of components produced have mean 6g and standard deviation 0.8g. If this claim is not rejected at the 5% level of significance based on the mean mass obtained from a random sample of 10 components, between what values must the mean mass of the 10 components in the sample lie? State any assumptions necessary for validity. [4]

(Source: YJC'06 P2/10)

Solution:
This falls under **case 3**.
Let $X$ be the r.v. "the mass, in g, of a component".

Assumption: The masses of components produced are normally distributed. ❖

$H_0: \mu = 6$
$H_1: \mu \neq 6$

Under $H_0$, $Z = \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0, 1)$ where $\mu = 6$, $\sigma = 0.8$, $n = 10$

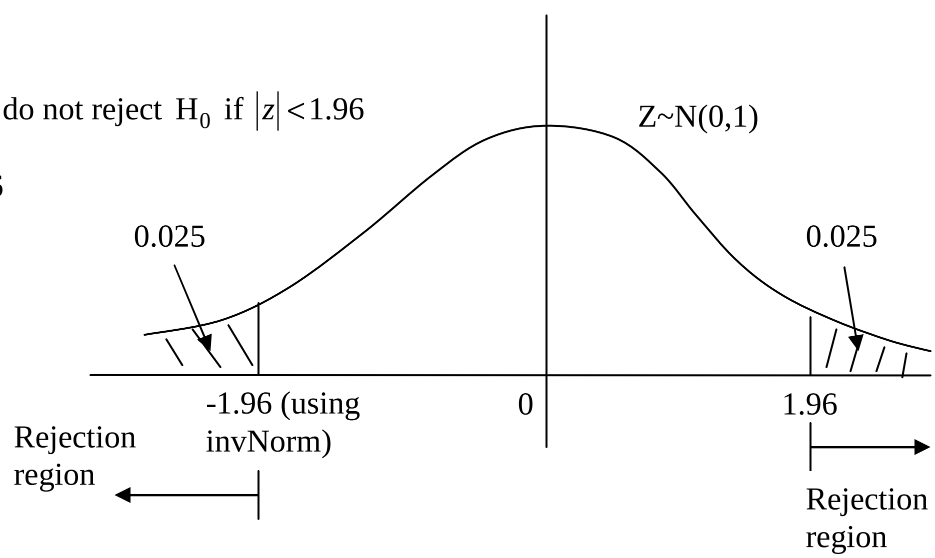
Level of significance: $\alpha = 5\%$

We reject $H_0$ if $|z| > 1.96$ and do not reject $H_0$ if $|z| < 1.96$

i.e. $-1.96 < \dfrac{\bar{x}-6}{0.8/\sqrt{10}} < 1.96$

$5.504 < \bar{x} < 6.496$

i.e. $5.50 < \bar{x} < 6.50$ ❖




**Note:** The population variance of $\sigma$ = 0.8 is claimed. So it is not *t*-test.

**Q3** A sample of $n$ observations is taken from a population with mean 5. The sample mean is found to be 5.1 and sample standard deviation 0.7. A 2-tailed normal test ($z$ – test) is conducted and the null hypothesis is rejected at the 5% significance level. Find the least value of $n$. [4]

(Source: IJC'07 P2/11)

<u>Solution:</u>

This falls under **case 1**.

$H_0 : \mu = 5$

$H_1 : \mu \neq 5$

Perform a 2 tailed ($z$-test) at 5% significance level

Since $H_0$ is rejected,

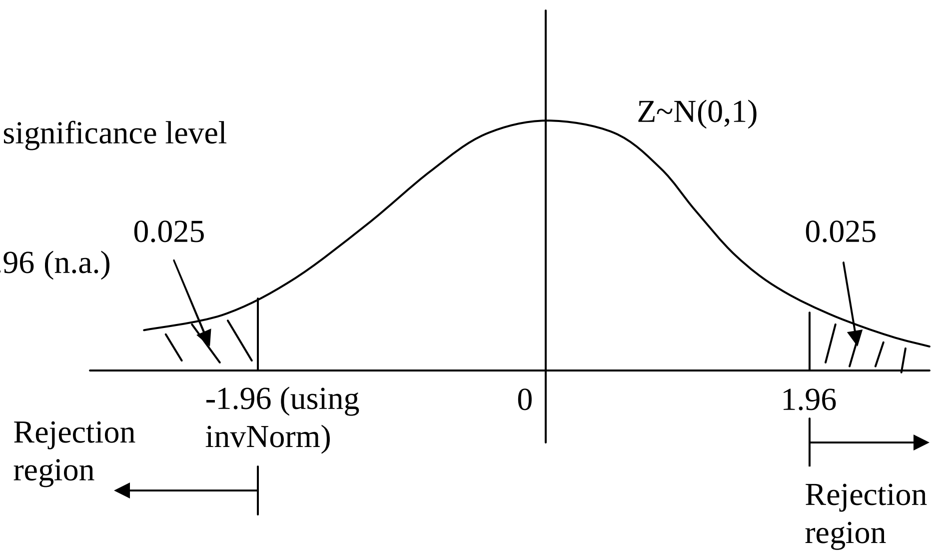
$$|Z| > 1.96$$

$$Z > 1.96 \quad \text{or} \quad Z < -1.96 \text{ (n.a.)}$$

$$\frac{5.1-5}{0.7/\sqrt{n-1}} > 1.96$$

$$n > 189.2384$$

Least $n$ = 190 ❖




**Note**: You must be careful here because sample standard deviation is given while $s$ is not. So you must use $s^2 = \frac{n}{n-1}(\text{sample variance})$ to get $s$ first. That is why the denominator becomes $\sqrt{n-1}$ and not $\sqrt{n}$.

i.e. $s^2 = \frac{n}{n-1}(\text{sample variance})$

$$= \frac{n}{n-1}(0.7)^2$$

Therefore $s = \frac{0.7\sqrt{n}}{\sqrt{n-1}}$

Thus $\frac{\overline{X}-\mu}{s/\sqrt{n}}$ becomes $\frac{\overline{X}-\mu}{0.7/\sqrt{n-1}}$

**Q4** In general, the marks obtained by students in their Mathematics examination may be assumed to be a random variable with mean 65 and variance 32.

After the introduction of a new programme of revision classes, it is found that there are improvements to the marks of some of the students.

A random sample of the marks of 50 students is taken and the mean is found to be 66.4. A test is conducted and it shows that there is insufficient evidence that the new programme of revision classes is effective. Find an inequality satisfied by the significance level of the test. [3]

State, with a reason, whether it is necessary to assume that the Mathematics examination marks follow a normal distribution. [1]

(Source: IJC'07 P2/11)

Solution:

This falls under a **new case** of $\alpha$ being an unknown. We look at the approach of solving it.

$H_0 : \mu = 65$

$H_1 : \mu > 65$

Perform a 1 tailed (*z*-test) at $\alpha\%$ significance level

Under $H_0$, $\bar{X} \sim N\left(65, \frac{32}{50}\right)$ by CLT.

Since $H_0$ is not rejected, *p*-value = $P\left(\bar{X} > 66.4\right) > \alpha\%$

$0.04006 > \alpha\%$

$\alpha\% < 4.01\%$ ❖

OR

Enter into GC: $\mu_0 = 65, \sigma = \sqrt{32}, \bar{x} = 66.4, n = 50$

Choose $\mu > \mu_0$

We obtain $p = 0.04006$. Since $H_0$ is not rejected,

$0.04006 > \alpha\%$

$\alpha\% < 4.01\%$

No. Since the sample size 50 is large, the mean marks of 50 students follow a normal distribution approximately (according to CLT). ❖