Chapter 6 (Statistics) : Hypothesis Testing

Objectives:

At the end of the chapter, you should have learnt about the following:

- Concepts of null hypothesis (H_0) and alternative hypothesis (H_1), test statistic, critical region, critical value, level of significance, and *p*-value
- Formulation of hypotheses and testing for a population mean based on:
  - a sample from a normal population of known variance
  - a large sample from any population
- 1-tailed and 2-tailed tests
- Interpretation of the results of a hypothesis test in the context of the question

Pre-requisite Knowledge

- Sampling Distribution of Means;
- Point Estimates

Content

6.1 Introduction to Hypothesis Testing

Are Coke cans really 330 ml? Are this type of dumbbells at least 10 kg? Will that slimming centre help customers lose up to 16 pounds in just 4 weeks as promised?

Claims abound in advertising and merchandise, as well as elsewhere e.g. social sciences. The most objective way to test whether a claim is valid is to take a (random) sample and measure it.

To be specific: In H1 Maths, we measure its sample mean ($\bar{x}$) and calculate how much it deviates from an assumed population mean (μ_0 , which could be based on known facts or some unverified claim). This deviation is expressed in terms of a *relative* score or probability, which we compare with a predetermined standard to reach a conclusion about whether the claim is valid. This whole procedure is known as **hypothesis testing**.

A **statistical hypothesis** is a statement, which may or may not be true, concerning one or more population parameters (e.g. mean, variance, proportion).

Hypothesis testing is the procedure used to test the validity of this statement based on evidence from random samples drawn from the population.

Let us look at some concrete examples:

<u>Example 1a</u>

A fizzy drink brand claims that the mean volume of its canned drinks (the volume is normally distributed with standard deviation 10ml) is 330ml. A consumer weighs a random carton of 24 drinks and found that it has a mean volume of 325ml. He suspects that the mean volume is less than what is claimed. Test the manufacturer's claim at the 5% level of significance.

Solution:

Let X be the volume (in ml) of a canned drink with population mean volume μ.

$X \sim \mathrm{N}(\mu, 10^2)$

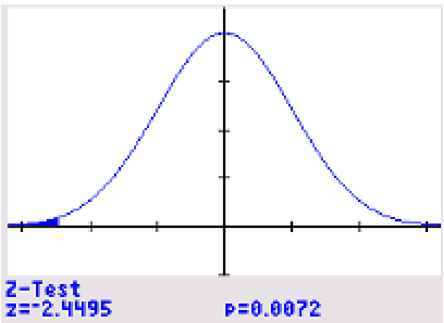
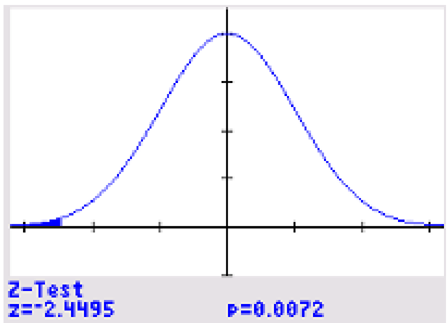
$\overline{X} \sim \mathrm{N}(330, \frac{10^2}{24})$

Test $\mathrm{H}_0 : \mu = 330$

vs. $\mathrm{H}_1 : \mu < 330$ at the 5% level of significance.

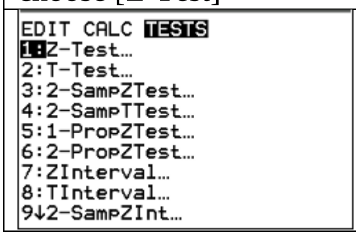
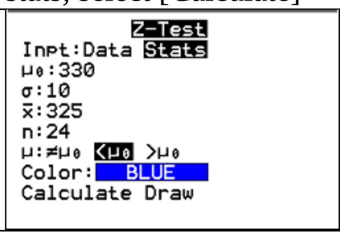
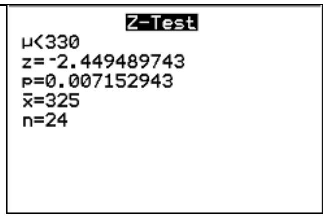
Under H_0, $Z = \dfrac{\overline{X} - \mu}{\sigma / \sqrt{n}} = \dfrac{\overline{X} - 330}{10/\sqrt{24}} \sim \mathrm{N}(0,1)$

Note to teachers:
These examples are meant to reduce the abstractness. Teachers can use Section 6.2 to elucidate the theory and demonstrate the calculations involved.

Option 1: Critical Value Method	**Option 2: *p*-value Method**
Using a one-tailed *z*-test, reject H_0 if $z \le -1.64485$.  Using GC, $\overline{x} = 325$ gives z-cal = –2.4495 < –1.64485 $\therefore$ We reject H_0 and conclude that at the 5% level of significance, there is sufficient evidence that the mean volume of a canned drink is less than 330ml.	Using a one-tailed *z*-test, reject H_0 if $p \le 0.05$.  Using GC, $\overline{x} = 325$ gives *p*-value = 0.00715 < 0.05 $\therefore$ We reject H_0 and conclude that at the 5% level of significance, there is sufficient evidence that the mean volume of a canned drink is less than 330ml.

Note: $p = \mathrm{P}\left(\overline{X} \le 325\right)$

$= \mathrm{P}\left(Z \le -2.4495\right)$

Press [Stats], choose [Z-Test]	Choose [Stats], key in stats, select [Calculate]	The output:
		

Example 1b (two-tailed test)

A fizzy drink brand claims that the mean volume of its canned drinks (the volume is normally distributed with standard deviation 10ml) is 330ml. A random carton of 24 drinks has a mean volume of 325ml. Test the manufacturer's claim at 5% level of significance.

Solution:

Let X be the volume (in ml) of a canned drink with population mean volume μ.

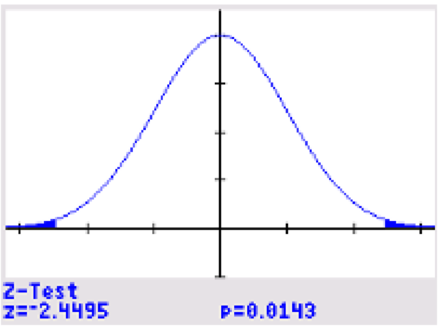
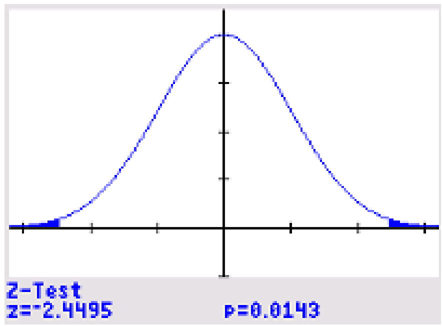
$X \sim \mathrm{N}(\mu, 10^2)$

$\overline{X} \sim \mathrm{N}(330, \frac{10^2}{24})$

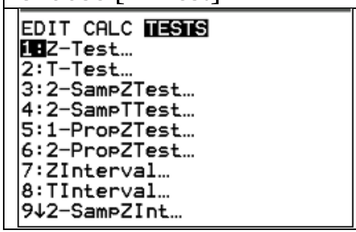
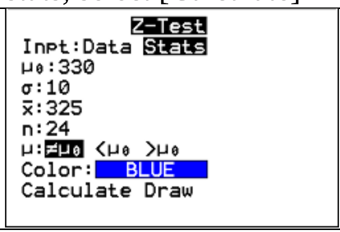
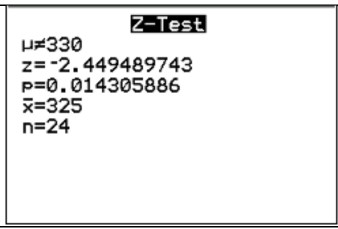
Test $\quad \mathrm{H}_0 : \mu = 330$

vs. $\quad \mathrm{H}_1 : \mu \neq 330 \quad$ at the 5% level of significance.

Under H_0, $Z = \dfrac{\overline{X} - \mu}{\sigma / \sqrt{n}} = \dfrac{\overline{X} - 330}{10/\sqrt{24}} \sim \mathrm{N}(0,1)$

Option 1: Critical Value Method	**Option 2: *p*-value Method**
Using a two-tailed *z*-test, reject H_0 if $z \le -1.95996$ or $z \ge 1.95996$.  Using GC, $\bar{x} = 325$ gives *z*-cal = –2.4495 < –1.95996 $\therefore$ We reject H_0 and conclude that at the 5% level of significance, there is sufficient evidence that the mean volume of a canned drink is not 330ml.	Using a two-tailed *z*-test, reject H_0 if $p \le 0.05$.  Using GC, $\bar{x} = 325$ gives *p*-value = 0.0143 < 0.05 $\therefore$ We reject H_0 and conclude that at the 5% level of significance, there is sufficient evidence that the mean volume of a canned drink is not 330ml.

Note: The implied message is that since the *p*-value is really low ($\le$ 0.05), it is unlikely that the mean volume is indeed 330ml.

Press [Stats], choose [Z-Test]	Choose [Stats], key in stats, select [Calculate]	The output:
		

6.2 Terms & definitions

6.2.1 Null Hypothesis and Alternative Hypothesis

Are we testing for (a) any change/difference from the (assumed) population mean μ_0,
or (b)/(c) a definite increase/decrease in the (assumed) population mean?

Null hypothesis (denoted by H_0)

H_0 is a statement that says that a **population parameter** takes a **specific value**.

Alternative hypothesis (denoted by H_1)

H_1 is a statement that says that the population parameter takes a value different, in some way, from the value given in H_0.

Test $\quad \mathrm{H}_0 : \mu = \mu_0$ (Assumption: population mean is μ_0)

vs. $\quad$ (a) $\mathrm{H}_1 : \mu \neq \mu_0$ $\quad$ (b) $\mathrm{H}_1 : \mu > \mu_0$ $\quad$ (c) $\mathrm{H}_1 : \mu < \mu_0$

In a two-tailed test, H_1 looks for any change in the population mean (can be increase or decrease).

In a one-tailed test, H_1 looks for a definite increase (or decrease) in the population mean.

Note:

1. The rejection of the null hypothesis will generally imply that the alternative hypothesis is accepted. In this case, we conclude that the null hypothesis is false.
2. Failure to reject the null hypothesis is not the same as accepting the null hypothesis. It only means there is insufficient statistical evidence to reject H_0 in favour of H_1 (i.e. H_0 may or may not be true). In other words, **we cannot conclude that we accept H$_0$**.

6.2.2 Test Statistic

The sample statistic used to test the hypotheses is called the **test statistic**. In the H1 Maths syllabus, we take the sample mean $\bar{X}$ and *standardise* it to obtain the standard normal distribution *Z*. In this way, any particular value $\bar{x}$ will have a corresponding standard score (also known as *z*-score), which is a measure of its number of standard deviations above or below μ. Recall that if we use a large sample size *n*, then $\bar{X}$ approximately follows a normal distribution by Central Limit Theorem (CLT), so a bell curve appears!

If we now assume our null hypothesis is true, then we get something familiar:

$$\text{Under } \mathrm{H}_0,\ \bar{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right) \text{ and so } Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim \mathrm{N}(0,1).$$

Note:

1. If the population *X* is already normal, then Central Limit Theorem is not required.
2. If σ^2 is unknown, we use the unbiased estimate s^2 (S_x^2 in GC) instead.

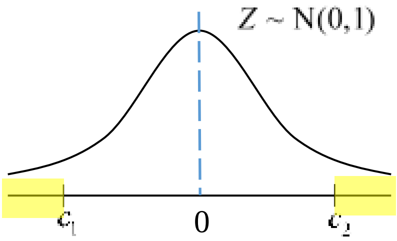
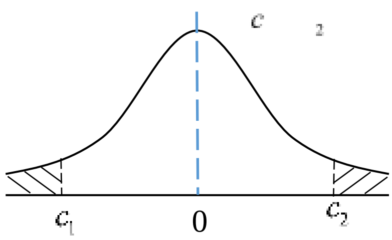
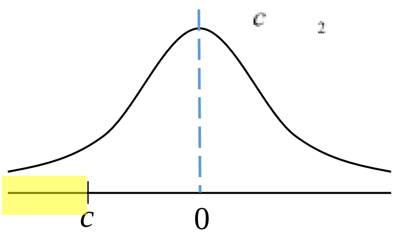
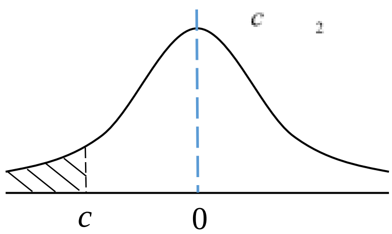
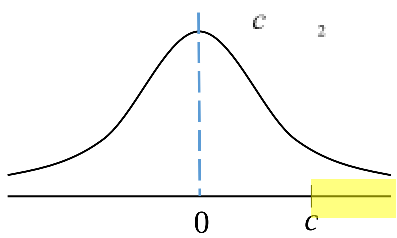
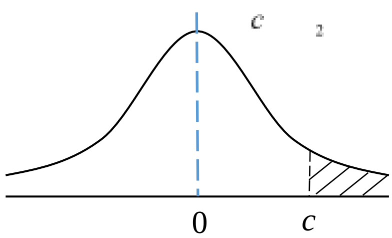
$$s^2 = \frac{n}{n-1}\left(\frac{\Sigma(x-\bar{x})^2}{n}\right) = \frac{1}{n-1}\left(\Sigma x^2 - \frac{(\Sigma x)^2}{n}\right) = \frac{n}{n-1} \times \text{sample variance}$$

Inside MF26 $\quad$ Not inside MF26

6.2.3 Comparing *z*-score and critical value; or *p*-value and level of significance

The extent the test statistic differs from μ is measured by its *z*-score and *p*-value:

***z*-score = the no. of standard deviations the test statistic is above/below μ**	***p*-value = the probability of the test statistic being that no. of standard deviations above/below μ**
The larger the absolute *z*-score, the greater the deviation.	The smaller the *p*-value, the greater the deviation.
If the *z*-score lies beyond a predetermined value (or two) called the **critical value(s)**, the test stat is deemed to be significantly far from μ. We say the test is **statistically significant**. We will then **reject** H_0.	If the *p*-value falls below a predetermined % called the **level of significance**, the test stat is deemed to be significantly far from μ. We say the test is **statistically significant**. We will then **reject** H_0.

	c – critical value – critical region	α% – level of significance Total striped area = α%
Two-tailed test		
One-tailed tests		
		

E.g. For a 5% level of significance (i.e. α = 5%),
Two-tailed test: $c_1 = -1.96$ and $c_2 = 1.96$; Critical region = $\{z \in \mathbb{R} : z \leq -1.96 \text{ or } z \geq 1.96\}$
Left-tailed test: $c = -1.645$; Critical region = $\{z \in \mathbb{R} : z \leq -1.645\}$
Right-tailed test: $c = 1.645$; Critical region = $\{z \in \mathbb{R} : z \geq 1.645\}$

The critical value(s) is found using [InvNorm] in the GC.

6.2.4 Formal definitions of *p*-value and level of significance

There are two formal definitions of the *p*-value of a test:

(1) The smallest level of significance at which H_0 can be rejected.
(2) The probability that the sample mean is at least as extreme as observed, assuming H_0 is true.

The **level of significance** of a test (denoted by α) is the probability of rejecting H_0 when H_0 is true.

Note:

1. The level of significance $\frac{\alpha}{100}$ is the conditional probability $\mathrm{P}(\text{reject } \mathrm{H}_0 \mid \mathrm{H}_0 \text{ is true})$.
2. The lower the level of significance (i.e. the smaller the value of α), the smaller the probability of making **this error** of rejecting H_0 when H_0 is true. α and *p*-value are tools that help us <u>quantify</u> and <u>control</u> **this error**.
3. This margin of error is typically set at 5% as that is an industrial and research standard. That is why many of our examples specify a 5% level of significance.
4. No hypothesis test can yield 100% accuracy. We could end up with a test statistic in the Critical Region 5% of the time when H_0 is true, simply due to random sampling error. The luck of the draw matters.

In Example 1a, why do we need to do a hypothesis test when 325ml is *clearly* less than 330ml?

The reason is simple: We only looked at a **sample**, not the entire population. Perhaps the population mean is indeed 330ml and the lower value of 325ml was due to sampling error. (If we take another carton of 24 cans and measured their volume, the new sample mean might be close to 330ml.) A hypothesis test assesses the likelihood that scenario will occur, without having to literally take another sample! Thus, we can make a decision backed by statistical theory.

In theory, if we repeatedly measure different cartons, we can plot a distribution of the sample means, called the **sampling distribution of means**. You might recall this theoretical distribution follows a normal distribution for large samples.

6.3 Carrying out a Hypothesis Test

This is the outline of the steps in hypothesis testing:

1. Define the appropriate random variables and population distribution.
2. State the probability distribution from which the observations were taken.
3. State the null hypothesis H_0 and alternative hypothesis H_1.
4. State the type of test and level of significance α.
5. Identify the test statistic and its distribution.
6. Either Compute the *z*-score and compare it with the critical value(s)
   or Compute the *p*-value and compare it with α.
7. Draw conclusions – reject or do not reject H_0.

With the availability of computational tools such as GC, it is more convenient to use the *p*-value method. However, the critical value method is useful in situations where there are unknowns to be found.

The critical value method requires standardising the test statistic to the standard normal variable *Z* (for H1 Maths). We shall illustrate the use of both the critical value and *p*-value methods in a few scenarios.

Hands On Practice

2016/PJC/Promo/7

The manager of a shopping mall claims that customers spend more than 4 hours shopping in the mall. To investigate this claim, the time, x hours, spent by 60 randomly chosen customers is recorded and summarised as follows.

$$\sum(x-4)=10.6 \text{ and } \sum(x-4)^2=63.2.$$

(i) Calculate unbiased estimates of the population mean and variance.

(ii) Carry out a test at the 2% level of significance to determine whether the manager's claim is true.

[Answer: $\bar{x}\approx 4.18$, $s^2 \approx 1.04$]

6.3.1 Sample size large/small, Population normal, σ^2 known

Recall:

If $X \sim N\left(\mu, \sigma^2\right)$, then $\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ for any sample size *n*.

Use **z-test** with test statistic $Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \sim \mathrm{N}(0,1)$.

<u>Example 2</u>

Climbing ropes previously produced by a manufacturer was known to have breaking strengths that are normally distributed with mean 170.2 kg and standard deviation 10.5 kg. A new component material is added to the ropes being produced. The manufacturer believes that this would increase the mean breaking strength without changing the standard deviation. A random sample of 50 of such new ropes produced is found to have a mean breaking strength of 172.4 kg. Perform a significance test at 5% level to decide whether this result provides sufficient evidence to confirm the manufacturer's belief that the mean breaking strength is increased. State your hypothesis and conclusion clearly.

Solution:

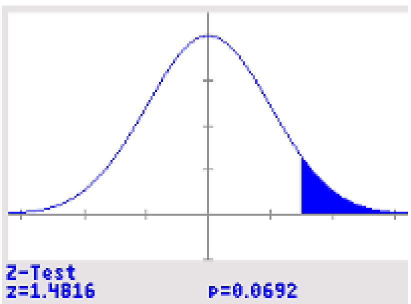
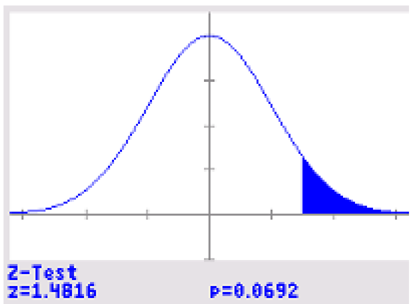
Let X be the breaking strength (in kg) of a climbing rope and μ be the population mean breaking strength.

Test $\mathrm{H}_0 : \mu = 170.2$

vs. $\mathrm{H}_1 : \mu > 170.2$ at the 5% level of significance.

Under H_0, $\overline{X} \sim \mathrm{N}(170.2, \frac{10.5^2}{50})$

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\overline{X} - 170.2}{10.5/\sqrt{50}} \sim \mathrm{N}(0,1).$$

Option 1: Critical Value Method	**Option 2: *p*-value Method**
Using a one-tailed *z*-test, reject H_0 if $z \geq 1.64485$.  Using GC, $\bar{x} = 172.4$ gives *z*-score = 1.4816 < 1.64485	Using a one-tailed *z*-test, reject H_0 if $p \leq 0.05$.  Using GC, $\bar{x} = 172.4$ gives *p*-value = 0.0692 > 0.05

$\therefore$ We <u>do not reject</u> H_0 and conclude that at the 5% level of significance, there is <u>insufficient</u> evidence that <u>the mean breaking strength has increased</u>.

6.3.2 Sample size large, Population non-normal, σ^2 known

Recall :

> Since n is large, Central Limit Theorem tells us that
>
> $$\overline{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right) \text{ approximately.}$$
>
> Use **z-test** with test statistic $Z = \dfrac{\overline{X} - \mu}{\sigma / \sqrt{n}} \sim \mathrm{N}(0,1)$ approximately.

<u>Example 3</u>

The mean IQ of Math professors was known to be 118, but is now claimed to be greater than 118. It is known that the population variance is 64. A sample (n = 50) yields $\overline{x} = 120$.
(a) Perform a hypothesis test on the claim at 5% significance level.
(b) Using the same data, what would be the conclusion for this claim at 5% significance level: The mean IQ of Math professors is <u>not equal</u> to 118.

Solution:

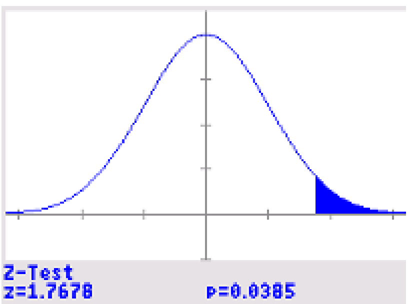
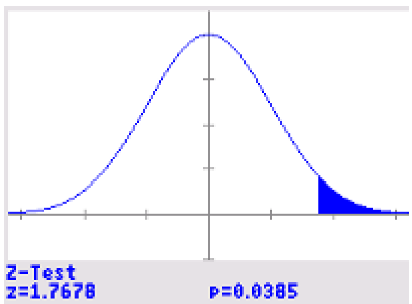
(a) Let X be the IQ of a Math professor and μ be the population mean IQ.

Test $\quad \mathrm{H_0} : \mu = 118$

vs. $\quad \mathrm{H_1} : \mu > 118 \quad$ at the 5% level of significance.

Under $\mathrm{H_0}$, since $n = 50$ is large, by Central Limit Theorem, $\overline{X} \sim \mathrm{N}(118, \frac{64}{50})$ approximately

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} = \frac{\overline{X} - 118}{8 / \sqrt{50}} \sim \mathrm{N}(0,1)$$

Option 1: Critical Value Method	**Option 2: *p*-value Method**
Using a one-tailed z-test, reject $\mathrm{H_0}$ if $z \geq 1.64485$.	Using a one-tailed z-test, reject $\mathrm{H_0}$ if $p \leq 0.05$.
	
Using GC, $\overline{x} = 120$ gives z-score = 1.7678 > 1.64485	Using GC, $\overline{x} = 120$ gives p-value = 0.0385 < 0.05

$\therefore$ We <u>reject</u> $\mathrm{H_0}$ and conclude that at the 5% level of significance, there is <u>sufficient</u> evidence that <u>the mean IQ of Math professors is more than 118</u>.

(b) Using a 2-tailed z-test, p-value = 2 × 0.0385
= 0.0771 > 0.05

$\therefore$ We <u>do not reject</u> $\mathrm{H_0}$ and conclude that at the 5% level of significance, there is <u>insufficient</u> evidence that <u>the mean IQ of Math professors is not 118</u>.

Example 4

The length of string in the balls of string made by a particular manufacturer has mean μ m and variance 27.4 m^2. The manufacturer claims that the mean length is at least 300 m. A random sample of 100 balls of string is taken and the sample mean is found to be 299.2 m. Test whether this provides significant evidence, at 3% level, that the manufacturer's claim overstates the value of μ.

Solution:

Let X be the length (in m) of a ball of string and μ be the population mean length.

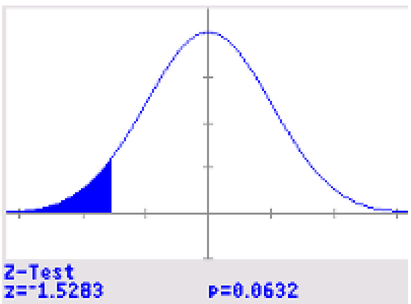
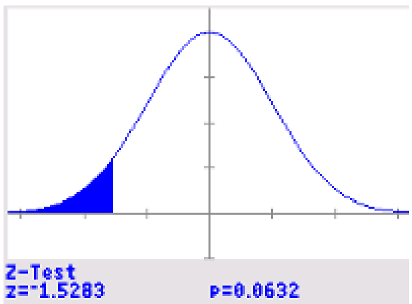
Test $\quad \mathrm{H_0} : \mu = 300$

vs. $\quad \mathrm{H_1} : \mu < 300 \quad$ at the 3% level of significance.

Under $\mathrm{H_0}$, since $n = 100$ is large, by Central Limit Theorem.

$\overline{X} \sim \mathrm{N}(300, \frac{27.4}{100})$ approximately

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} = \frac{\overline{X} - 300}{\sqrt{27.4 / 10}} \sim \mathrm{N}(0,1)$$

Option 1: Critical Value Method	**Option 2: *p*-value Method**
Using a one-tailed *z*-test, reject $\mathrm{H_0}$ if $z \le -1.8808$.  Using GC, $\bar{x} = 299.2$ gives *z*-score = –1.5283 > –1.8808	Using a one-tailed *z*-test, reject $\mathrm{H_0}$ if $p \le 0.03$.  Using GC, $\bar{x} = 299.2$ gives *p*-value = 0.0632 > 0.03

$\therefore$ We do not reject $\mathrm{H_0}$ and conclude that at the 3% level of significance, there is insufficient evidence that the manufacturer's claim overstates the value of μ.

6.3.3 Sample size large, Population non-normal, σ^2 unknown

Recall :

We can estimate the population variance using the unbiased estimate s^2.   Then by Central Limit Theorem, $\overline{X} \sim \mathrm{N}\left(\mu, \frac{s^2}{n}\right)$ approximately, since sample size is large.   Use ***z*-test** with test statistic $Z = \dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim \mathrm{N}(0,1)$.

<u>Example 5</u>

A horticulturist takes a random sample of 80 bean seeds of a particular variety, and sows them under standard laboratory conditions. The number of the seeds that germinate is 64. The crop weight, x kg, of each of the 64 plants is measured, with the following summarised results: $\Sigma x = 303.4$, $\Sigma x^2 = 1615.96$. Find the unbiased estimate of the population mean and variance of crop weight per plant. The horticulturist wishes to test the hypothesis that the mean crop weight per plant is 5 kg, against the alternative hypothesis that the mean crop weight per plant is less than 5 kg. Carry out the test at the 10 % level of significance.

Find the smallest level of significance at which the test would result in rejection of the null hypothesis.

Solution:

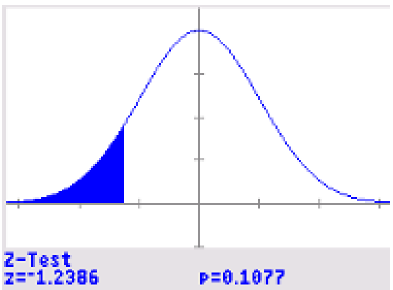
Let X be the crop weight (in kg) of a plant and μ be the population mean weight.

Unbiased estimate of population mean = $\overline{x} = \dfrac{303.4}{64} \approx 4.7406 = 4.74$ (3 s.f.)

Unbiased estimate of population variance, $s^2 = \dfrac{1}{63}\left[1615.96 - \dfrac{303.4^2}{64}\right] \approx 2.8199 = 2.82$ (3 s.f.)

Test $\quad \mathrm{H}_0 : \mu = 5$

vs. $\quad \mathrm{H}_1 : \mu < 5 \quad$ at the 10% level of significance.



Under H_0, since $n = 64$ is large, by Central Limit Theorem.

$\overline{X} \sim \mathrm{N}(5, \dfrac{2.8199}{64})$ approximately

$$Z = \frac{\overline{X}-\mu}{S/\sqrt{n}} = \frac{\overline{X}-5}{\sqrt{2.8199}/8} \sim \mathrm{N}(0,1)$$

Using a one-tailed z-test,
reject H_0 if $p \le 0.10$.

Using GC, $\overline{x} = 4.74$
gives p-value = 0.1077 > 0.10

$\therefore$ We do not reject H_0 and conclude that at the 10% level of significance, there is insufficient evidence that the mean crop weight of the plants is less than 5 kg.

If H_0 is rejected at the α % level, then p-value $\le$ α %.

$$\frac{\alpha}{100} \ge 0.1077 \Rightarrow \alpha \ge 10.77$$

The smallest significance level is 10.8%. (3 s.f.)

6.4 Further Examples

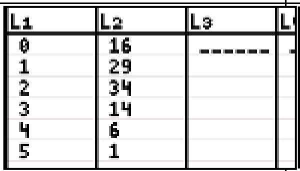
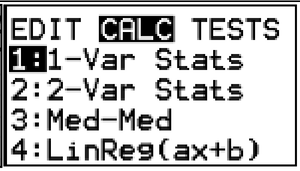
Example 6

The table below shows the number of handphones owned by 100 households.

Number of handphones per household, x	0	1	2	3	4	5
Number of households with a certain number of handphones, f	16	29	34	14	6	1

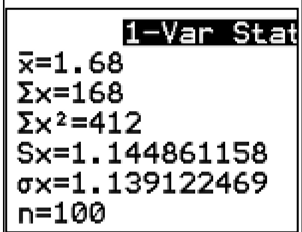
a) Find the mean and standard deviation of the data.
b) Test the claim that the population mean is not equal to 1.60 at the 5% level of significance.

Solution:

Press [Stats], choose [Edit], [1]	Key in the values under L_1 and L_2	Press [Stats], select [Calc], [1]	Key in List: L_1 and FreqList: L_2
			

(a) Using GC,
mean of data = $\bar{x} = 1.68$
standard deviation of data ≈ 1.1391 (σ_x in GC) = 1.14 (3 s.f.)

Output:



(b) Let μ be the population mean number of handphones in a household.

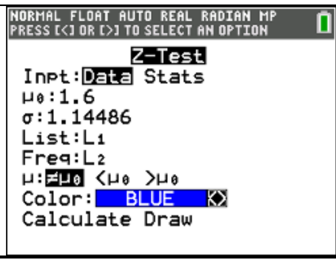
Test $H_0 : \mu = 1.60$
vs. $H_1 : \mu \neq 1.60$ at the 5% level of significance.
Under H_0, since $n = 100$ is large, by Central Limit Theorem.

$$\bar{X} \sim N(1.6, \frac{1.14486^2}{100}) \text{ approximately}$$

$$Z = \frac{\bar{X} - \mu}{s/\sqrt{n}} = \frac{\bar{X} - 1.60}{1.14486/10} \sim N(0,1)$$

Using a two-tailed z-test,
reject H_0 if $p \leq 0.05$.
$\bar{x} = 1.68$ gives p-value = 0.4547 > 0.05

$\therefore$ We do not reject H_0 and conclude that at the 5% level of significance, there is insufficient evidence that the population mean is not equal to 1.60.

Choose [Data] instead. Key in Stats, Select [Calculate]	

<u>Example 7</u> [2015/H2 Math 9740/Prelims/YJC/II/8 (modified)]

Under ordinary conditions, the mean time required by a person to complete an arithmetic sum is 50 s. Sixty people were randomly selected to do the sum in a room with a temperature set at $15°\text{C}$. The time taken for this sample had an average of 55 s and a standard deviation of 19.3 s.

(i) Test, at the 5% significance level, whether there has been an increase in the mean time required by a person to complete the sum.

(ii) Explain whether any assumptions about the population are needed in order for the test to be valid.

(iii) The same sample is now used to carry out a test at the 5% significance level, of whether the mean time required by a person to complete the sum has changed. State with a reason using **(i)**, whether there is any change in the conclusion of this test.

Solution:

(i) Let X be the length of time (in s) required to complete the sum and μ be the population mean time.

Test $\quad \text{H}_0 : \mu = 50$

vs. $\quad \text{H}_1 : \mu > 50 \quad$ at the 5 % significance level.

Unbiased estimate of the population variance, $s^2 = \dfrac{60}{59}(19.3)^2 \approx 378.80$

Under H_0, since $n = 60$ is large, by Central Limit Theorem.

$\overline{X} \sim \text{N}(50, \dfrac{378.8}{60})$ approximately

$$Z = \frac{\overline{X} - \mu}{s/\sqrt{n}} = \frac{\overline{X} - 50}{\sqrt{378.80}/\sqrt{60}} \sim \text{N}(0,1)$$

Using a one-tailed z-test,
reject H_0 if $p \le 0.05$.

$\overline{x} = 55$ gives p-value = 0.0233 < 0.05

$\therefore$ We reject H_0 and conclude that at the 5% significance level, there is sufficient evidence that the mean time required to complete the sum has increased.

(ii) Since the sample size (60) is large, Central Limit Theorem can be used to approximate the sample mean time required to complete the sum to a normal distribution. Hence no assumption about the population distribution is necessary.

(iii) For a two-tailed test, p-value = 2(0.0233) (from **(i)**)
= 0.0466 < 0.05

$\therefore$ We still reject H_0. There is no change in the conclusion.

<u>Example 8</u> [2014/H2 Math 9740/ACJC/II/Q11 (part)]

The mean mathematics examination mark for the whole school was 55 last year. Let the random variable X denote the mathematics examination mark obtained by a student this year.

A mathematics teacher has finished marking all the 120 scripts of the mathematics examination this year for his tutorial classes. The standard deviation of his 120 students' marks is 15. He carried out a test, at the 5% significance level, to determine whether the mean mark for the mathematics examination for the whole school this year has exceeded that of last year.

Find the range of values of $\bar{x}$ for which the result of the test would be to reject the null hypothesis, where $\bar{x}$ is the mean mark of his 120 students. (Answers obtained by trial and improvement from a calculator will obtain no marks.)

Solution:

Let μ be the population mean maths exam mark obtained by a student this year.

Test $\quad \mathrm{H}_0 : \mu = 55$

vs. $\quad \mathrm{H}_1 : \mu > 55 \quad$ at the 5% level of significance.

Unbiased estimate of population variance, $s^2 = \dfrac{120}{119} \times 15^2 \approx 226.89$

Under H_0, since $n = 120$ is large, by Central Limit Theorem.

$\overline{X} \sim \mathrm{N}(55, \dfrac{226.89}{120})$ approximately

$$Z = \frac{\overline{X} - \mu}{s/\sqrt{n}} = \frac{\overline{X} - 55}{\sqrt{226.89}/\sqrt{120}} \sim \mathrm{N}(0,1)$$

Using a one-tailed z-test,
reject H_0 if $z \geq 1.64485$.

$$\frac{\bar{x} - 55}{\sqrt{\frac{226.89}{120}}} \geq 1.64485$$

$$\bar{x} \geq 57.262$$

$$\bar{x} \geq 57.3 \text{ (3 s.f.)}$$

<u>Hands on practice</u>

At an early stage in analysing the marks scored by the large number of candidates in an examination paper, the Examining Board takes a random sample of 250 candidates and finds that the marks, x, of these candidates give $\Sigma x = 11872$ and $\Sigma x^2 = 646\,193$. Using the figures obtained in this sample, the null hypothesis $\mu = 49.5$ is tested against the alternative hypothesis $\mu < 49.5$ at the α % significance level.

Determine the set of values of α for which the null hypothesis is rejected in favour of the alternative hypothesis.

[Answer: $\bar{x} = 47.488$, $s^2 \approx 330.99$, $\alpha \geq 4.02$]

Exercise

1. **[N2001/2/Q10]**

A random sample of 90 batteries, used in a particular model of mobile phone, is tested and the "standby-time" is measured. The results are summarised by

$$\sum x = 3040.8 \text{ and } \sum x^2 = 115773.66.$$

Test, at the 1 % significance level, whether the mean standby-time is less than 36.0 h.

This type of battery is advertised as having a 'talk-time" of 5 hours. In a test at the 5% significance it is found that there is significant evidence that the population mean talk-time is less than 5 hours. Using only this information, and giving a reason in each case, state whether each of the following statements is (i) necessarily true, (ii) necessarily false, (iii) neither necessarily true nor necessarily false.

(a) There is significant evidence at the 10% significance level that the population mean talk-time is less than 5 hours.

(b) There is significant evidence at the 5% significance level that the population mean talk-time is not 5 hours.

Solution:

Unbiased estimate of the population mean, $\bar{x} = \dfrac{3040.8}{90} \approx 33.787$

Unbiased estimate of the population variance, $s^2 = \dfrac{1}{89}\left[115773.66 - \dfrac{3040.8^2}{90}\right] \approx 146.463$

Let X be the standby-time (in h) of a battery in this model of mobile phone and μ be the population mean time.

Test $\quad \mathrm{H}_0 : \mu = 36 \quad$ vs. $\quad \mathrm{H}_1 : \mu < 36 \quad$ at the 1% level of significance.

Under H_0, since n = 90 is large, by Central Limit Theorem.

$\overline{X} \sim \mathrm{N}(36, \dfrac{146.463}{90})$ approximately

$$Z = \frac{\overline{X} - \mu}{s/\sqrt{n}} = \frac{\overline{X} - 36}{\sqrt{146.463}/\sqrt{90}} \sim \mathrm{N}(0,1)$$

Using a one-tailed z-test,
reject H_0 if $p \le 0.01$.

$\bar{x} = 33.787$ gives p-value = 0.0414 > 0.01

$\therefore$ We do not reject H_0 and conclude that at the 1% level of significance, there is insufficient evidence that the mean standby-time is less than 36h.

(a) Necessarily true. This is still a one-tailed test, with the same p-value. $p \le 0.05 \Rightarrow p < 0.10$, so evidence that is significant at the 5% level is also significant at the 10% level.

(b) Neither necessarily true nor necessarily false. This is now a two-tailed test, with twice the p-value. If $p \le 0.05$, then $2p$ may or may not be less than or equal to 0.05. (E.g. It is true if $p = 0.02$ but false if $p = 0.04$)

2. **[2018/H1 Math/A Level/Q11]** A website states that the mean length of adults in a particular species of fish is 30 cm. A scientist claims that the true mean length is greater than 30 cm. To test this claim, a random sample of 100 adult fish of this species is captured from a lake. The lengths of the fish, *x* cm, are summarised by

$$\sum(x-30)=15,\ \ \sum(x-30)^2=82$$

(i) Find unbiased estimates of the population mean and variance.

(ii) Test at the 2.5% significance level whether the scientist's claim is supported by the data.

(iii) State, with a reason, whether it is necessary to assume that the lengths of these fish are distributed normally for the test to be valid.

A new random sample of 100 adult fish of this species is captured from a different lake. The mean length of the fish in this sample is *m* cm and the population variance is 0.9 cm^2. A test at the 10% significance level supports the scientist's claim that the mean length of the fish is greater than 30 cm.

(iv) Find the range of possible values of *m*.

Solution:

(i) Unbiased estimate of the population mean, $\bar{x}=\frac{15}{100}+30=30.15$

Unbiased estimate of the pop. variance, $s^2=\frac{1}{99}\left[82-\frac{15^2}{100}\right]\approx 0.80556=0.806$ (3 s.f.)

(ii) Let *X* be the length (in cm) of an adult fish in this particular species and μ be the population mean length.

Test $\quad H_0:\mu=30 \quad$ vs. $\quad H_1:\mu>30 \quad$ at the 2.5% level of significance.

Under H_0, since $n=100$ is large, by Central Limit Theorem.

$\bar{X}\sim N(30,\dfrac{0.80556}{100})$ approximately

$Z=\dfrac{\bar{X}-30}{\sqrt{0.80556}/10}\sim N(0,1)$ approximately by Central Limit Theorem.

Using a one-tailed *z*-test, reject H_0 if $p\le 0.025$.

$\bar{x}=30.15$ gives *p*-value = 0.0473 > 0.025

$\therefore$ We do not reject H_0 and conclude that at the 2.5% significance level, there is insufficient evidence that the mean length of adult fish in this species is greater than 30cm. The scientist's claim is not supported by the data.

(iii) Not necessary since the sample size of 100 is large. By Central Limit theorem, the sample mean lengths of adult fish will be approximately normal.

(iv) Under H_0, since $n=100$ is large, by Central Limit Theorem.

$\bar{X}\sim N(30,\dfrac{0.9}{100})$

$Z=\dfrac{\bar{X}-30}{\sqrt{0.9}/10}\sim N(0,1)$

To support the scientist's claim (i.e. reject H_0), $z \geq 1.28155$.

$$\frac{m-30}{\sqrt{0.9}/10} \geq 1.28155$$

$$m \geq 30.122$$

$$m \geq 30.1 \text{ (3 s.f.)}$$

3. Under usual conditions, a certain type of seed produces fruit with mean mass 1.84 kg and standard deviation 0.13 kg. A crop of this fruit is grown under new conditions and each fruit in a random sample of 50 of the fruits produced is weighed. The mean mass of the sample is 1.88 kg.

(i) Test at the 3% significance level whether the data indicate the mean mass of this crop differs from 1.84 kg.

(ii) If the test had been for an increase in the mean at the 1% significance level, what would have been the result of the test?

(iii) State, giving a reason, whether it is necessary to assume that the mass of the fruit follows a normal distribution.

Solution:

Let X be the random variable "the mass of fruits" and μ denote "the population mean mass of fruits".

(i) To test $H_o: \mu = 1.84$ against $H_1: \mu \neq 1.84$

Using 2-tailed test at 3% significance level (i.e. $\alpha = 0.03$)

Under H_o, $\bar{X} \sim N(1.84, \frac{(0.13)^2}{50})$ approximately by Central Limit Theorem since the sample size = 50 is large.

Using GC, p–value = 0.0296. Since p-value < 0.03, we reject H_o and conclude that there is sufficient evidence that the population mean mass of the crop differs from 1.84 kg at 3% significance level.

(ii) To test $H_o: \mu = 1.84$ against $H_1: \mu > 1.84$

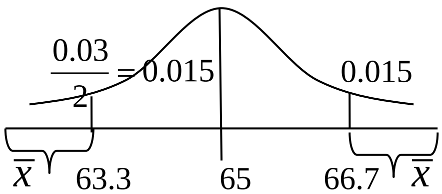
Using GC, p–value = 0.0148. Since the p–value > 0.01, we do not reject H_o and conclude that there is insufficient evidence that there is an increase in the population mean mass of the crop at 1% significance level.

(iii) Not necessary. Since the sample size is large, Central Limit Theorem applies to approximate the distribution of the sample mean to be normal.

4. A random variable X is known to have a normal distribution with variance 25. The mean of the distribution of X is denoted by μ. A random sample of 40 observations of X has mean $\bar{x}$. In a 2-tail test of the null hypothesis $\mu = 65$ at the 3% significance level, the alternative hypothesis is accepted. Find the range of values of $\bar{x}$.

Solution:

Test $\quad \mathrm{H}_0 : \mu = 65 \quad$ vs. $\quad \mathrm{H}_1 : \mu \neq 65 \quad$ at the 3% level of significance.



Under H_0, $\overline{X} \sim \mathrm{N}(65, \frac{25}{40})$

$$Z = \frac{\overline{X} - 65}{5/\sqrt{40}} \sim \mathrm{N}(0,1)$$

Using a two-tailed z-test, H_0 is rejected

$\Rightarrow \quad z \leq -2.1701 \quad$ or $\quad z \geq 2.1701$

$$\frac{\overline{x} - 65}{5/\sqrt{40}} \leq -2.1701 \quad \text{or} \quad \frac{\overline{x} - 65}{5/\sqrt{40}} \geq 2.1701$$

$\overline{x} \leq 63.3 \quad$ or $\quad \overline{x} \geq 66.7$ (3 s.f.)

5. **[N2008/H1/10]**

A consumer association is testing the lifetime of a particular type of battery that is claimed to have a lifetime of 150 hours. A random sample of 70 batteries of this type is tested and the lifetime, x hours, of each battery is measured. The results are summarised by

$$\sum x = 10317, \qquad \sum x^2 = 1540231.$$

The population mean lifetime is denoted by μ hours. The null hypothesis $\mu = 150$ is to be tested against the alternative hypothesis $\mu < 150$. Find the p-value of the test and state the meaning of this p-value in the context of the question.

A second random sample of 50 batteries of this type is tested and the lifetime, y hours, of each battery is measured, with results summarised by

$$\sum y = 7331, \ \sum y^2 = 1100565.$$

Combining the two samples into a single sample, carry out a test, at the 10% significance level, of the same null and alternative hypotheses.

Probing questions:

Q1. For the first part, what can you deduce about the value of $\overline{x}$, is this value greater or lesser than the population mean? Why?

Solution: The value of $\overline{x}$ is expected to be less than 150. We can know this as the alternative hypothesis is $\mu < 150$.

Q2. For the 2 samples, why we did not add up the population mean and variance and divide by 2? What are the essential conditions for this?

Solution: This method of solving is only true if the 2 samples consist of the same number of items. Hence, we need to add up the population mean and variance and divide by the total population so that the value that we have gotten will be the average.

Q3. For the 2 samples, what can you can deduce the p-value or z-value? Is this value larger or lesser than each sample p-value or z-value and why?

Solution: We cannot deduce anything for the 2 samples from each sample in terms of *p*-value or *z*-value as the population will be different. Hence, we need to recalculate again for *p*-value or *z*-value for 2 samples.

> Examiners' Report:
> (i) Most candidates were able to find the value of p, although many incorporated some premature approximation into their calculation.
> (ii) Few could interpret the meaning of p correctly, many giving an explanation of when to reject or accept the null hypothesis.
> (iii) The second part of this question was poorly answered. Most candidates added means and variances from two separate samples instead of combining the two samples into a single large sample. Others simply used only the second sample mean and variance.
> (iv) The final conclusion was often confused, with inconsistent statements such as 'p > 0.1 so do not reject the null hypothesis' and 'mean lifetime is less than 150'.

Solution:

Unbiased estimate of population mean, $\bar{x} = \dfrac{10317}{70} = 147.386$

Unbiased estimate of population variance, $s^2 = \dfrac{1}{69}\left[1540231 - \dfrac{10317^2}{70}\right] \approx 284.82$

Let X be the lifetime (in h) of a battery of the particular type and μ be the population mean lifetime.

Test $\mathrm{H}_0 : \mu = 150$ vs. $\mathrm{H}_1 : \mu < 150$ at the α % level of significance.

Under H_0, since n = 70 is large, by Central Limit Theorem.

$\bar{X} \sim \mathrm{N}(150, \dfrac{284.82}{70})$ approximately

$$Z = \frac{\bar{X} - \mu}{s/\sqrt{n}} = \frac{\bar{X} - 150}{\sqrt{284.82}/\sqrt{70}} \sim \mathrm{N}(0,1)$$

Using a one-tailed *z*-test,

$\bar{x} = 147.386$ gives *p*-value = 0.0975

p-value is the probability that the sample mean lifetime of battery in hours is less than 147.386, given that the population mean lifetime of battery in hours is 150.

For the combined sample,

Unbiased estimate of population mean $= \dfrac{\sum x + \sum y}{n_x + n_y} = \dfrac{10317 + 7331}{70 + 50} = 147.067$

Unbiased estimate of population variance

$$= \frac{1}{n-1}\left[\left(\sum x^2 + \sum y^2\right) - \frac{\left(\sum x + \sum y\right)^2}{n}\right] = \frac{1}{119}\left[2640796 - \frac{17648^2}{120}\right] = 381.206$$

where $n = n_x + n_y = 70 + 50 = 120$

Under H_0, since n = 120 is large, by Central Limit Theorem.

$\overline{X} \sim N(150, \frac{381.206}{120})$ approximately

Using a one-tailed z-test,
reject H_0 if $p \le 0.10$.

$\bar{x} = 147.067$ gives p-value = 0.0499 < 0.10

$\therefore$ We reject H_0 and conclude that at the 10% significance level, there is sufficient evidence that the population mean lifetime is less than 150 hours.

Practice Questions

1. State whether each of the following statements is True or False.

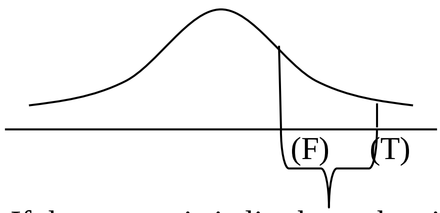
(a) If a test is significant at 5% level, then it must be significant at 1% level.

(b) When a statistician selects a particular level of significance for a test, he is effectively setting the probability of rejecting the null hypothesis when in fact it is true.

(c) One can never prove the truth of a statistical (null) hypothesis.

Solution:

(a) False. A statistic significant at 5% level is not necessarily significant at the 1% level (refer to diagram).



If the test statistic lies here, then it is significant at 5% but not at 1%.
Note: If the statement is reversed, then it is true (If a test is significant at the 1% level, then it must be significant at 5% level.).

(b) True by the definition of level of significance.

(c) True. In statistical inference there is always some probability, however small, of drawing the wrong conclusion. The fact that a hypothesis is consistent with a set of data does not mean that it is correct; at the same time, if it is not consistent with the data set it does not mean that it is incorrect.

2. **[N2011/2/Q10 Modified]**

In a factory, the time in minutes for an employee to install an electronic component is a normally-distributed continuous random variable T. The standard deviation of T is 5.0

and under ordinary conditions the expected value of T is 38.0. After background music is introduced into the factory, a sample of n components is taken and the mean time taken for randomly chosen employees to install them is found to be $\bar{t}$ minutes. A test is carried out, at the 5% significance level, to determine whether the mean time taken to install a component has been reduced.

(i) State appropriate hypotheses for the test, defining any symbols you use.

(ii) Given that $n = 50$, state the set of values of $\bar{t}$ for which the result of the test would be to reject the null hypothesis.

(iii) It is given instead that $\bar{t} = 37.1$, and the result of the test is that the null hypothesis is not rejected. Obtain an inequality involving n, and hence find the set of values that n can take.

Examiners' Report:

(i) Most candidates correctly define H_0 and H_1 in terms of μ. Although many candidates went on to explain what H_0 and H_1 were (which were not required), very few defined μ and a very small minority defined it correctly as the population mean time taken for employees to install one component.

(ii) A few candidates used 0.05 rather than a z-value. Others used +1.645 rather than -1.645. Still others used the wrong inequality or multiplied by a factor $\sqrt{\frac{n}{n-1}}$.

(iii) Candidates found this part rather challenging. Candidates made sign errors manipulating an inequality involving negatives and square root. For full credit, candidates were expected to realise that n was an integer.

Solution:

(i) Let μ be the population mean time taken for an employee to install one component.

Test $\mathrm{H_0} : \mu = 38.0$ vs. $\mathrm{H_1} : \mu < 38.0$ at the 5% level of significance.

(ii) Under $\mathrm{H_0}$, $Z = \dfrac{\overline{T} - 38}{5/\sqrt{50}} \sim \mathrm{N}(0,1)$

Reject $\mathrm{H_0}$ if $z \le -1.64485$.

$$\frac{\bar{t} - 38}{5/\sqrt{50}} \le -1.64485$$

$$\bar{t} \le 36.837$$

$$\therefore \left\{\bar{t} \in \mathbb{R} : \bar{t} \le 36.8\right\}$$

(iii) Now $Z = \dfrac{\overline{T} - 38}{5/\sqrt{n}} \sim \mathrm{N}(0,1)$

When $\bar{t} = 37.1$, H_0 is not rejected.

$$\therefore z > -1.64485$$

$$\frac{37.1 - 38}{5/\sqrt{n}} > -1.64485$$

$$0.54716 < \frac{5}{\sqrt{n}}$$

$$0.29939 < \frac{25}{n}$$

$$n < 83.5 \text{ (3 s.f.)}$$

$$\therefore \left\{ n \in \mathbb{Z}^+ : n \le 83 \right\}$$

3. [2010/IJC/II/11 (part)]

An engineer takes a sample of 20 guitar strings manufactured by a company, whose guitar strings have tensile strength that is normally distributed with mean μ kpsi (kilo-pounds per square inch) and standard deviation 4.7 kpsi. The null hypothesis $\mu = 430$ is being tested against the alternative hypothesis $\mu \ne 430$ at 5% level of significance. Find the range of values of the sample mean for which the null hypothesis is rejected, giving 2 decimal places in your answer.

Solution:

Let X be the tensile strength (in kpsi) of a guitar string manufactured by the company.

Test $H_0 : \mu = 430$

vs. $H_1 : \mu \ne 430$ at the 5% level of significance.

Under H_0, $Z = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \dfrac{\overline{X} - 430}{4.7/\sqrt{20}} \sim \mathrm{N}(0,1)$

Using a two-tailed z-test,

reject H_0 if $z \le -1.95996$ or $z \ge 1.95996$.

$$\frac{\bar{x} - 430}{4.7/\sqrt{20}} \le -1.95996 \quad \text{or} \quad \frac{\bar{x} - 430}{4.7/\sqrt{20}} \ge 1.95996$$

$$\bar{x} \le 427.94 \quad \text{or} \quad \bar{x} \ge 432.06$$

4. [2014/Prelim/AJC/Q12]

A company claims that the average lifetime of batteries it produces is 55 hours. A retailer suspects that the claim is inaccurate. To test the claim, the retailer tests a random sample of 100 batteries, and the lifetime x hours are summarized as follows:

$$\sum(x - 55) = -38, \ \sum(x - 55)^2 = 451.$$

(i) Find the unbiased estimates of the population mean and variance.

(ii) Test, at the 5% significance level, whether the company's claim is valid.

(iii) If the retailer has suspected that the average lifetime of batteries is less than 55 hours, state with a reason whether the conclusion of the test in part (ii) would remain the same.

An improvement is subsequently made to the manufacturing process and the company claims that the average lifetime is now μ_0 hours instead. A test on 8 batteries is conducted and the lifetime in hours, are recorded as follows:

56.3 54.8 54.5 54.4 53.9 55.5 54.6 54.9

It is known that the lifetime of batteries follows a Normal distribution with a standard deviation of 3.9.

(iv) Using the data from the sample of 8 batteries, find the least value of μ_0 to the nearest hour if there is evidence at 10% level of significance that the company has overstated the average lifetime of batteries.

Solution:

(i) unbiased estimate of the population mean $=\dfrac{-38}{100}+55=54.62$

unbiased estimate of the population variance $=\dfrac{1}{99}\left[451-\dfrac{(-38)^2}{100}\right]\approx 4.40969=4.41$ (3 s.f.)

(ii) Let X be the lifetime (in hours) of a battery and μ be the population mean lifetime.

Test $\mathrm{H}_0: \mu=55$ (company's claim)

vs. $\mathrm{H}_1: \mu\neq 55$ (retailer's suspicion) at the 5% level of significance.

Under H_0, since n = 100 is large,

$$Z=\frac{\bar{X}-\mu}{s/\sqrt{n}}=\frac{\bar{X}-55}{\sqrt{4.40969/10}}\sim \mathrm{N}(0,1) \text{ approximately by Central Limit Theorem.}$$

Using a two-tailed z-test, reject H_0 if $p\leq 0.05$.

$\bar{x}=54.62$ gives p-value = 0.0704 > 0.05

$\therefore$ We do not reject H_0 and conclude that at the 5% significance level, there is insufficient evidence that the company's claim is invalid.

(iii) For the one-tailed test,

$\bar{x}=54.62$ gives p-value = $\dfrac{0.0704}{2}=0.0352<0.05$

The conclusion would change. We now reject H_0 and conclude that at the 5% significance level, there is sufficient evidence that the company's claim is invalid.

(iv) Observed sample mean $\bar{x}=\dfrac{56.3+54.8+54.5+54.4+53.9+55.5+54.6+54.9}{8}$

$=54.8625$

Test $\quad H_0 : \mu = \mu_0$

vs. $\quad H_1 : \mu < \mu_0$ $\quad$ at the 10% level of significance.

Under H_0, $Z = \dfrac{\overline{X} - \mu_0}{3.9/\sqrt{8}} \sim N(0,1)$.

Using a one-tailed *z*-test, reject H_0 if $z \le -1.28155$.

If company has overstated the average lifetime, then H_0 is rejected.

$$\frac{54.8625 - \mu_0}{3.9/\sqrt{8}} \le -1.28155$$

$$\mu_0 \ge 56.6 \ \text{(3 s.f.)}$$

The least value of μ_0 is 57.

6.5 Summary and Checklist

Decision chart:

Population variance σ^2

- Known/ **Given** in question.
  - ***z*-test**
    - Population is **Normal**.

      Test statistic is $\overline{X}$.

      Under H_0, $\overline{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$
    - Distribution of population is not normal or **NOT** given and *n* is large

      Test statistic is $\overline{X}$.

      Under H_0, $\overline{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$ approximately by **<u>CLT</u>**.
- Unknown/ **NOT given** in question. Then we need to **calculate the unbiased estimate** of population variance, s^2.
  - **Large sample size *n***
    - ***z*-test**
      - Population is **Normal**.

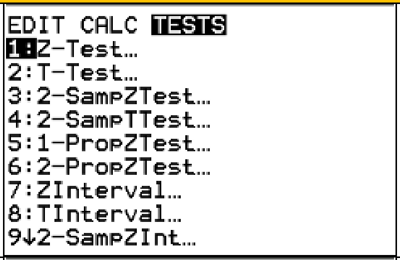
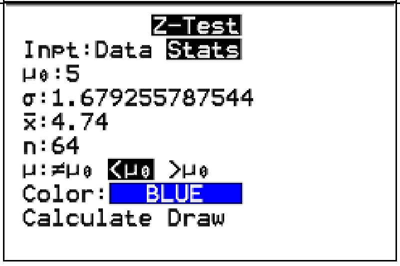
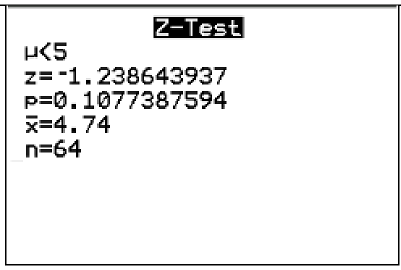
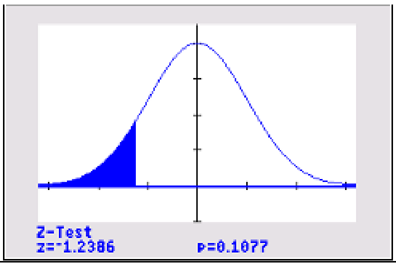
        Test statistic is $\overline{X}$.

        Under H_0, $\overline{X} \sim N\left(\mu, \dfrac{s^2}{n}\right)$
      - Distribution of population is not normal or **NOT** given.

        Test statistic is $\overline{X}$.

        Under H_0, $\overline{X} \sim N\left(\mu, \dfrac{s^2}{n}\right)$ approximately by **<u>CLT</u>**
  - **Small sample size *n***
    - ***Not in Syllabus***

Use of GC to calculate *p*-value:

Keystrokes	Screenshot	Remark
Press [Stat], and select [TESTS]. Select [1: *Z*-test].		
There are two ways to input values: Choose [Stats] if summary data (e.g. $\bar{x}$, σ and *n*) is given. Choose [Data] if raw data is given e.g. in a table.		μ_o = value of μ in H_0 σ = standard deviation of the population $\bar{x}$ = sample mean. n = sample size For two-tailed test, choose $\neq \mu_o$ For one-tailed test, choose either $>\mu_o$ or $<\mu_o$
Press [Calculate]: Or press [Draw] to obtain a standard normal curve with the shaded area equal to the *p*-value.	 	Note: *z* is the *z*-score *p* is the *p*-value

Concepts & Skills

I am able to:

- ☐ formulate the null hypothesis (H_0) and alternative hypothesis (H_1) for a given problem
- ☐ define following terms in the context of the given problem :
  - significance level
  - *p*-value
- ☐ perform a *z*-test when testing population mean in the following cases :
  - population is normal and variance is known
  - population is not normal and variance is known and sample size is large
  - population variance is unknown and sample size is large
- ☐ state the test statistic for the test & its distribution under H_0
- ☐ know how to carry out a hypothesis test using both the *p*-value method and critical value method
- ☐ interpret the results of a hypothesis test in the context of the problem