DP IB Maths: AI SL

Your notes

4.7 Hypothesis Testing

Contents

4.7.1 Hypothesis Testing

Your notes

Language of Hypothesis Testing

What is a hypothesis test?

- A hypothesis test uses a **sample of data** in an experiment to test a **statement** made about the **population**
  - The statement is either about a **population parameter** or the distribution of the **population**
- The hypothesis test will look at the probability of observed outcomes happening under set conditions
- The probability found will be compared against a given **significance level** to determine whether there is **evidence to support** the statement being made

What are the key terms used in statistical hypothesis testing?

- Every hypothesis test **must** begin with a clear **null hypothesis** (what we believe to already be true) and **alternative hypothesis** (how we believe the data pattern or probability distribution might have changed)
- A **hypothesis** is an assumption that is made about a particular **population parameter** or the **distribution of the population**
  - A **population parameter** is a numerical characteristic which helps define a population
    - Such as the mean value of the population
  - The **null hypothesis** is denoted H_0 and sets out the assumed population parameter or distribution given that no change has happened
  - The **alternative hypothesis** is denoted H_1 and sets out how we think the population parameter or distribution could have changed
    - A **one-tailed test** is used for testing the distribution or testing whether the parameter has increased (or decreased)
    - A **two-tailed test** is used for testing whether the parameter has changed (either increased or decreased)
  - When a hypothesis test is carried out, the null hypothesis is **assumed to be true** and this assumption will either be **accepted** or **rejected**
    - When a null hypothesis is accepted or rejected a **statistical inference** is made
- A hypothesis test will always be carried out at an appropriate **significance level**
  - The significance level sets the **smallest probability** that an event could have occurred by chance
    - Any probability smaller than the significance level would suggest that the event is unlikely to have happened by chance
  - The **significance level** must be set **before** the hypothesis test is carried out
  - The **significance level** will usually be 1%, 5% or 10%, however it may vary

Conclusions of Hypothesis Testing

Your notes

How do I decide whether to reject or accept the null hypothesis?

- A sample of the population is taken and the **test statistic** is calculated **using the observations** from the sample
  - Your GDC will calculate the test statistic for you
- To decide whether or not to reject the null hypothesis you first need either the ***p*-value** or the **critical region**
- The ***p* - value** is the probability of a value being **at least as extreme** as the test statistic, assuming that the null hypothesis is true
  - Your GDC will give you the *p*-value
  - If the ***p*-value is less than the significance level** then the **null hypothesis** would be **rejected**
- The **critical region** is the range of values of the test statistic which will lead to the **null hypothesis** being **rejected**
  - If the **test statistic** falls within the **critical region** then the **null hypothesis** would be **rejected**
- The **critical value** is the boundary of the critical region
  - It is the least extreme value that would lead to the rejection of the null hypothesis
  - The **critical value** is determined by the **significance level**
    - In your exam you will be given the critical value if it is needed

How should a conclusion be written for a hypothesis test?

- Your conclusion **must** be written in the **context** of the question
- Use the **wording in the question** to help you write your conclusion
  - If **rejecting the null** hypothesis your conclusion should state that there is **sufficient evidence** to suggest that the null hypothesis is unlikely to be true
  - If **accepting the null** hypothesis your conclusion should state that there is **not enough evidence** to suggest that the null hypothesis is unlikely to be true
- Your conclusion **must not** be definitive
  - There is a chance that the test has led to an **incorrect conclusion**
  - The outcome is **dependent on the sample**
    - a **different sample** might lead to a **different outcome**
- The conclusion of a two-tailed test can state if there is evidence of a change
  - You should not state whether this change is an increase or decrease

Examiner Tip

- Accepting the null hypothesis does **not** mean that you are saying it is true
  - You are simply saying there is not enough evidence to reject it

Your notes

4.7.2 Chi-squared Test for Independence

Chi-Squared Test for Independence

What is a chi-squared test for independence?

- A chi-squared (χ^2) **test for independence** is a hypothesis test used to test whether **two variables are independent** of each other
  - This is sometimes called a χ^2 **two-way test**
- This is an example of a **goodness of fit** test
  - We are testing whether the data fits the model that the variables are independent
- The chi-squared (χ^2) distribution is used for this test
- You will use a **contingency table**
  - This is a two-way table that shows the **observed frequencies** for the different combinations of the two variables
    - For example: if the two variables are hair colour and eye colour then the contingency table will show the frequencies of the different combinations

What are the degrees of freedom?

- There will be a **minimum number of expected values** you would need to know in order to be able to calculate all the expected values
- This minimum number is called the **degrees of freedom** and is often denoted by ν
- For a **test for independence** with an $m \times n$ contingency table
  - $\nu = (m-1) \times (n-1)$
  - For example: If there are 5 rows and 3 columns then you only need to know **2 of the values** in **4 of the rows** as the rest can be calculated using the totals

What are the steps for a chi-squared test for independence?

- **STEP 1**: Write the **hypotheses**
  - H_0 : Variable X is independent of variable Y
  - H_1 : Variable X is not independent of variable Y
    - Make sure you clearly write what the variables are and don't just call them X and Y
- **STEP 2**: Calculate the **degrees of freedom** for the test
  - For an $m \times n$ contingency table
  - Degrees of freedom is $\nu = (m-1) \times (n-1)$
- **STEP 3**: Enter your **observed frequencies** into your GDC using the option for a 2-way test
  - Enter these as a matrix
  - Your GDC will give you a matrix of the **expected values** (assuming the variables are independent)
  - Your GDC will also give you the χ^2 statistic and its p-value
  - The χ^2 statistic is denoted as χ^2_{calc}
- **STEP 4**: Decide whether there is evidence to reject the null hypothesis

Your notes

- EITHER compare the χ^2 **statistic** with the given **critical value**
  - If χ^2 statistic **>** critical value then **reject** H_0
  - If χ^2 statistic **<** critical value then **accept** H_0
- OR compare the ***p*-value** with the given **significance level**
  - If *p*-value **<** significance level then **reject** H_0
  - If *p*-value **>** significance level then **accept** H_0
- **STEP 5**: Write your **conclusion**
  - If you **reject** H_0
    - There is sufficient evidence to suggest that variable *X* is not independent of variable *Y*
    - Therefore this suggests they are **associated**
  - If you **accept** H_0
    - There is insufficient evidence to suggest that variable *X* is not independent of variable *Y*
    - Therefore this suggests they are **independent**

How do I calculate the chi-squared statistic?

- You are **expected** to be able to use your **GDC** to calculate the χ^2 statistic by inputting the matrix of the observed frequencies
- Seeing how it is done by hand might deepen your understanding but you are **not expected** to use this method
- **STEP 1**: For each **observed frequency** O_i calculate its **expected frequency** E_i
  - Assuming the variables are independent
    - $E_i = \text{P}(X = x) \times \text{P}(Y = y) \times \text{Total}$
    - Which simplifies to $E_i = \frac{\text{Row Total} \times \text{Column Total}}{\text{Overall Total}}$
- **STEP 2**: Calculate the χ^2 statistic using the formula
  - $$\chi^2_{calc} = \sum \frac{(O_i - E_i)^2}{E_i}$$
  - You do not need to learn this formula as your GDC calculates it for you
- To calculate the *p*-value you would find the probability of a value being bigger than your χ^2 statistic using a χ^2 distribution with ν degrees of freedom

Examiner Tip

Note for Internal Assessments (IA)

- If you use a χ^2 test in your IA then beware that the outcome may not be accurate if:
  - Any of the **expected values are less than 5**
  - There is **only 1 degree of freedom**
    - This means it is a 2×2 contingency table
- Note that **none of these cases will occur in the exam**

Your notes

Worked example

At a school in Paris, it is believed that favourite film genre is related to favourite subject. 500 students were asked to indicate their favourite film genre and favourite subject from a selection and the results are indicated in the table below.

	Comedy	Action	Romance	Thriller
Maths	51	52	37	55
Sports	59	63	41	33
Geography	35	31	28	15

It is decided to test this hypothesis by using a χ^2 test for independence at the 1% significance level.

The critical value is 16.812.

a) State the null and alternative hypotheses for this test.

H_0 : Favourite subject is independent of favourite film genre
H_1 : Favourite subject is not independent of favourite film genre

b) Write down the number of degrees of freedom for this table.

$\nu = (\text{rows} - 1) \times (\text{columns} - 1) = (3-1) \times (4-1)$

$\nu = 6$

c) Calculate the χ^2 test statistic for this data.

Your notes

Type matrix into GDC

χ^2 statistic $= 12.817...$

$\chi^2_{calc} = 12.8$ (3 sf)

d) Write down the conclusion to the test. Give a reason for your answer.

$12.8 < 16.812$

Accept H_0 as χ^2 statistic < critical value.
There is insufficient evidence to suggest that favourite subject is not independent of favourite film genre. Therefore this suggests they are independent.

4.7.3 Goodness of Fit Test

Your notes

Chi-Squared GOF: Uniform

What is a chi-squared goodness of fit test for a given distribution?

- A chi-squared (χ^2) **goodness of fit test** is used to test data from a sample which suggests that the population has a given distribution
- This could be that:
  - the proportions of the population for different categories follows a **given ratio**
  - the population follows a **uniform distribution**
    - This means all outcomes are **equally likely**

What are the steps for a chi-squared goodness of fit test for a given distribution?

- **STEP 1**: Write the **hypotheses**
  - H_0 : Variable *X* can be modelled by the given distribution
  - H_1 : Variable *X* cannot be modelled by the given distribution
    - Make sure you clearly write what the variable is and don't just call it *X*
- **STEP 2**: Calculate the **expected frequencies**
  - Split the total frequency using the given ratio
  - For a uniform distribution: divide the total frequency *N* by the number of possible outcomes *k*
- **STEP 3**: Calculate the **degrees of freedom** for the test
  - For *k* possible outcomes
  - Degrees of freedom is $v = k - 1$
- **STEP 4**: Enter the **frequencies** and the **degrees of freedom** into your GDC
  - Enter the observed and expected frequencies as two separate lists
  - Your GDC will then give you the χ^2 statistic and its *p*-value
  - The χ^2 statistic is denoted as χ^2_{calc}
- **STEP 5**: Decide **whether** there is evidence to **reject the null** hypothesis
  - EITHER compare the χ^2 **statistic** with the given **critical value**
    - If χ^2 statistic **>** critical value then **reject** H_0
    - If χ^2 statistic **<** critical value then **accept** H_0
  - OR compare the ***p*-value** with the given **significance level**
    - If *p*-value **<** significance level then **reject** H_0
    - If *p*-value **>** significance level then **accept** H_0
- **STEP 6**: Write your **conclusion**
  - If you **reject** H_0
    - There is sufficient evidence to suggest that variable *X* does not follow the given distribution
    - Therefore this suggests that the data is **not distributed as claimed**
  - If you **accept** H_0
    - There is insufficient evidence to suggest that variable *X* does not follow the given distribution
    - Therefore this suggests that the data is **distributed as claimed**

Your notes

Worked example

A car salesman is interested in how his sales are distributed and records his sales results over a period of six weeks. The data is shown in the table.

Week	1	2	3	4	5	6
Number of sales	15	17	11	21	14	12

A χ^2 goodness of fit test is to be performed on the data at the 5% significance level to find out whether the data fits a uniform distribution.

a) Find the expected frequency of sales for each week if the data were uniformly distributed.

If uniformly distributed all expected frequencies are equal

Expected frequency $= \frac{15+17+11+21+14+12}{6}$

Expected frequency $= 15$

b) Write down the null and alternative hypotheses.

H_0 : Number of sales can be modelled by a uniform distribution

H_1 : Number of sales can not be modelled by a uniform distribution

c) Write down the number of degrees of freedom for this test.

$\nu = 6 - 1$

$\nu = 5$

d) Calculate the p-value.

Your notes

Type two lists into GDC

Observed	15	17	11	21	14	12
Expected	15	15	15	15	15	15

$p = 0.4933...$

$p = 0.493$ (3sf)

e) State the conclusion of the test. Give a reason for your answer.

$0.493 > 0.05$

Accept H_0 as p-value > significance level
There is insufficient evidence to suggest that number of sales can not be modelled by a uniform distribution. Therefore this suggests it is uniformly distributed.

Your notes

Chi-Squared GOF: Binomial

What is a chi-squared goodness of fit test for a binomial distribution?

- A chi-squared (χ^2) **goodness of fit test** is used to test data from a sample suggesting that the population has a **binomial distribution**
  - You will be given the value of *p* for the binomial distribution

What are the steps for a chi-squared goodness of fit test for a binomial distribution?

- **STEP 1**: Write the **hypotheses**
  - H_0 : Variable *X* can be modelled by the binomial distribution $\text{B}(n, p)$
  - H_1 : Variable *X* cannot be modelled by the binomial distribution $\text{B}(n, p)$
    - Make sure you clearly write what the variable is and don't just call it *X*
    - State the values of *n* and *p* clearly
- **STEP 2**: Calculate the **expected frequencies**
  - Find the probability of the outcome using the binomial distribution $\text{P}(X = x)$
  - Multiply the probability by the total frequency $\text{P}(X = x) \times N$
- **STEP 3**: Calculate the **degrees of freedom** for the test
  - For *k* outcomes
  - Degrees of freedom is $\nu = k - 1$
- **STEP 4**: Enter the **frequencies** and the **degrees of freedom** into your GDC
  - Enter the observed and expected frequencies as two separate lists
  - Your GDC will then give you the χ^2 statistic and its *p*-value
  - The χ^2 statistic is denoted as χ^2_{calc}
- **STEP 5**: Decide whether there is **evidence** to **reject the null** hypothesis
  - EITHER compare the **χ^2 statistic** with the given **critical value**
    - If χ^2 statistic **>** critical value then **reject H_0**
    - If χ^2 statistic **<** critical value then **accept H_0**
  - OR compare the *p*-value with the given significance level
    - If *p*-value **<** significance level then **reject H_0**
    - If *p*-value **>** significance level then **accept H_0**
- **STEP 6**: Write your **conclusion**
  - If you **reject H_0**
    - There is sufficient evidence to suggest that variable *X* does not follow the binomial distribution $\text{B}(n, p)$
    - Therefore this suggests that the data **does not follow** $\text{B}(n, p)$
  - If you **accept H_0**
    - There is insufficient evidence to suggest that variable *X* does not follow the binomial distribution $\text{B}(n, p)$
    - Therefore this suggests that the data **follows** $\text{B}(n, p)$

Your notes

Worked example

A stage in a video game has three boss battles. 1000 people try this stage of the video game and the number of bosses defeated by each player is recorded.

Number of bosses defeated	0	1	2	3
Frequency	490	384	111	15

A χ^2 goodness of fit test at the 5% significance level is used to decide whether the number of bosses defeated can be modelled by a binomial distribution with a 20% probability of success.

a) State the null and alternative hypotheses.

H_0: Number of bosses defeated can be modelled by the binomial distribution B(3, 0.2)
H_1: Number of bosses defeated can not be modelled by the binomial distribution B(3, 0.2)

b) Assuming the binomial distribution holds, find the expected number of people that would defeat exactly one boss.

Let $X \sim B(3, 0.2)$
Using GDC $P(X = 1) = 0.384$
Expected $1000 \times 0.384 = 384$

Expected frequency of 1 $= 384$

c) Calculate the p-value for the test.

Your notes

Find the other expected frequencies

For 0: $1000 \times P(X=0) = 1000 \times 0.512 = 512$

For 2: $1000 \times P(X=2) = 1000 \times 0.096 = 96$

For 3: $1000 \times P(X=3) = 1000 \times 0.008 = 8$

Type two lists into GDC

Observed	490	384	111	15
Expected	512	384	96	8

$\nu = 4 - 1 = 3$

$p = 0.02426\ldots$

$p = 0.0243$ (3sf)

d) State the conclusion of the test. Give a reason for your answer.

$0.0243 < 0.05$

Reject H_0 as p-value < significance level. There is sufficient evidence to suggest that the number of bosses defeated can not be modelled by the binomial distribution $B(3, 0.2)$.

Your notes

Chi-Squared GOF: Normal

What is a chi-squared goodness of fit test for a normal distribution?

- A chi-squared (χ^2) **goodness of fit test** is used to test data from a sample suggesting that the population has a **normal distribution**
  - You will be given the value of μ and σ for the normal distribution

What are the steps for a chi-squared goodness of fit test for a normal distribution?

- **STEP 1**: Write the **hypotheses**
  - H_0 : Variable X can be modelled by the normal distribution $N(\mu, \sigma^2)$
  - H_1 : Variable X cannot be modelled by the normal distribution $N(\mu, \sigma^2)$
    - Make sure you clearly write what the variable is and don't just call it X
    - State the values of μ and σ clearly
- **STEP 2**: Calculate the **expected frequencies**
  - Find the probability of the outcome using the normal distribution $P(a < X < b)$
    - Beware of unbounded inequalities $P(X < b)$ or $P(X > a)$ for the class intervals on the 'ends'
  - Multiply the probability by the total frequency $P(a < X < b) \times N$
- **STEP 3**: Calculate the **degrees of freedom** for the test
  - For k class intervals
  - Degrees of freedom is $\nu = k - 1$
- **STEP 4**: Enter the **frequencies** and the **degrees of freedom** into your GDC
  - Enter the observed and expected frequencies as two separate lists
  - Your GDC will then give you the χ^2 statistic and its p-value
  - The χ^2 statistic is denoted as χ^2_{calc}
- **STEP 5**: Decide whether there is **evidence** to **reject the null** hypothesis
  - EITHER compare the **χ^2 statistic** with the given **critical value**
    - If χ^2 statistic **>** critical value then **reject** H_0
    - If χ^2 statistic **<** critical value then **accept** H_0
  - OR compare the ***p*-value** with the given **significance level**
    - If p-value **<** significance level then **reject** H_0
    - If p-value **>** significance level then **accept** H_0
- **STEP 6**: Write your **conclusion**
  - If you **reject** H_0
    - There is sufficient evidence to suggest that variable X does not follow the normal distribution $N(\mu, \sigma^2)$
    - Therefore this suggests that the data **does not follow** $N(\mu, \sigma^2)$
  - If you **accept** H_0

- There is insufficient evidence to suggest that variable *X* does not follow the normal distribution $\mathrm{N}(\mu, \sigma^2)$
- Therefore this suggests that the data **follows** $\mathrm{N}(\mu, \sigma^2)$

Your notes

Your notes

Worked example

300 marbled ducks in Quacktown are weighed and the results are shown in the table below.

Mass (g)	Frequency
$m < 470$	10
$470 \le m < 520$	158
$520 \le m < 570$	123
$m \ge 570$	9

A χ^2 goodness of fit test at the 10% significance level is used to decide whether the mass of a marbled duck can be modelled by a normal distribution with mean 520 g and standard deviation 30 g.

a) Calculate the expected frequencies, giving your answers correct to 2 decimal places.

Let $X \sim N(520, 30^2)$

300 × probability

Mass (g)	Probability	Expected frequency
$m < 470$	0.047790...	14.34
$470 \le m < 520$	0.452209...	135.66
$520 \le m < 570$	0.452209...	135.66
$m \ge 570$	0.047790...	14.34

b) Write down the null and alternative hypotheses.

H_0: Mass of the marbled ducks can be modelled by the normal distribution $N(520, 30^2)$

H_1: Mass of the marbled ducks can not be modelled by the normal distribution $N(520, 30^2)$

c) Calculate the χ^2 statistic.

Enter the observed and expected frequencies into GDC $\nu = 4 - 1 = 3$

χ^2 statistic $= 8.162 \ldots$

$\chi^2_{calc} = 8.16$ (3sf)

d) Given that the critical value is 6.251, state the conclusion of the test. Give a reason for your answer.

$8.16 > 6.251$

Reject H_0 as χ^2 statistic > critical value.
There is sufficient evidence to suggest that the mass of the marbled ducks can not be modelled by the normal distribution $N(520, 30^2)$.

Your notes

Your notes

4.7.4 The t-test

Two-Sample Tests

What is a *t*-test?

- A *t*-test is used to **compare the means** of **two normally distributed** populations
- In the exam the **population variance will always be unknown**

What assumptions are needed for the *t*-test?

- The underlying distribution for **each variable** must be **normal**
- In the exam you will need to assume the variance for the two groups are equal
  - You will need to use the **pooled two-sample** *t*-test

What are the steps for a pooled two-sample *t*-test?

- **STEP 1**: Write the **hypotheses**
  - $H_0 : \mu_x = \mu_y$
    - Where μ_x and μ_y are the **population means**
    - Make sure you make it clear which mean corresponds to each population
    - In words this means the **two population means are equal**
  - $H_1 : \mu_x < \mu_y$ or $H_1 : \mu_x > \mu_y$ or $H_1 : \mu_x \neq \mu_y$
    - The alternative hypothesis will depend on what is being tested (see sections for one-tailed and two-tailed tests)
- **STEP 2**: Enter the data into your GDC
  - Enter two lists of data – one for each sample
  - Choose the pooled option
  - Your GDC will then give you the *p*-value
- **STEP 3**: Decide whether there is **evidence to reject** the **null** hypothesis
  - Compare the ***p*-value** with the given **significance level**
    - If *p*-value < significance level then **reject H_0**
    - If *p*-value > significance level then **accept H_0**
- **STEP 4**: Write your **conclusion**
  - If you **reject H_0**
    - There is sufficient evidence to suggest that the population mean of X is bigger than/smaller than/different to the population mean of Y
    - This will depend on the alternative hypothesis
  - If you **accept H_0**
    - There is insufficient evidence to suggest that the population mean of X is bigger than/small than/different to the population mean of Y
    - Therefore this suggests that the **population means are equal**

One-tailed Tests

Your notes

How do I perform a one-tailed *t*-test?

- A **one-tailed test** is used to test one of the two following cases:
  - The population mean of X is **bigger** than the population mean of Y
    - The alternative hypothesis will be: $H_1 : \mu_x > \mu_y$
    - Look out for words such as increase, bigger, higher, etc
  - The population mean of X is **smaller** than the population mean of Y
    - The alternative hypothesis will be: $H_1 : \mu_x < \mu_y$
    - Look out for words such as decrease, smaller, lower, etc
- If you **reject the null** hypothesis then
  - This suggests that the population mean of X is **bigger** than the population mean of Y
    - If the alternative hypothesis is $H_1 : \mu_x > \mu_y$
  - This suggests that the population mean of X is **smaller** than the population mean of Y
    - If the alternative hypothesis is $H_1 : \mu_x < \mu_y$

Your notes

Worked example

The times (in minutes) for children and adults to complete a puzzle are recorded below.

Children	3.1	2.7	3.5	3.1	2.9	3.2	3.0	2.9		
Adults	3.1	3.6	3.5	3.6	2.9	3.6	3.4	3.6	3.7	3.0

The creator of the puzzle claims children are generally faster at solving the puzzle than adults. A t-test is to be performed at a 1% significance level.

a) Write down the null and alternative hypotheses.

Let μ_c be the population mean for children's times

and μ_A be the population mean for adults' times

$H_0: \mu_c = \mu_A$

$H_1: \mu_c < \mu_A$

It is claimed that children are quicker

b) Find the p-value for this test.

Enter the data as two lists in GDC

Use 2-sample pooled t-test

$p = 0.007259...$

$p = 0.00726$ (3sf)

c) State whether the creator's claim is supported by the test. Give a reason for your answer.

Your notes

$0.00726 < 0.01$

Reject H_0 as p-value < significance level. There is sufficient evidence to suggest that children are generally faster at solving the puzzle than adults. This supports the creator's claim.

Two-tailed Tests

Your notes

How do I perform a two-tailed *t*-test?

- A **two-tailed test** is used to test the following case:
  - The population mean of *X* is **different** to the population mean of *Y*
    - The alternative hypothesis will be: $H_1 : \mu_x \neq \mu_y$
    - Look out for words such as change, different, not the same, etc
- If you **reject the null** hypothesis then
  - This suggests that the population mean of *X* is **different** to the population mean of *Y*
  - You can not state which one is bigger as you were not testing for that
    - All you can conclude is that there is evidence that the means are not equal
    - To test whether a specific one is bigger you would need to use a one-tailed test

Your notes

Worked example

In a school all students must study either French or Spanish as well as maths. 18 students in a maths class complete a test and their scores are recorded along with which language they study.

Studies French	61	82	77	80	99	69	75	71	81
Studies Spanish	74	79	83	66	95	79	82	81	85

The maths teacher wants to investigate whether the scores are different between the students studying each language. A t-test is to be performed at a 10% significance level.

a) Write down the null and alternative hypotheses.

Let μ_F be the population mean for scores of students of French and μ_S be the population mean for scores of students of Spanish

$H_0: \mu_F = \mu_S$
$H_1: \mu_F \neq \mu_S$

Testing for a difference

b) Find the p-value for this test.

Enter the data as two lists in GDC
Use 2-sample pooled t-test
$p = 0.47391...$

$p = 0.474$ (3sf)

c) Write down the conclusion to the test. Give a reason for your answer.

$0.474 > 0.1$

Accept H_0 as p-value > significance level.
There is insufficient evidence to suggest that there is a difference between the scores.