DP IB Maths: AI HL

4.12 Further Hypothesis Testing

Contents

4.12.1 Hypothesis Testing for Mean (One Sample)

One-Sample z-tests

What is a one-sample *z*-test?

- A **one-sample *z*-test** is used to **test the mean (μ)** of a **normally distributed** population
  - You use a *z*-test when the **population variance (σ^2) is known**
- The **mean of a sample** of size *n* is calculated $\overline{X}$ and a normal distribution is used to test the **test statistic**
- $\overline{X}$ can be used as the test statistic
  - In this case you would use the distribution $\overline{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$
    - Remember when using this distribution that the standard deviation is $\frac{\sigma}{\sqrt{n}}$
- $z = \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$ can be used as the test statistic
  - In this case you would use the distribution $Z \sim \mathrm{N}(0, 1^2)$
    - This is a more old-fashioned approach but your GDC still might tell you the *z*-value when you do the test
    - You will not need to use this method in the exam as your GDC should be capable of doing the other method

What are the steps for performing a one-sample *z*-test on my GDC?

- **STEP 1**: **Write the hypotheses**
  - $H_0 : \mu = \mu_0$
    - Clearly state that μ represents the **population mean**
    - μ_0 is the **assumed population mean**
  - For a **one-tailed** test $H_1 : \mu < \mu_0$ or $H_1 : \mu > \mu_0$
  - For a **two-tailed** test: $H_1 : \mu \neq \mu_0$
    - The alternative hypothesis will depend on what is being tested
- **STEP 2**: **Enter the data** into your GDC and choose the **one-sample *z*-test**
  - If you have the raw data
    - Enter the data as a list
    - Enter the value of σ
  - If you have summary statistics
    - Enter the values of $\overline{X}$, σ and n
  - Your GDC will give you the *p*-value
- **STEP 3**: **Decide** whether there is **evidence to reject the null hypothesis**

- If the p-value < significance level then reject H_0
- **STEP 4**: **Write your conclusion**
  - If you **reject H_0** then there is **evidence** to suggest that...
    - The mean has decreased (for $H_1 : \mu < \mu_0$)
    - The mean has increased (for $H_1 : \mu > \mu_0$)
    - The mean has changed (for $H_1 : \mu \neq \mu_0$)
  - If you **accept H_0** then there is **insufficient evidence** to reject the null hypothesis which suggests that...
    - The mean has not decreased (for $H_1 : \mu < \mu_0$)
    - The mean has not increased (for $H_1 : \mu > \mu_0$)
    - The mean has not changed (for $H_1 : \mu \neq \mu_0$)

How do I find the *p*-value for a one-sample *z*-test using a normal distribution?

- The p-value is determined by the **test statistic** $\overline{X}$
- For $H_1 : \mu < \mu_0$ the p-value is $\mathrm{P}\left(\overline{X} < \overline{x} \mid \mu = \mu_0\right)$
- For $H_1 : \mu > \mu_0$ the p-value is $\mathrm{P}\left(\overline{X} > \overline{x} \mid \mu = \mu_0\right)$
- For $H_1 : \mu \neq \mu_0$ the p-value is $\mathrm{P}\left(\left|\overline{X} - \mu_0\right| > \left|x - \mu_0\right| \mid \mu = \mu_0\right)$
  - If $\overline{x} < \mu_0$ then this can be calculated easier by $2 \times \mathrm{P}\left(\overline{X} < \overline{x} \mid \mu = \mu_0\right)$
  - If $\overline{x} > \mu_0$ then this can be calculated easier by $2 \times \mathrm{P}\left(\overline{X} > \overline{x} \mid \mu = \mu_0\right)$

How do I find the critical value and critical region for a one-sample *z*-test?

- The critical region is determined by the **significance level** $\alpha\%$
  - For $H_1 : \mu < \mu_0$ the critical region is $\overline{X} < c$ where $\mathrm{P}\left(\overline{X} < c \mid \mu = \mu_0\right) = \alpha\%$
  - For $H_1 : \mu > \mu_0$ the critical region is $\overline{X} > c$ where $\mathrm{P}\left(\overline{X} > c \mid \mu = \mu_0\right) = \alpha\%$
  - For $H_1 : \mu \neq \mu_0$ the critical regions are $\overline{X} < c_1$ and $\overline{X} > c_2$ where

    $$\mathrm{P}\left(\overline{X} < c_1 \mid \mu = \mu_0\right) = \mathrm{P}\left(\overline{X} > c_2 \mid \mu = \mu_0\right) = \frac{1}{2}\alpha\%$$
- The critical value(s) can be found using the inverse normal distribution function
  - When rounding the critical value(s) you should choose:
    - The **lower bound** for the inequalities $\overline{X} < c$
    - The **upper bound** for the inequalities $\overline{X} > c$
  - This is so that the probability **does not exceed the significance level**

Worked example

The mass of a Burmese cat, C, follows a normal distribution with mean 4.2 kg and a standard deviation 1.3 kg. Kamala, a cat breeder, claims that Burmese cats weigh more than the average if they live in a household which contains young children. To test her claim, Kamala takes a random sample of 25 cats that live in households containing young children.

a) State the null and alternative hypotheses to test Kamala's claim.

Let μ be the population mean for the mass of Burmese cats

$$\boxed{\begin{array}{l} H_0: \mu = 4.2 \\ H_1: \mu > 4.2 \end{array}}$$

Testing for an increase

b) Using a 5% level of significance, find the critical region for this test.

The population variance is known so use a z-test

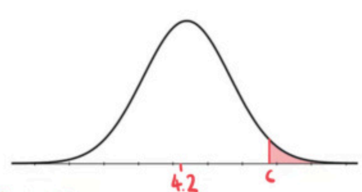
Find the distribution of the sample means $N\left(\mu, \frac{\sigma^2}{n}\right)$

$\bar{C} \sim N\left(\mu, \frac{1.3^2}{25}\right)$

Square the standard deviation

The sample size

Critical region is $\bar{C} > c$ where $P(\bar{C} > c \mid \mu = 4.2) = 0.05$



Use inverse normal:

$\mu = 4.2$

$\sigma = \sqrt{\frac{1.3^2}{25}} = 0.26$

$P(\bar{C} < c) = 0.95$

$c = 4.6276...$

$$\boxed{\text{Critical region } \bar{C} > 4.63}$$

c) Kamala calculates the mean of the 25 cats included in her sample to be 4.65 kg. Determine the conclusion of the test.

$4.65 > 4.6276...$ so 4.65 is in critical region

Reject H_0 as test statistic is in critical region. There is sufficient evidence to suggest that Burmese cats weigh more if they live in a household which contains young children.

One-Sample t-tests

What is a one-sample *t*-test?

- A **one-sample *t*-test** is used to **test the mean (μ)** of a **normally distributed** population
  - You use a *t*-test when the **population variance (σ^2) is unknown**
  - You need to use the unbiased estimate for the population variance (s_{n-1}^2)
- The **mean of a sample** of size *n* is calculated $\overline{X}$ and a *t*-distribution is used to test it
  - *t*-distributions are similar to normal distributions
    - As the sample size gets larger the *t*-distribution tends towards the standard normal distribution
- You won't be expected to find the critical value
  - The *p*-value can be found using the test function on your GDC

What are the steps for performing a one-sample *t*-test on my GDC?

- **STEP 1**: **Write the hypotheses**
  - $H_0 : \mu = \mu_0$
    - Clearly state that μ represents the **population mean**
    - μ_0 is the **assumed population mean**
  - For a **one-tailed** test $H_1 : \mu < \mu_0$ or $H_1 : \mu > \mu_0$
  - For a **two-tailed** test: $H_1 : \mu \neq \mu_0$
    - The alternative hypothesis will depend on what is being tested
- **STEP 2**: **Enter the data** into your GDC and choose the **one-sample *t*-test**
  - If you have the raw data
    - Enter the data as a list
  - If you have summary statistics
    - Enter the values of $\overline{X}$, s_{n-1} (sometimes written as s_x on a GDC) and *n*
  - Your GDC will give you the *p*-value
- **STEP 3**: **Decide** whether there is **evidence to reject the null hypothesis**
  - If the *p*-value < significance level then reject H_0
- **STEP 4**: **Write your conclusion**
  - If you **reject H_0** then there is **evidence** to suggest that...
    - The mean has decreased (for $H_1 : \mu < \mu_0$)
    - The mean has increased (for $H_1 : \mu > \mu_0$)
    - The mean has changed (for $H_1 : \mu \neq \mu_0$)
  - If you **accept H_0** then there is **insufficient evidence** to reject the null hypothesis which suggests that...
    - The mean has not decreased (for $H_1 : \mu < \mu_0$)
    - The mean has not increased (for $H_1 : \mu > \mu_0$)
    - The mean has not changed (for $H_1 : \mu \neq \mu_0$)

Worked example

The IQ of a student at Calculus High can be modelled as a normal distribution with mean 126. The headteacher decides to play classical music during lunchtimes and suspects that this has caused a change in the average IQ of the students.

a) State the null and alternative hypotheses to test the headteacher's suspicion.

Let μ be the population mean for the IQ of a student at Calculus High

$\boxed{\begin{array}{l} H_0: \mu = 126 \\ H_1: \mu \neq 126 \end{array}}$ Testing for a change

b) The headteacher selects 15 students and asks them to complete an IQ test. The mean score for the sample is 127.1 and the sample variance is 14.7. Calculate the unbiased estimate for the population variance s^2_{n-1}.

Formula booklet

Unbiased estimate of population variance s^2_{n-1}	$s^2_{n-1} = \frac{n}{n-1} s^2_n$

$s^2_{n-1} = \frac{15}{14} \times 14.7$

$\boxed{s^2_{n-1} = 15.75}$

c) Calculate the p-value for the test.

The population variance is unknown so use a t-test

Enter summary statistics into GDC using one-sample t-test

$\bar{x} = 127.1 \quad s_{n-1} = \sqrt{15.75} \quad n = 15$

$p = 0.3012...$

$\boxed{p = 0.301 \text{ (3sf)}}$

d) State whether the headteacher's suspicion is supported by the test.

$0.3012... > 0.1$

Accept H_0 as p-value > significance level.
There is insufficient evidence to support the headteacher's suspicion.

4.12.2 Hypothesis Testing for Mean (Two Sample)

Two-Sample Tests

What is a two-sample test?

- A **two-sample test** is used to **compare the means (μ_1 & μ_2)** of **two normally distributed** populations
  - You use a ***z*-test** when the **population variances (σ_1^2 & σ_2^2) are known**
  - You use a ***t*-test** when the **population variances are unknown**
    - In this course you will assume the **variances are equal** and use a **pooled sample** for a *t*-test
    - In a pooled sample the data from both samples are used to estimate the population variance

What are the steps for performing a two-sample test on my GDC?

- **STEP 1**: **Write the hypotheses**
  - $H_0 : \mu_1 = \mu_2$
    - Clearly state that μ_1 & μ_2 represent the **population means**
    - Make sure you make it clear which mean corresponds to which population
    - In words this means that the two population means are equal
  - For a **one-tailed** test $H_1 : \mu_1 < \mu_2$ or $H_1 : \mu_1 > \mu_2$
  - For a **two-tailed** test: $H_1 : \mu_1 \neq \mu_2$
    - The alternative hypothesis will depend on what is being tested
- **STEP 2**: Decide if it is a ***z*-test or a *t*-test**
  - If the populations variances are **known** then use a ***z*-test**
  - If the populations variances are **unknown** then use a ***t*-test**
    - Assume the variances are equal and use a pooled sample
- **STEP 3**: **Enter the data** into your GDC and choose **two-sample *z*-test or two-sample *t*-test**
  - If you have the raw data
    - Enter the data as a list
    - Enter the values of σ_1 & σ_2 if a *z*-test
    - Choose the pooled option if a *t*-test
  - If you have summary statistics (only for a *z*-test)
    - Enter the values of $\bar{x}_1$, $\bar{x}_2$, σ_1, σ_2, n_1 & n_2
  - Your GDC will give you the *p*-value
- **STEP 4**: **Decide** whether there is **evidence to reject the null hypothesis**
  - If the *p*-value < significance level then reject H_0
- **STEP 5**: **Write your conclusion**
  - If you **reject H_0** then there is evidence to suggest that...
    - The mean of the 1st population is smaller (for $H_1 : \mu_1 < \mu_2$)
    - The mean of the 1st population is bigger (for $H_1 : \mu_1 > \mu_2$)
    - The means of the two populations are different (for $H_1 : \mu_1 \neq \mu_2$)
  - If you **accept H_0** then there is **insufficient evidence** to reject the null hypothesis which suggests that...

- The mean of the 1st population is not smaller (for $H_1 : \mu_1 < \mu_2$)
- The mean of the 1st population is not bigger (for $H_1 : \mu_1 > \mu_2$)
- The means of the two populations are not different (for $H : \mu \neq \mu$)

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

$$\boxed{\begin{array}{l} H_0: \mu_c = \mu_A \\ H_1: \mu_c < \mu_A \end{array}}$$

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
$p = 0.007259...$

$$\boxed{p = 0.00726 \text{ (3sf)}}$$

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

$0.00726 < 0.01$

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.

Paired t-tests

What is a paired *t*-test?

- A paired test is where you take **two samples** but **each data point from one sample can be paired with a data point from the other sample**
  - These are used when one group of members are used twice and the two results for each member are paired
    - It could be to compare the sample before and after introducing a new factor
    - It could be to compare the sample under two different conditions
- For this test you use the differences between the pairs and treat them as one sample
  - As the variance of the differences is unlikely to be known you will use a *t*-test
  - For a paired test you need to assume the differences are normally distributed
    - You don't need to assume the populations are normally distributed

What are the steps for performing a paired *t*-test on my GDC?

- **STEP 1**: **Write the hypotheses**
  - $H_0: \mu_D = 0$
    - Clearly state that μ_D represents the **population mean of the differences**
    - In words this means the population mean has not changed
  - For a **one-tailed** test $H_1: \mu_D < 0$ or $H_1: \mu_D > 0$
  - For a **two-tailed** test: $H_1: \mu_D \neq 0$
    - The alternative hypothesis will depend on what is being tested
- **STEP 2**: **Enter the data** into your GDC and choose the **one-sample *t*-test**
  - Enter the differences as a list
    - Be consistent with the order in which you subtract paired values
  - Your GDC will give you the *p*-value
- **STEP 3**: **Decide** whether there is **evidence to reject the null hypothesis**
  - If the *p*-value < significance level then reject H_0
- **STEP 4**: **Write your conclusion**
  - If you **reject H_0** then there is evidence to suggest that...
    - The mean has decreased (for $H_1: \mu_D < 0$)
    - The mean has increased (for $H_1: \mu_D > 0$)
    - The mean has changed (for $H_1: \mu_D \neq 0$)
  - If you **accept H_0** then there is **insufficient evidence** to reject the null which suggests that...
    - The mean has not decreased (for $H_1: \mu_D < 0$)
    - The mean has not increased (for $H_1: \mu_D > 0$)
    - The mean has not changed (for $H_1: \mu_D \neq 0$)

Worked example

In a school all students must study French and Spanish. 9 students are selected and complete a test in both subjects, the standardised scores are shown below

Student	1	2	3	4	5	6	7	8	9
French score	61	82	77	80	99	69	75	71	81
Spanish score	74	79	83	66	95	79	82	81	85

The headteacher wants to investigate whether there is a difference in the students' scores between the two subjects. A paired *t*-test is to be performed at a 10% significance level.

a) Write down the null and alternative hypotheses.

Let D be the French score minus the Spanish score for the students. Let μ_D be the mean difference for the whole population of students.

$\boxed{\begin{array}{l} H_0: \mu_D = 0 \\ H_1: \mu \neq 0 \end{array}}$ Testing for a difference in scores

b) Find the *p*-value for this test.

Calculate the difference for each student d = French − Spanish

Student	1	2	3	4	5	6	7	8	9
d	-13	3	-6	14	4	-10	-7	-10	-4

Enter the differences into the GDC and use a t-test

$p = 0.2958...$

$\boxed{p = 0.296 \text{ (3sf)}}$

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

$0.2958... > 0.1$

$\boxed{\text{Accept } H_0 \text{ as p-value > significance level. There is insufficient evidence to suggest that there is a difference in scores.}}$

4.12.3 Binomial Hypothesis Testing

Binomial Hypothesis Testing

What is a hypothesis test using a binomial distribution?

- You can use **a binomial distribution** to test whether the **proportion** of a population with a specified characteristic has **increased** or **decreased**
  - These tests will always be **one-tailed**
  - You will not be expected to perform a two-tailed hypothesis test with the binomial distribution
- A sample will be taken and the **test statistic** x will be the **number of members with the characteristic** which will be tested using the distribution $X \sim \mathrm{B}(n, p)$
  - This can be thought of as the number of successes

What are the steps for a hypothesis test of a binomial proportion?

- **STEP 1**: **Write the hypotheses**
  - $H_0 : p = p_0$
    - Clearly state that p represents the **population proportion**
    - p_0 is the **assumed population proportion**
  - $H_1 : p < p_0$ or $H_1 : p > p_0$
- **STEP 2**: Calculate the ***p*-value** or find the **critical region**
  - See below
- **STEP 3**: **Decide** whether there is **evidence to reject the null hypothesis**
  - If the p-value < significance level then reject H_0
  - If the test statistic is in the critical region then reject H_0
- **STEP 4: Write your conclusion**
  - If you **reject H_0** then there is evidence to suggest that...
    - The population proportion has decreased (for $H_1 : p < p_0$)
    - The population proportion has increased (for $H_1 : p > p_0$)
  - If you **accept H_0** then there is **insufficient evidence** to reject the null hypothesis which suggests that...
    - The population proportion has not decreased (for $H_1 : p < p_0$)
    - The population proportion has not increased (for $H_1 : p > p_0$)

How do I calculate the *p*-value?

- The p-value is determined by the **test statistic** x
- The p-value is the probability that 'a value being **at least as extreme** as the test statistic' would occur if **null hypothesis were true**
  - For $H_1 : p < p_0$ the p-value is $\mathrm{P}(X \le x | p = p_0)$
  - For $H_1 : p > p_0$ the p-value is $\mathrm{P}(X \ge x | p = p_0)$

How do I find the critical value and critical region?

- The critical value and critical region are determined by the **significance level** $\alpha\%$
- Your calculator might have an **inverse binomial function** that works just like the inverse normal function
  - You need to use this value to find the critical value
  - The value given by the inverse binomial function is normally one away from the actual critical value
- For $H_1: p < p_0$ the critical region is $X \leq c$ where c is the critical value
  - c is the **largest integer** such that $P(X \leq c | p = p_0) \leq \alpha\%$
    - Check that $P(X \leq c + 1 | p = p_0) > \alpha\%$
- For $H_1: p > p_0$ the critical region is $X \geq c$ where c is the critical value
  - c is the **smallest integer** such that $P(X \geq c | p = p_0) \leq \alpha\%$
    - Check that $P(X \geq c - 1 | p = p_0) > \alpha\%$

Worked example

The existing treatment for a disease is known to be effective in 85% of cases. Dr Sabir develops a new treatment which she claims is more effective than the existing one. To test her claim she uses the new treatment on a random sample of 60 patients with the disease and finds that the treatment was effective for 57 of them.

a) State null and alternative hypotheses to test Dr Sabir's claim.

Let p be the proportion of the population for which the new treatment is effective.

$H_0: p = 0.85$

$H_1: p > 0.85$ Testing for an increase

b) Perform the test using a 1% significance level. Clearly state the conclusion in context.

Let $X \sim B(60, p)$ be the number of people in the sample for which the new treatment is effective

Find the p-value and compare to the significance level

$p = P(X \geq 57 \mid p = 0.85) = 0.01483... > 0.01$

Accept H_0 as p-value > significance level.
There is insufficient evidence to suggest that the new treatment is more effective than the existing treatment.

4.12.4 Poisson Hypothesis Testing

Poisson Hypothesis Testing

What is a hypothesis test using a Poisson distribution?

- You can use **a Poisson distribution** to test whether the **mean number of occurrences** for a **given time period** within a population has **increased** or **decreased**
  - These tests will always be **one-tailed**
  - You will not be expected to perform a two-tailed hypothesis test with the Poisson distribution
- A sample will be taken and the **test statistic** x will be the **number of occurrences** which will be tested using the distribution $X \sim \mathrm{Po}(m)$

What are the steps for a hypothesis test of a Poisson proportion?

- **STEP 1**: **Write the hypotheses**
  - $H_0 : m = m_0$
    - Clearly state that m represents the **mean number of occurrences** for the **given time period**
    - m_0 is the **assumed mean number of occurrences**
    - You might have to use **proportion** to find m_0
  - $H_1 : m < m_0$ or $H_1 : m > m_0$
- **STEP 2**: Calculate the **p-value** or find the **critical region**
  - See below
- **STEP 3**: **Decide** whether there is **evidence to reject the null hypothesis**
  - If the p-value < significance level then reject H_0
  - If the test statistic is in the critical region then reject H_0
- **STEP 4**: **Write your conclusion**
  - If you **reject H_0** then there is evidence to suggest that...
    - The mean number of occurrences has decreased (for $H_1 : m < m_0$)
    - The mean number of occurrences has increased (for $H_1 : m > m_0$)
  - If you **accept H_0** then there is **insufficient evidence** to reject the null hypothesis which suggests that...
    - The mean number of occurrences has not decreased (for $H_1 : m < m_0$)
    - The mean number of occurrences has not increased (for $H_1 : m > m_0$)

How do I calculate the p-value?

- The p-value is determined by the **test statistic** x
- The p-value is the probability that 'a value being **at least as extreme** as the test statistic' would occur if **null hypothesis were true**
  - For $H_1 : m < m_0$ the p-value is $\mathrm{P}(X \le x \mid m = m_0)$
  - For $H_1 : m > m_0$ the p-value is $\mathrm{P}(X \ge x \mid m = m_0)$

How do I find the critical value and critical region?

- The critical value and critical region are determined by the **significance level** $\alpha\%$
- Your calculator might have an **inverse Poisson function** that works just like the inverse normal function
  - You need to use this value to find the critical value
  - The value given by the inverse Poisson function is normally one away from the actual critical value
- For $H_1: m < m_0$ the critical region is $X \leq c$ where c is the critical value
  - c is the **largest integer** such that $P(X \leq c \mid m = m_0) \leq \alpha\%$
    - Check that $P(X \leq c+1 \mid m = m_0) > \alpha\%$
- For $H_1: m > m_0$ the critical region is $X \geq c$ where c is the critical value
  - c is the **smallest integer** such that $P(X \geq c \mid m = m_0) \leq \alpha\%$
    - Check that $P(X \geq c-1 \mid m = m_0) > \alpha\%$

Worked example

The owner of a website claims that his website receives an average of 120 hits per hour. An interested purchaser believes the website receives on average fewer hits than they claim. The owner chooses a 10–minute period and observes that the website receives 11 hits. It is assumed that the number of hits the website receives in any given time period follows a Poisson Distribution.

a) State null and alternative hypotheses to test the purchaser's claim.

Let m be the mean number of hits in a 10-minute period

120 hits in an hour $\Rightarrow$ 20 hits in a 10-minute period

$\boxed{\begin{array}{l} H_0: m = 20 \\ H_1: m < 20 \end{array}}$ Testing for fewer hits

b) Find the critical region for a hypothesis test at the 5% significance level.

Let $X \sim Po(m)$ be the number of hits in a 10-minute period

The critical value c is the largest value such that

$P(X \leq c \mid m = 20) \leq 0.05$

You can use the inverse Poisson function on the GDC to decide which value to check first

$P(X \leq 13 \mid m = 20) = 0.0661... > 0.05$ Too big so reduce the region

$P(X \leq 12 \mid m = 20) = 0.0390... < 0.05$

$\boxed{\text{Critical region } X \leq 12}$

c) Perform the test using a 5% significance level. Clearly state the conclusion in context.

$11 < 12$ so 11 is in the critical region

$\boxed{\text{Reject } H_0 \text{ as test statistic is in critical region. There is sufficient evidence to suggest that the website receives on average fewer hits than they claim.}}$

4.12.5 Hypothesis Testing for Correlation

Hypothesis Testing for Correlation

What is a hypothesis test for correlation?

- You can use a ***t*-test** to test whether there is **linear correlation** between two normally distributed variables
  - If specifically testing for positive (or negative) linear correlation then a **one-tailed test** is used
  - If testing for any linear correlation then a **two-tailed test** is used
- A sample will be taken and the **raw data** will be given
  - You might be asked to calculate the **PMCC (Pearson's product-moment correlation coefficient)**

What are the steps for a hypothesis test for correlation?

- **STEP 1**: **Write the hypotheses**
  - $H_0 : \rho = 0$
    - Clearly state that ρ represents **population correlation coefficient** between the two variables
    - In words this means there is no correlation
  - $H_1 : \rho < 0$, $H_1 : \rho > 0$ or $H_1 : \rho \neq 0$
- **STEP 2**: Calculate the ***p*-value** or the **PMCC**
  - Choose a t-test for linear regression
  - Enter the data as two lists into GDC
- **STEP 3**: **Decide** whether there is evidence to **reject the null hypothesis**
  - If the p-value < significance level then reject H_0
  - If the absolute value of the PMCC is bigger than the absolute value of the critical value then reject H_0
    - If you are expected to use the PMCC you will be **given the critical value** in the exam
- **STEP 4**: **Write your conclusion**
  - If you **reject H_0** then there is evidence to suggest that...
    - There is a negative linear correlation between the two variables (for $H_1 : \rho < 0$)
    - There is a positive linear correlation between the two variables (for $H_1 : \rho > 0$)
    - There is a linear correlation between the two variables (for $H_1 : \rho \neq 0$)
  - If you **accept H_0** then there is **insufficient evidence** to reject the null hypothesis which suggests that...
    - There is not a negative linear correlation between the two variables (for $H_1 : \rho < 0$)
    - There is not a positive linear correlation between the two variables (for $H_1 : \rho > 0$)
    - There is not a linear correlation between the two variables (for $H_1 : \rho \neq 0$)

Worked example

Jessica wants to test whether there is any linear correlation between the distance she runs in a day, d km, and the amount of sleep she has the night after her run, t hours. Over the period of a month she takes a random sample of 9 days, the results are recorded in the table.

Distance (d km)	1.2	2.3	1.5	1.3	2.5	1.8	1.9	2.0	1.1
Sleep (t hours)	7.9	8.1	7.6	7.3	8.1	8.4	7.8	7.9	6.8

a) Write down null and alternative hypotheses that Jessica can use for her test.

Let ρ be the correlation coefficient between Jessica's distances and the hours of sleep she gets.

$H_0 : \rho = 0$

$H_1 : \rho \neq 0$

Testing for any linear correlation

b) Perform the hypothesis test for linear correlation using a 5% significance level. Clearly state your conclusion.

Type the data in GDC and select a t-test for linear regression

$p = 0.03833... < 0.05$

Reject H_0 as p-value < significance level.
There is sufficient evidence to suggest that there is a linear correlation between the distance that Jessica runs and the hours she sleeps.

4.12.6 Type I & Type II Errors

Type I & Type II Errors

What are Type I & Type II errors?

- There are **four possible outcomes** of a hypothesis test:
  - H_0 is **false** and H_0 is **rejected**
  - H_0 is **true** and H_0 is **not rejected**
    - The test is **accurate** for these two outcomes
  - H_0 is **true** and H_0 is **rejected**
  - H_0 is **false** and H_0 is **not rejected**
    - The test has led to an **error** for these two outcomes
- A **Type I error** occurs when a hypothesis test gives **sufficient evidence to reject H_0** despite it **being true**
  - This is sometimes called a "**false positive**"
  - In a court case this would be when the defendant is found **guilty despite being innocent**
- A **Type II error** is when a hypothesis test gives **insufficient evidence to reject H_0** despite it **being false**
  - This is sometimes called a "**false negative**"
  - In a court case this would be when the defendant is found **innocent despite being guilty**

How do I find the probabilities of a Type I or Type II error occurring?

- You should calculate the probability of errors occurring **before a sample is taken**
- The probabilities are **determined by the critical region**
  - Equally it is **determined by the significance level** $\alpha\%$
  - Critical regions are determined such that**:**
    - They keep the **probability of a Type I error less than or equal to** the **significance level**
    - They **maximise** the **probability of a Type I error**

- The probability of a **Type I error** occurring is equal to the probability of **being in the critical region** if H_0 were true
  - P(Type I error) = P(being in the critical region | H_0 is true)
  - For a continuous distribution (normal, t, χ^2)
    - P(Type I error) = α%
  - For a discrete distribution (binomial, Poisson)
    - P(Type I error) ≤ α%
- The probability of a **Type II** error occurring is equal to the probability of **not being in the critical region** given the actual value of the population parameter
  - P(Type II error) = P(not being in critical region | actual population parameter)
  - You need to know the actual population parameter in order to find the probability of a Type II error
- Once a sample has been taken you can determine which error could have occurred
  - If you **rejected H_0** then you could have made a **Type I error**
  - If you **accepted H_0** then you could have made a **Type II error**

Can I reduce the probabilities of making a Type I or Type II error?

- You can **reduce** the probability of a **Type I** error by **reducing the significance level**
  - However this will **increase** the probability of a **Type II error**
- You can **reduce** the probability of a **Type II** error by **increasing the significance level**
  - However this will **increase** the probability of a **Type I error**
- The only way to **reduce both** probabilities is by **increasing the size of the sample**

Worked example

Lucy can hit the target 70% of the time when she throws an axe with her right hand. She claims that the proportion, p, of her throws that hit the target is higher than 70% when she uses her left hand. Lucy uses the hypotheses $\mathrm{H}_0 : p = 0.7$ and $\mathrm{H}_0 : p > 0.7$ to test her claim. Lucy makes 100 throws and will reject the null hypothesis if the axe hits the target more than 77 times.

a) find the probability of a Type I error.

Let $X \sim B(100, p)$ be the number of times Lucy hits the target when using her left hand.

P(Type I error) = P(being in critical region | H_0 is true)

$$\begin{aligned} P(\text{Type I error}) &= P(X > 77 \mid p = 0.7) \\ &= P(78 \leq X \leq 100 \mid p = 0.7) \\ &= 0.04786\ldots \end{aligned}$$

$$\boxed{P(\text{Type I error}) = 0.0479 \ (3sf)}$$

b) Given that Lucy actually hits the target 80% of the time with her left hand, find the probability of a Type II error.

P(Type II error) = P(not being in critical region | true population parameter)

$$\begin{aligned} P(\text{Type II error}) &= P(X \leq 77 \mid p = 0.8) \\ &= P(0 \leq X \leq 77 \mid p = 0.8) \\ &= 0.2610\ldots \end{aligned}$$

$$\boxed{P(\text{Type II error}) = 0.261 \ (3sf)}$$