Question 1

Vasily is a professional chess player. Over many years of competing, the mean number of minutes he has spent per move is 3.71. After beginning to work with a new training partner, Vasily believes that the mean number of minutes he spends per move has decreased. In his next tournament Vasily makes a total of 510 moves in his chess games, and the mean number of minutes per move is found to be 3.62.

Let the random variable X represent the amount of time Vasily spends per move after beginning to work with his new training partner. It is known from past experience that the standard deviation for the amount of time Vasily spends per chess move is 1.02 minutes.

(a) State the null and alternative hypotheses for a hypothesis test to test Vasily's belief.

[1]

(b) Write down the distribution of the sample mean $\bar{X}$ that may be used to test Vasily's belief. Be sure to justify your answer.

[2]

(c) Use the p-value to test whether there is sufficient evidence at the 10% significance level to support Vasily's belief that the mean number of minutes he spends per move has decreased.

[3]

(d) Find the critical region for the test.

[2]

(a) Let μ be the mean number of minutes per move

$\boxed{\begin{array}{l} H_0 : \mu = 3.71 \\ H_1 : \mu < 3.71 \end{array}}$ < as testing for decrease

Vasily is a professional chess player. Over many years of competing, the mean number of minutes he has spent per move is 3.71. After beginning to work with a new training partner, Vasily believes that the mean number of minutes he spends per move has decreased. In his next tournament Vasily makes a total of 510 moves in his chess games, and the mean number of minutes per move is found to be 3.62. $n = 510$

Let the random variable X represent the amount of time Vasily spends per move after beginning to work with his new training partner. It is known from past experience that the standard deviation for the amount of time Vasily spends per chess move is 1.02 minutes.

$\sigma = 1.02$

(a) State the null and alternative hypotheses for a hypothesis test to test Vasily's belief.

[1]

(b) Write down the distribution of the sample mean $\bar{X}$ that may be used to test Vasily's belief. Be sure to justify your answer.

[2]

(c) Use the p-value to test whether there is sufficient evidence at the 10% significance level to support Vasily's belief that the mean number of minutes he spends per move has decreased.

[3]

(d) Find the critical region for the test.

[2]

(b) Central limit theorem states:

Let X be a random variable with mean μ and variance σ^2. If n independent observations of X are taken then $\bar{X}$ is approximately $N(\mu, \frac{\sigma^2}{n})$ provided n is sufficiently large

$\boxed{\text{As 510 is large the Central limit theorem states } \bar{X} \text{ can be approximated as } \bar{X} \sim N(\mu, \frac{1.02^2}{510})}$

Vasily is a professional chess player. Over many years of competing, the mean number of minutes he has spent per move is 3.71. After beginning to work with a new training partner, Vasily believes that the mean number of minutes he spends per move has decreased. In his next tournament Vasily makes a total of 510 moves in his chess games, and the mean number of minutes per move is found to be 3.62.

Let the random variable X represent the amount of time Vasily spends per move after beginning to work with his new training partner. It is known from past experience that the standard deviation for the amount of time Vasily spends per chess move is 1.02 minutes.

(a) State the null and alternative hypotheses for a hypothesis test to test Vasily's belief.

[1]

(b) Write down the distribution of the sample mean $\bar{X}$ that may be used to test Vasily's belief. Be sure to justify your answer.

[2]

(c) Use the p-value to test whether there is sufficient evidence at the 10% significance level to support Vasily's belief that the mean number of minutes he spends per move has decreased.

[3]

(d) Find the critical region for the test.

[2]

(c) As H_1 is looking for a decrease : $p = P(\bar{X} < \bar{x})$

$p = P(\bar{X} < 3.62 \mid \mu = 3.71)$ Note $\sigma = \frac{1.02}{\sqrt{510}}$

$= 0.02315...$

Reject H_0 if p-value < significance level

$p = 0.02315... < 0.1$ so there is sufficient evidence at the 10% significance level to reject H_0. This suggests that the mean number of minutes Vasily spends per move has decreased.

The p-value can be found directly from GDC using one-sample z-test (as population σ is known)

$\bar{x} = 3.62$ $\sigma = 1.02$ $n = 510$

$p = 0.02315...$

Vasily is a professional chess player. Over many years of competing, the mean number of minutes he has spent per move is 3.71. After beginning to work with a new training partner, Vasily believes that the mean number of minutes he spends per move has decreased. In his next tournament Vasily makes a total of 510 moves in his chess games, and the mean number of minutes per move is found to be 3.62.

Let the random variable X represent the amount of time Vasily spends per move after beginning to work with his new training partner. It is known from past experience that the standard deviation for the amount of time Vasily spends per chess move is 1.02 minutes.

(a) State the null and alternative hypotheses for a hypothesis test to test Vasily's belief.

[1]

(b) Write down the distribution of the sample mean $\bar{X}$ that may be used to test Vasily's belief. Be sure to justify your answer.

[2]

(c) Use the p-value to test whether there is sufficient evidence at the 10% significance level to support Vasily's belief that the mean number of minutes he spends per move has decreased.

[3]

(d) Find the critical region for the test.

[2]

(d) Critical region is $\bar{X} < c$ where $P(\bar{X} < c) = 0.1$

$P(\bar{X} < c \mid \mu = 37.1) = 0.1$

Using inverse normal distribution on GDC :

$c = 3.65211...$

Critical region $\bar{X} < 3.65$ (3sf)

Question 2

Gavio is the owner of a company that farms black soldier fly maggots to be made into insect protein powder for use in recipes in trendy restaurants. The weight of the maggots is known to be normally distributed, with a mean of 0.104 grams and a standard deviation of 0.039 grams.

Gavio has begun to use a new type of feed for the maggots on his farm. Because he is curious whether the new feed has resulted in a change in the average weight of his maggots, he selects 100 maggots raised on the new feed. It may be assumed that the use of the new feed has not changed the standard deviation of the maggots' weights.

(a) Explain why a two-tailed hypothesis test is appropriate in this situation.

[1]

The test is to be conducted at the 5% significance level.

(b) (i) Explain what a Type I error is.

(ii) Write down the probability that Gavio's test will result in a Type I error.

[2]

(c) The 100 maggots raised on the new feed have a total weight of 11.158 grams. Conduct a hypothesis test to determine whether there is sufficient evidence at the 5% level to suggest that the mean weight of the maggots has changed.

[5]

(d) Suggest a change that Gavio might make to increase his certainty in the results of the test.

[1]

(a) The test is for a change (either an increase or a decrease) so a two-tailed test should be used.

Gavio is the owner of a company that farms black soldier fly maggots to be made into insect protein powder for use in recipes in trendy restaurants. The weight of the maggots is known to be normally distributed, with a mean of 0.104 grams and a standard deviation of 0.039 grams.

Gavio has begun to use a new type of feed for the maggots on his farm. Because he is curious whether the new feed has resulted in a change in the average weight of his maggots, he selects 100 maggots raised on the new feed. It may be assumed that the use of the new feed has not changed the standard deviation of the maggots' weights.

(a) Explain why a two-tailed hypothesis test is appropriate in this situation.

[1]

The test is to be conducted at the 5% significance level.

(b) (i) Explain what a Type I error is.

(ii) Write down the probability that Gavio's test will result in a Type I error.

[2]

(c) The 100 maggots raised on the new feed have a total weight of 11.158 grams. Conduct a hypothesis test to determine whether there is sufficient evidence at the 5% level to suggest that the mean weight of the maggots has changed.

[5]

(d) Suggest a change that Gavio might make to increase his certainty in the results of the test.

[1]

(b) (i) A Type I error would involve rejecting the null hypothesis when it is true.
For this test that would be saying the mean weight has changed when it has not.

(ii) For a continuous distribution the probability of a Type I error is equal to the significance level.

$P(\text{Type I error}) = 0.05$

EXAM PAPERS PRACTICE

Gavio is the owner of a company that farms black soldier fly maggots to be made into insect protein powder for use in recipes in trendy restaurants. The weight of the maggots is known to be normally distributed, with a mean of 0.104 grams and a standard deviation of 0.039 grams.

Gavio has begun to use a new type of feed for the maggots on his farm. Because he is curious whether the new feed has resulted in a change in the average weight of his maggots, he selects 100 maggots raised on the new feed. It may be assumed that the use of the new feed has not changed the standard deviation of the maggots' weights.

(a) Explain why a two-tailed hypothesis test is appropriate in this situation.

[1]

The test is to be conducted at the 5% significance level.

(b) (i) Explain what a Type I error is.

(ii) Write down the probability that Gavio's test will result in a Type I error.

[2]

(c) The 100 maggots raised on the new feed have a total weight of 11.158 grams. Conduct a hypothesis test to determine whether there is sufficient evidence at the 5% level to suggest that the mean weight of the maggots has changed.

[5]

(d) Suggest a change that Gavio might make to increase his certainty in the results of the test.

[1]

(c) Let μ be the mean weight of maggots

$H_0: \mu = 0.104$
$H_1: \mu \neq 0.104$

Standard deviation is known so use one-sample z-test

$\bar{x} = \frac{11.158}{100} = 0.11158 \quad \sigma = 0.039 \quad n = 100$

$p = 0.05194...$

Reject H_0 if p-value < significance level

$p = 0.05194... > 0.05$ so insufficient evidence to reject H_0. Accept H_0. This suggests the mean weight has not changed.

Critical region can be found using $\bar{X} \sim N(\mu, \frac{0.039^2}{100})$
$\bar{X} < 0.09635...$ or $\bar{X} > 0.11164...$

You can also calculate the p-value directly
$p = 2 \times P(\bar{X} > 0.11158 \mid \mu = 0.104) = 2 \times 0.02597... = 0.05194..$

Gavio is the owner of a company that farms black soldier fly maggots to be made into insect protein powder for use in recipes in trendy restaurants. The weight of the maggots is known to be normally distributed, with a mean of 0.104 grams and a standard deviation of 0.039 grams.

Gavio has begun to use a new type of feed for the maggots on his farm. Because he is curious whether the new feed has resulted in a change in the average weight of his maggots, he selects 100 maggots raised on the new feed. It may be assumed that the use of the new feed has not changed the standard deviation of the maggots' weights.

(a) Explain why a two-tailed hypothesis test is appropriate in this situation.

[1]

The test is to be conducted at the 5% significance level.

(b) (i) Explain what a Type I error is.

(ii) Write down the probability that Gavio's test will result in a Type I error.

[2]

(c) The 100 maggots raised on the new feed have a total weight of 11.158 grams. Conduct a hypothesis test to determine whether there is sufficient evidence at the 5% level to suggest that the mean weight of the maggots has changed.

[5]

(d) Suggest a change that Gavio might make to increase his certainty in the results of the test.

[1]

(d) Use a larger sample as this makes the standard deviation of $\bar{X}$ smaller. This will make the critical regions bigger.

Question 3

The lengths of rock songs follow a normal distribution. Miggy believes that the average length rock songs is longer than the average metal song which is 253 seconds. Miggy wants to test his belief using a hypothesis test with a 10% significance level, he uses the null hypothesis $H_0 : \mu = 253$. Miggy takes a random sample of 9 rock songs and records their lengths, in seconds, in the table below.

313	146	222	284	219	265	416	205	390

(a) (i) State why Miggy should use a t-test.

(ii) Write down the alternative hypothesis for Miggy's test.

[2]

(b) Find the p-value for the test.

[2]

(c) State, with a reason, whether the test supports Miggy's belief.

[2]

(a) (i) The population variance is unknown so a t-test should be used.

(ii) Longer than 253 $\Rightarrow \mu > 253$

$\boxed{H_1 : \mu > 253}$

The lengths of rock songs follow a normal distribution. Miggy believes that the average length rock songs is longer than the average metal song which is 253 seconds. Miggy wants to test his belief using a hypothesis test with a 10% significance level, he uses the null hypothesis $H_0 : \mu = 253$. Miggy takes a random sample of 9 rock songs and records their lengths, in seconds, in the table below.

313	146	222	284	219	265	416	205	390

(a) (i) State why Miggy should use a t-test.

(ii) Write down the alternative hypothesis for Miggy's test.

[2]

(b) Find the p-value for the test.

[2]

(c) State, with a reason, whether the test supports Miggy's belief.

[2]

(b) Type data into into GDC and select 1 sample t-test

$p = 0.25430...$

$\boxed{p = 0.254 \text{ (3sf)}}$

Your GDC will also give you the summary statistics

$\bar{x} = 273.33...$, $s_{n-1} = 88.17...$, $n = 9$

The lengths of rock songs follow a normal distribution. Miggy believes that the average length rock songs is longer than the average metal song which is 253 seconds. Miggy wants to test his belief using a hypothesis test with a 10% significance level, he uses the null hypothesis $H_0 : \mu = 253$. Miggy takes a random sample of 9 rock songs and records their lengths, in seconds, in the table below.

313	146	222	284	219	265	416	205	390

(a) (i) State why Miggy should use a t-test.

(ii) Write down the alternative hypothesis for Miggy's test.

[2]

(b) Find the p-value for the test.

[2]

(c) State, with a reason, whether the test supports Miggy's belief.

[2]

(c) $p = 0.25430...$

Miggy's belief is supported if H_0 is rejected

$0.25430... > 0.1$

Therefore insufficient evidence to reject H_0.

The test does not support Miggy's belief.

Question 4

Kaarina and Hannu are big fans of Lobfickle brand jelly beans. The weights of individual Lobfickle jelly beans are known to be normally distributed. Hannu insists that the mean weight of a Lobfickle jelly bean is 1.20 grams, claiming that he read that on the internet somewhere. Kaarina suspects that the mean weight is less than that.

To test her suspicion, Kaarina takes a sample of 10 Lobfickle jelly beans and records their weights in grams. The table below shows her results:

1.09	1.15	1.22	1.15	1.30	1.11	1.13	1.14	1.11	1.20

Kaarina conducts a hypothesis test with this sample, using a significance level of 5%.

(a) Conduct Kaarina's proposed test, and determine the results using the stated significance level.

[5]

(a) Let μ be the mean weight of a Lobfickle jelly bean

$H_0: \mu = 1.20$
$H_1: \mu < 1.20$

Population variance is unknown so use a t-test.
Type the data into GDC one-sample t-test.
$p = 0.038619...$
Reject H_0 if p-value < significance level

$p = 0.0386... < 0.05$ so sufficient evidence to reject H_0. This suggests that the mean weight of Lobfickle jelly beans is less than 1.20 grams.

Note your GDC will also give you: $\bar{x} = 1.16$
$s_{n-1} = 0.06342..$
$n = 10$

Hannu says that Lobfickle jelly beans are too important to risk making a mistake here. Therefore he claims that a 1% level of significance should have been used for the test instead.

(b) Write down the conclusion of the hypothesis test if Hannu's proposed level of significance had been used instead.

[2]

Kaarina and Hannu are big fans of Lobfickle brand jelly beans. The weights of individual Lobfickle jelly beans are known to be normally distributed. Hannu insists that the mean weight of a Lobfickle jelly bean is 1.20 grams, claiming that he read that on the internet somewhere. Kaarina suspects that the mean weight is less than that.

To test her suspicion, Kaarina takes a sample of 10 Lobfickle jelly beans and records their weights in grams. The table below shows her results:

1.09	1.15	1.22	1.15	1.30	1.11	1.13	1.14	1.11	1.20

Kaarina conducts a hypothesis test with this sample, using a significance level of 5%.

(a) Conduct Kaarina's proposed test, and determine the results using the stated significance level.

[5]

Hannu says that Lobfickle jelly beans are too important to risk making a mistake here. Therefore he claims that a 1% level of significance should have been used for the test instead.

(b) Write down the conclusion of the hypothesis test if Hannu's proposed level of significance had been used instead.

[2]

(b) $p = 0.0386... > 0.01$ so insufficient evidence to reject H_0. This suggests that the mean weight of Lobfickle jelly beans is not less than 1.20 grams.

Question 5

Balik is a fisheries biologist studying fish living in two different river systems, A and B. A new industrial plant has been discharging waste into river system B for the past several years, and Balik is concerned that this is having an effect on the weight of a species of trout that lives in that river.

To test his theory, Balik collected adult samples of that trout species from each of the river systems. He recorded the weights in kg of the fish in each of his samples before returning them safely to their respective rivers. The following table summarises his results:

A	0.875	0.347	0.741	0.612	0.598	0.679	0.912	0.481	0.522	0.492
B	0.413	0.765	0.294	0.341	0.472	0.683	0.385	0.466	0.341	0.479

(a) Find the means of the weights of the fish sampled from each of the river systems.

[2]

A two-sample t-test at the 5% significance level is to be employed to analyse this data.

(b) Write down the null and alternative hypotheses for the test.

[1]

(c) Calculate the p-value for the test.

[3]

(d) Write down the conclusion to the test.

[2]

(a) Enter the data into the GDC

$\bar{x}_A = 0.6259$

$\bar{x}_B = 0.4639$

$\boxed{\text{Mean of sample from A: } \bar{x}_A = 0.626 \text{ kg (3sf)}}$

$\boxed{\text{Mean of sample from B: } \bar{x} = 0.464 \text{ kg (3sf)}}$

Balik is a fisheries biologist studying fish living in two different river systems, A and B. A new industrial plant has been discharging waste into river system B for the past several years, and Balik is concerned that this is having an effect on the weight of a species of trout that lives in that river.

To test his theory, Balik collected adult samples of that trout species from each of the river systems. He recorded the weights in kg of the fish in each of his samples before returning them safely to their respective rivers. The following table summarises his results:

A	0.875	0.347	0.741	0.612	0.598	0.679	0.912	0.481	0.522	0.492
B	0.413	0.765	0.294	0.341	0.472	0.683	0.385	0.466	0.341	0.479

(a) Find the means of the weights of the fish sampled from each of the river systems.

[2]

A two-sample t-test at the 5% significance level is to be employed to analyse this data.

(b) Write down the null and alternative hypotheses for the test.

[1]

(c) Calculate the p-value for the test.

[3]

(d) Write down the conclusion to the test.

[2]

(b) Let μ_A be the mean weight of the population of fish in river A

Let μ_B be the mean weight of the population of fish in river B

$H_0: \mu_A = \mu_B$

$H_1: \mu_A \neq \mu_B$ Testing for an "effect" (increase or decrease)

Balik is a fisheries biologist studying fish living in two different river systems, A and B. A new industrial plant has been discharging waste into river system B for the past several years, and Balik is concerned that this is having an effect on the weight of a species of trout that lives in that river.

To test his theory, Balik collected adult samples of that trout species from each of the river systems. He recorded the weights in kg of the fish in each of his samples before returning them safely to their respective rivers. The following table summarises his results:

A	0.875	0.347	0.741	0.612	0.598	0.679	0.912	0.481	0.522	0.492
B	0.413	0.765	0.294	0.341	0.472	0.683	0.385	0.466	0.341	0.479

(a) Find the means of the weights of the fish sampled from each of the river systems.

[2]

A two-sample t-test at the 5% significance level is to be employed to analyse this data.

(b) Write down the null and alternative hypotheses for the test.

[1]

(c) Calculate the p-value for the test.

[3]

(d) Write down the conclusion to the test.

[2]

(c) Perform two-sample pooled t-test

$p = 0.042460...$

$\boxed{p = 0.0425 \text{ (3sf)}}$

Balik is a fisheries biologist studying fish living in two different river systems, A and B. A new industrial plant has been discharging waste into river system B for the past several years, and Balik is concerned that this is having an effect on the weight of a species of trout that lives in that river.

To test his theory, Balik collected adult samples of that trout species from each of the river systems. He recorded the weights in kg of the fish in each of his samples before returning them safely to their respective rivers. The following table summarises his results:

A	0.875	0.347	0.741	0.612	0.598	0.679	0.912	0.481	0.522	0.492
B	0.413	0.765	0.294	0.341	0.472	0.683	0.385	0.466	0.341	0.479

(a) Find the means of the weights of the fish sampled from each of the river systems.

[2]

A two-sample t-test at the 5% significance level is to be employed to analyse this data.

(b) Write down the null and alternative hypotheses for the test.

[1]

(c) Calculate the p-value for the test.

[3]

(d) Write down the conclusion to the test.

[2]

(d) Reject H_0 if p-value < significance level

$p = 0.0424... < 0.05$ so sufficient evidence to reject H_0. This suggests that the mean weights of the fish in the two populations are not equal.

Note we can not conclude from this test which population has the larger mean. A one-tailed test would be needed to investigate that

Question 6

An alternative wellness and lifestyle coach has developed a new tutoring program which he insists will improve students' mathematics exam results. Students taking part in the program are encouraged to count the numbers of cute kittens in each of a series of images that are flashed before them while they listen to recordings by the lifestyle coach describing how awesome and clever he is. The coach claims that the practice with counting will improve the students' mathematics results because mathematics is all about numbers and counting things.

Seven students who participated in the program were tested at the start of the program, and again once the program was completed. Their results in the tests were as follows:

Student	1	2	3	4	5	6	7
Score before program	67	52	73	49	88	64	61
Score after program	69	55	72	52	80	65	67

(a) Complete the following table showing the change in the students' scores from before and after completing the program:

Student	1	2	3	4	5	6	7
Change in scores	2	3	−1	3	−8	1	6

[2]

(b) Calculate the p-value for a t-test at a 10% significance level on the table of differences from part (a), being sure to state your null and alternative hypotheses.

[3]

(c) Write down the conclusion to the test.

[2]

An alternative wellness and lifestyle coach has developed a new tutoring program which he insists will improve students' mathematics exam results. Students taking part in the program are encouraged to count the numbers of cute kittens in each of a series of images that are flashed before them while they listen to recordings by the lifestyle coach describing how awesome and clever he is. The coach claims that the practice with counting will improve the students' mathematics results because mathematics is all about numbers and counting things.

Seven students who participated in the program were tested at the start of the program, and again once the program was completed. Their results in the tests were as follows:

Student	1	2	3	4	5	6	7
Score before program	67	52	73	49	88	64	61
Score after program	69	55	72	52	80	65	67

(a) Complete the following table showing the change in the students' scores from before and after completing the program:

Student	1	2	3	4	5	6	7
Change in scores	2	3	−1	3	−8	1	6

[2]

(b) Calculate the p-value for a t-test at a 10% significance level on the table of differences from part (a), being sure to state your null and alternative hypotheses.

[3]

(c) Write down the conclusion to the test.

[2]

(b) Let μ_D be the mean change in scores for the population

$$\boxed{\begin{aligned} H_0 &: \mu_D = 0 \\ H_1 &: \mu_D > 0 \end{aligned}}$$

Testing for an increase

Enter changes into GDC and use one-sample t-test

$p = 0.31428...$

$\boxed{p = 0.314 \ (3sf)}$

An alternative wellness and lifestyle coach has developed a new tutoring program which he insists will improve students' mathematics exam results. Students taking part in the program are encouraged to count the numbers of cute kittens in each of a series of images that are flashed before them while they listen to recordings by the lifestyle coach describing how awesome and clever he is. The coach claims that the practice with counting will improve the students' mathematics results because mathematics is all about numbers and counting things.

Seven students who participated in the program were tested at the start of the program, and again once the program was completed. Their results in the tests were as follows:

Student	1	2	3	4	5	6	7
Score before program	67	52	73	49	88	64	61
Score after program	69	55	72	52	80	65	67

(a) Complete the following table showing the change in the students' scores from before and after completing the program:

Student	1	2	3	4	5	6	7
Change in scores	2	3	−1	3	−8	1	6

[2]

(b) Calculate the p-value for a t-test at a 10% significance level on the table of differences from part (a), being sure to state your null and alternative hypotheses.

[3]

(c) Write down the conclusion to the test.

[2]

(c) Reject H_0 if p-value < significance level

$\boxed{p = 0.314... > 0.1 \text{ so insufficient evidence to reject } H_0\text{. This suggests that the program does not lead to an improvement in scores.}}$

Question 7

The number of spam emails that Arturo receives per day is modelled by a Poisson distribution with a mean of 25 spam emails per day.

After changing the settings on his spam filter, Arturo decides to test whether the new settings have reduced the number of spam emails he receives. To do this he records the number of spam emails he receives over a period of one week. He decides to use a 5% level of significance for his test.

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

[1]

(b) (i) Find the critical value and the critical region for Arturo's test.

(ii) Hence find the probability that Arturo will make a Type I error in determining the conclusion of his test.

[5]

During the 1-week period, Arturo receives 149 spam emails.

(c) State Arturo's conclusion to his test, being sure to justify your answer.

[2]

(a) Let m be the mean number of spam emails in one week

$\boxed{\begin{aligned} H_0 &: m = 175 \\ H_1 &: m < 175 \end{aligned}}$ 25×7

Testing for a reduction

The number of spam emails that Arturo receives per day is modelled by a Poisson distribution with a mean of 25 spam emails per day.

After changing the settings on his spam filter, Arturo decides to test whether the new settings have reduced the number of spam emails he receives. To do this he records the number of spam emails he receives over a period of one week. He decides to use a 5% level of significance for his test.

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

[1]

(b) (i) Find the critical value and the critical region for Arturo's test.

(ii) Hence find the probability that Arturo will make a Type I error in determining the conclusion of his test.

[5]

During the 1-week period, Arturo receives 149 spam emails.

(c) State Arturo's conclusion to his test, being sure to justify your answer.

[2]

(b)(i) Let $X \sim Po(m)$ be the number of spam emails in a week

The critical value is the largest integer c such that $P(X \le c)$ is less than or equal to the significance level.

Your GDC might have an inverse Poisson function but always check the value

$P(X \le 154 \mid m = 175) = 0.058368... > 0.05$

$P(X \le 153 \mid m = 175) = 0.049730... < 0.05$

Critical value $X = 153$
Critical region $X \le 153$

(ii) Type I error is rejecting H_0 when it is true

P(Type I error) = P(X is in critical region)

$P(X \le 153 \mid m = 175) = 0.049730...$

P(Type I error) = 0.0497 (3sf)

The number of spam emails that Arturo receives per day is modelled by a Poisson distribution with a mean of 25 spam emails per day.

After changing the settings on his spam filter, Arturo decides to test whether the new settings have reduced the number of spam emails he receives. To do this he records the number of spam emails he receives over a period of one week. He decides to use a 5% level of significance for his test.

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

[1]

(b) (i) Find the critical value and the critical region for Arturo's test.

(ii) Hence find the probability that Arturo will make a Type I error in determining the conclusion of his test.

[5]

During the 1-week period, Arturo receives 149 spam emails.

(c) State Arturo's conclusion to his test, being sure to justify your answer.

[2]

(c) Reject H_0 if test statistic is in the critical region

$149 < 153$ so 149 is in critical region. There is sufficient evidence to reject H_0. This suggests that the new settings have reduced the number of spam emails.

Question 8

Based on historical records, the number of shooting stars that may be seen per hour by observers during an annual meteor shower can be modelled by a Poisson distribution with a mean of 60 shooting stars per hour.

Due to recent astronomical events, Zlata believes that this year's shower will be heavier than normal, with observers thus able to see a greater number of shooting stars per hour. To test her belief, she decides to record the number of shooting stars that she sees over a single 2-hour viewing period. If she observes more than 138 shooting stars during that period she will reject the historical mean.

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

[1]

(b) Find the probability that Zlata will make a Type I error in the conclusion of her test.

[2]

Zlata's colleague Lyaksandro believes that this year the actual mean number of shooting stars that an observer may expect to see per hour is 85.

(c) Explain what a Type II error is.

[1]

(d) If Lyaksandro is correct, find the probability that Zlata's test will result in a Type II error.

[3]

(a) Let m be the mean number of shooting stars in 2 hours

$$\boxed{\begin{aligned} H_0 &: m = 120 \\ H_1 &: m > 120 \end{aligned}}$$

2×60

Testing for a greater number

Based on historical records, the number of shooting stars that may be seen per hour by observers during an annual meteor shower can be modelled by a Poisson distribution with a mean of 60 shooting stars per hour.

Due to recent astronomical events, Zlata believes that this year's shower will be heavier than normal, with observers thus able to see a greater number of shooting stars per hour. To test her belief, she decides to record the number of shooting stars that she sees over a single 2-hour viewing period. If she observes more than 138 shooting stars during that period she will reject the historical mean.

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

[1]

(b) Find the probability that Zlata will make a Type I error in the conclusion of her test.

[2]

Zlata's colleague Lyaksandro believes that this year the actual mean number of shooting stars that an observer may expect to see per hour is 85.

(c) Explain what a Type II error is.

[1]

(d) If Lyaksandro is correct, find the probability that Zlata's test will result in a Type II error.

[3]

(b) Let X be the number of shooting stars in 2 hours

Type I error is rejecting H_0 when it is true

$P(\text{Type I error}) = P(X \text{ is in critical region})$

$P(X > 138 \mid m = 120) = P(X \geq 139 \mid m = 120)$

$= 0.048171...$

$$\boxed{P(\text{Type I error}) = 0.0482 \text{ (3sf)}}$$

EXAM PAPERS PRACTICE

Based on historical records, the number of shooting stars that may be seen per hour by observers during an annual meteor shower can be modelled by a Poisson distribution with a mean of 60 shooting stars per hour.

Due to recent astronomical events, Zlata believes that this year's shower will be heavier than normal, with observers thus able to see a greater number of shooting stars per hour. To test her belief, she decides to record the number of shooting stars that she sees over a single 2-hour viewing period. If she observes more than 138 shooting stars during that period she will reject the historical mean.

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

[1]

(b) Find the probability that Zlata will make a Type I error in the conclusion of her test.

[2]

Zlata's colleague Lyaksandro believes that this year the actual mean number of shooting stars that an observer may expect to see per hour is 85.

(c) Explain what a Type II error is.

[1]

(d) If Lyaksandro is correct, find the probability that Zlata's test will result in a Type II error.

[3]

(c) A Type II error is where H_0 is accepted when it is not true.
For this case it would be saying the number of shooting stars has not increased when it has.

Based on historical records, the number of shooting stars that may be seen per hour by observers during an annual meteor shower can be modelled by a Poisson distribution with a mean of 60 shooting stars per hour.

Due to recent astronomical events, Zlata believes that this year's shower will be heavier than normal, with observers thus able to see a greater number of shooting stars per hour. To test her belief, she decides to record the number of shooting stars that she sees over a single 2-hour viewing period. If she observes more than 138 shooting stars during that period she will reject the historical mean.

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

[1]

(b) Find the probability that Zlata will make a Type I error in the conclusion of her test.

[2]

Zlata's colleague Lyaksandro believes that this year the actual mean number of shooting stars that an observer may expect to see per hour is 85.

(c) Explain what a Type II error is.

[1]

(d) If Lyaksandro is correct, find the probability that Zlata's test will result in a Type II error.

[3]

(d) H_0 is accepted if X is not in critical region

H_0 is accepted if $X \leq 138$

$P(\text{Type II error}) = P(X \leq 138 \mid m = 170)$ $\quad 2 \times 85$

$= 0.0065002...$

$\boxed{P(\text{Type II error}) = 0.00650 \quad (3sf)}$

Question 9

In order to test the hypotheses $H_0 : p = 0.65$, $H_1 : p > 0.65$, where p is the probability of success for a binomial random variable X, 24 observations of X are made.

(a) Assuming the null hypothesis is true, determine the expected number of successes out of 24 observations.

[2]

The test is to be conducted at the 10% significance level.

(b) (i) Determine the critical region for the test.

(ii) Write down the probability that the test will result in a Type I error.

[4]

Out of the 24 observations, there are 19 successes.

(c) State the conclusions of the hypothesis test.

[2]

(d) Determine what the critical regions for the test would have been had the test instead been conducted at a significance level of

(i) 5%

(ii) 1%.

[3]

(a) If H_0 is true $X \sim B(24, 0.65)$

Formula booklet

$24 \times 0.65 = 15.6$

$\boxed{\text{Expected number of successes} = 15.6}$

In order to test the hypotheses $H_0 : p = 0.65$, $H_1 : p > 0.65$, where p is the probability of success for a binomial random variable X, 24 observations of X are made.

(a) Assuming the null hypothesis is true, determine the expected number of successes out of 24 observations.

[2]

The test is to be conducted at the 10% significance level.

(b) (i) Determine the critical region for the test.

(ii) Write down the probability that the test will result in a Type I error.

[4]

Out of the 24 observations, there are 19 successes.

(c) State the conclusions of the hypothesis test.

[2]

(d) Determine what the critical regions for the test would have been had the test instead been conducted at a significance level of

(i) 5%

(ii) 1%.

[3]

(b)(i) The critical value is the smallest integer c such that $P(X \geq c)$ is less than or equal to the significance level.

Your GDC might have an inverse binomial function but always check the value

$P(X \geq 19 \mid p = 0.65) = 0.104411... > 0.1$

$P(X \geq 20 \mid p = 0.65) = 0.042163... < 0.1$

$\boxed{\text{Critical region} \quad X \geq 20}$

ii) Type I error is rejecting H_0 when it is true

$P(\text{Type I error}) = P(X \text{ is in critical region})$

$P(X \geq 20 \mid p = 0.65) = 0.042163...$

$\boxed{P(\text{Type I error}) = 0.0422 \ (3sf)}$

In order to test the hypotheses $H_0 : p = 0.65,\ H_1 : p > 0.65,$ where p is the probability of success for a binomial random variable X, 24 observations of X are made.

(a) Assuming the null hypothesis is true, determine the expected number of successes out of 24 observations.

[2]

The test is to be conducted at the 10% significance level.

(b) (i) Determine the critical region for the test.

(ii) Write down the probability that the test will result in a Type I error.

[4]

Out of the 24 observations, there are 19 successes.

(c) State the conclusions of the hypothesis test.

[2]

(d) Determine what the critical regions for the test would have been had the test instead been conducted at a significance level of

(i) 5%

(ii) 1%.

[3]

(c) Reject H_0 if test statistic is in the critical region

$19 < 20$ so 19 is not in critical region.
There is insufficient evidence to reject H_0.
This suggests that the probability of success is not greater than 0.65.

In order to test the hypotheses $H_0 : p = 0.65,\ H_1 : p > 0.65,$ where p is the probability of success for a binomial random variable X, 24 observations of X are made.

(a) Assuming the null hypothesis is true, determine the expected number of successes out of 24 observations.

[2]

The test is to be conducted at the 10% significance level.

(b) (i) Determine the critical region for the test.

(ii) Write down the probability that the test will result in a Type I error.

[4]

Out of the 24 observations, there are 19 successes.

(c) State the conclusions of the hypothesis test.

[2]

(d) Determine what the critical regions for the test would have been had the test instead been conducted at a significance level of

(i) 5%

(ii) 1%.

[3]

(d)(i) $P(X \geq 19 \mid p = 0.65) = 0.104411... > 0.05$

$P(X \geq 20 \mid p = 0.65) = 0.042163... < 0.05$

At 5% level : critical region $X \geq 20$

(ii) $P(X \geq 21 \mid p = 0.65) = 0.01326... > 0.01$

$P(X \geq 22 \mid p = 0.65) = 0.00303... < 0.01$

At 1% level : critical region $X \geq 22$

EXAM PAPERS PRACTICE

Question 10

A national institute of health is attempting to assess the efficacy of a new treatment for a disease.

When given the current best treatment, 87% of patients recover fully from the effects of the disease. The institute wishes to know whether the percentage of patients who fully recover is greater when the new treatment is given instead.

(a) Explain why a significance level of 1% or lower would be appropriate for the institute's test.

[2]

The institute decides to give the new treatment to 1000 patients with the disease and to record the number, X, who fully recover. The test is to be conducted with a 1% significance level.

(b) (i) State the null and alternative hypotheses for the test.

(ii) Determine the critical region for the test.

[4]

Of the 1000 patients given the new treatment, 903 fully recover.

(c) State the conclusions of the hypothesis test.

[2]

(a) The significance level is the greatest value for the probability of incorrectly rejecting the null hypothesis.

In a health context we want there to be a very small probability of the test saying that a treatment has made an improvement when it has not.

A national institute of health is attempting to assess the efficacy of a new treatment for a disease.

When given the current best treatment, 87% of patients recover fully from the effects of the disease. The institute wishes to know whether the percentage of patients who fully recover is greater when the new treatment is given instead.

(a) Explain why a significance level of 1% or lower would be appropriate for the institute's test.

[2]

The institute decides to give the new treatment to 1000 patients with the disease and to record the number, X, who fully recover. The test is to be conducted with a 1% significance level.

(b) (i) State the null and alternative hypotheses for the test.

(ii) Determine the critical region for the test.

[4]

Of the 1000 patients given the new treatment, 903 fully recover.

(c) State the conclusions of the hypothesis test.

[2]

(b)(i) Let p be the proportion that fully recover

$H_0 : p = 0.87$

$H_1 : p > 0.87$

Test to see if percentage is greater

(ii) The critical value is the smallest integer c such that $P(X \geq c)$ is less than or equal to the significance level.

Your GDC might have an inverse binomial function but always check the value

$P(X \geq 894 \mid p = 0.87) = 0.011933... > 0.01$

$P(X \geq 895 \mid p = 0.87) = 0.009165... \quad 0.01$

Critical region $X \geq 895$

A national institute of health is attempting to assess the efficacy of a new treatment for a disease.

When given the current best treatment, 87% of patients recover fully from the effects of the disease. The institute wishes to know whether the percentage of patients who fully recover is greater when the new treatment is given instead.

(a) Explain why a significance level of 1% or lower would be appropriate for the institute's test.

[2]

The institute decides to give the new treatment to 1000 patients with the disease and to record the number, X, who fully recover. The test is to be conducted with a 1% significance level.

(b) (i) State the null and alternative hypotheses for the test.

(ii) Determine the critical region for the test.

[4]

Of the 1000 patients given the new treatment, 903 fully recover.

(c) State the conclusions of the hypothesis test.

[2]

(c) Reject H_0 if test statistic is in the critical region

903 > 895 so 903 is in critical region. There is sufficient evidence to reject H_0. This suggests that the new treatment has a higher percentage of patients who fully recover.

Question 11

Hrothgar is a professional mathematician who believes that there is a strong positive correlation between a person's happiness and the amount of time that person spends solving complex mathematics problems. To test his theory Hrothgar collects data from 10 people on how much time they spend solving complex mathematics problems, along with each person's score on a standardised 'level of happiness' test. The results are shown in the table below:

Person	1	2	3	4	5	6	7	8	9	10
Maths (hours/day)	2.3	5.8	1.0	12.2	0	3.4	9.6	0.5	4.9	15.7
Happiness	3.3	8.1	1.9	4.2	1.1	4.8	3.2	1.5	7.0	0.8

(a) Draw a scatter diagram for the data in the table.

[3]

(b) Test at the 5% level whether it is reasonable to assume a positive linear correlation between the two variables, being sure to state your null and alternative hypotheses. You may assume that happiness scores follow a normal distribution.

[3]

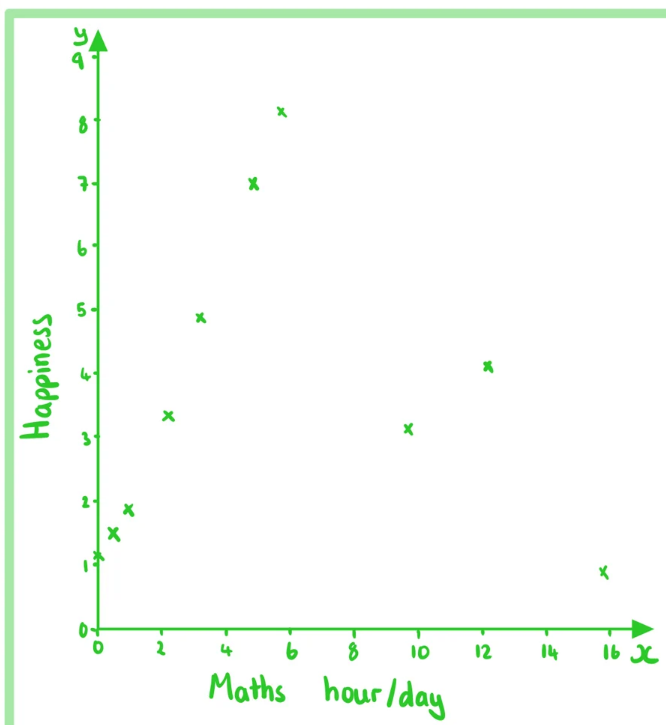
(c) State whether or not it would be appropriate to calculate a least squares regression line for the data in the table. Be sure to justify your answer.

[1]

(d) By interpreting the results of parts (a) and (b) above, remark on what the data set suggests with regard to the validity of Hrothgar's belief.

[2]

(a)



EXAM PAPERS PRACTICE

Hrothgar is a professional mathematician who believes that there is a strong positive correlation between a person's happiness and the amount of time that person spends solving complex mathematics problems. To test his theory Hrothgar collects data from 10 people on how much time they spend solving complex mathematics problems, along with each person's score on a standardised 'level of happiness' test. The results are shown in the table below:

Person	1	2	3	4	5	6	7	8	9	10
Maths (hours/day)	2.3	5.8	1.0	12.2	0	3.4	9.6	0.5	4.9	15.7
Happiness	3.3	8.1	1.9	4.2	1.1	4.8	3.2	1.5	7.0	0.8

(a) Draw a scatter diagram for the data in the table.

[3]

(b) Test at the 5% level whether it is reasonable to assume a positive linear correlation between the two variables, being sure to state your null and alternative hypotheses. You may assume that happiness scores follow a normal distribution.

[3]

(c) State whether or not it would be appropriate to calculate a least squares regression line for the data in the table. Be sure to justify your answer.

[1]

(d) By interpreting the results of parts (a) and (b) above, remark on what the data set suggests with regard to the validity of Hrothgar's belief.

[2]

(b) Let ρ be the population PMCC between maths and happiness

$H_0 : \rho = 0$

$H_1 : \rho > 0$ Testing for positive linear correlation

Enter data into GDC and use linear regression t-test

$p = 0.468097...$

Reject H_0 if p-value < significance level

$p = 0.468... > 0.05$ so insufficient evidence to reject H_0. This suggests that the there is not a positive linear correlation between maths and happiness.

Hrothgar is a professional mathematician who believes that there is a strong positive correlation between a person's happiness and the amount of time that person spends solving complex mathematics problems. To test his theory Hrothgar collects data from 10 people on how much time they spend solving complex mathematics problems, along with each person's score on a standardised 'level of happiness' test. The results are shown in the table below:

Person	1	2	3	4	5	6	7	8	9	10
Maths (hours/day)	2.3	5.8	1.0	12.2	0	3.4	9.6	0.5	4.9	15.7
Happiness	3.3	8.1	1.9	4.2	1.1	4.8	3.2	1.5	7.0	0.8

(a) Draw a scatter diagram for the data in the table.

[3]

(b) Test at the 5% level whether it is reasonable to assume a positive linear correlation between the two variables, being sure to state your null and alternative hypotheses. You may assume that happiness scores follow a normal distribution.

[3]

(c) State whether or not it would be appropriate to calculate a least squares regression line for the data in the table. Be sure to justify your answer.

[1]

(d) By interpreting the results of parts (a) and (b) above, remark on what the data set suggests with regard to the validity of Hrothgar's belief.

[2]

(c) The test suggests that there is no positive linear correlation, therefore a least squares regression line is not appropriate.

Hrothgar is a professional mathematician who believes that there is a strong positive correlation between a person's happiness and the amount of time that person spends solving complex mathematics problems. To test his theory Hrothgar collects data from 10 people on how much time they spend solving complex mathematics problems, along with each person's score on a standardised 'level of happiness' test. The results are shown in the table below:

Person	1	2	3	4	5	6	7	8	9	10
Maths (hours/day)	2.3	5.8	1.0	12.2	0	3.4	9.6	0.5	4.9	15.7
Happiness	3.3	8.1	1.9	4.2	1.1	4.8	3.2	1.5	7.0	0.8

(a) Draw a scatter diagram for the data in the table.

[3]

(b) Test at the 5% level whether it is reasonable to assume a positive linear correlation between the two variables, being sure to state your null and alternative hypotheses. You may assume that happiness scores follow a normal distribution.

[3]

(c) State whether or not it would be appropriate to calculate a least squares regression line for the data in the table. Be sure to justify your answer.

[1]

(d) By interpreting the results of parts (a) and (b) above, remark on what the data set suggests with regard to the validity of Hrothgar's belief.

[2]

(d) The test suggests there is no positive linear correlation. However the scatter diagrams suggests there is a positive linear correlation up to a limit, after which point there appears not to be positive linear correlation. Therefore Hrothgar's theory is not valid for all values but it might be true for a subset of the values.