# 4.12 Further Hypothesis Testing

## Question Paper

| Course | DP IB Maths |
|---|---|
| Section | 4. Statistics & Probability |
| Topic | 4.12 Further Hypothesis Testing |
| Difficulty | Medium |

**Time allowed:** 110

**Score:** /91

**Percentage:** /100

## Question 1a

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 mark]

## Question 1b

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

[2 marks]

## Question 1c

(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 marks]

## Question 1d

(d)  Find the critical region for the test.

**[2 marks]**

## Question 2a

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 mark]**

## Question 2b

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 marks]**

## Question 2c

(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 marks]**

## Question 2d

(d)

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

**[1 mark]**

## Question 3a

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.

**[5 marks]**

## Question 3b

b) Find the p-value for the test.

[2 marks]

## Question 3c

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

[2 marks]

## Question 4a

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 marks]

## Question 4b

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 marks]**

## Question 5a

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 marks]**

## Question 5b

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 mark]**

## Question 5c

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

**[3 marks]**

## Question 5d

(d)
Write down the conclusion to the test.

**[2 marks]**

## Question 6a

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 | | –1 | | | | |

**[2 marks]**

## Question 6b

(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 marks]**

## Question 6c

(c)
Write down the conclusion to the test.

**[2 marks]**

## Question 7a

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 mark]**

## Question 7b

(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 marks]**

## Question 7c

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 marks]**

## Question 8a

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 mark]**

## Question 8b

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

**[2 marks]**

## Question 8c

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 mark]**

## Question 8d

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

**[3 marks]**

## Question 9a

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 marks]**

## Question 9b

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 marks]**

## Question 9c

Out of the 24 observations, there are 19 successes.

(c)
State the conclusions of the hypothesis test.

**[2 marks]**

## Question 9d

(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 marks]**

## Question 10a

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 marks]**

## Question 10b

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 marks]**

## Question 10c

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

(c)
State the conclusions of the hypothesis test.

**[2 marks]**

## Question 11a

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 marks]**

## Question 11b

(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 marks]**

## Question 11c

(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 mark]**

## Question 11d

(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 marks]**