

4.12 Further Hypothesis Testing

Question Paper

Course	DPIB Maths
Section	4. Statistics & Probability
Topic	4.12 Further Hypothesis Testing
Difficulty	Very Hard

Time allowed: 120
Score: /92
Percentage: /100

Question 1a

A company currently gets an average of 25.3 views per minute for their website. Alex, a website developer, has been hired to increase the average number of views per hour for the company. After Alex make some changes, the director of the company, Reema, wants to test, using a 10% significance level, whether the average number of views per hour has increased. Reema monitors the site for 10 minutes and models the number of visits during this time using a Poisson distribution.

a)

(i) State the assumptions that are needed to use a Poisson model.

(ii) Write down null and alternative hypotheses for Reema's test.

[3 marks]

Question 1b

b)

Find the probability that Reema's test will lead to a Type I error.

[3 marks]

Question 1c

c)

Describe one way in which Reema could improve the reliability of her test.

[1 mark]

Question 1d

Alex claims the site now gets on average between 26.2 and 31.4 views per minute.

- d)
Assuming Alex's claim is correct, find the maximum probability that Reema's test will lead to a Type II error.

[3 marks]

Question 2a

At a funfair, Howard runs a duck game where there is a number of rubber ducks floating in water, each duck has a coloured sticker underneath. Players use a rod to hook one of the toy ducks and check the colour of the sticker. Players only win a prize if the colour of the sticker is purple. The duck is returned to the water after each game.

Previously it was known that Howard had 80 toy ducks and 28 of them had a purple sticker. Howard increases the number of ducks in the game. Donald, a regular player of the game, wants to test whether the chance of winning a prize has decreased so he conducts a hypothesis test with a 5% significance level. He observes the game being played 150 times and records the number of times a prize is won.

- a)
Write down the null and alternative hypotheses for Donald's test.

[2 marks]

Question 2b

- b)
Given that Howard added 10 ducks to the game, none of which have a purple sticker, find the probability that Donald will make a Type II error in his test conclusion.

[6 marks]

Question 2c

Out of the 150 games that Donald observed, 45 of them resulted in the player winning a prize.

c)

State, with a reason, the conclusion to Donald's test.

[2 marks]

Question 2d

d)

Donald repeats this test 200 times in total. Find the expected number of times that Donald will make the correct conclusion in his test.

[2 marks]

Question 3a

Wei, an economist, is investigating how the value of a vintage car changes over time. The table below shows the age of the car, t years, and its value, $\text{£}P$.

Age (t years)	0	2	3	4.5	6	7	8.5	10
Value ($\text{£}P$)	16 200	16 543	17 587	17 996	20 255	19 789	20 884	22 307

Wei believes that the relationship between the age and value will mean that there is a positive correlation between the variables t and $\ln P$.

a)

(i)

Complete the table below, giving all values correct to 3 decimal places.

t	0	2	3	4.5	6	7	8.5	10
$\ln P$		9.714	9.775	9.798	9.916		9.947	

(ii)

Use a hypothesis test with a 5% significance level to show that there is evidence of a positive linear correlation between the variables t and $\ln P$. State the hypotheses clearly and justify your conclusion.

(iii)

Hence state whether an exponential or power model would best represent the value of the car as its age varies.

[6 marks]

Question 3b

b)

Find the equation of the least squares regression model for P on t using the model identified in (a) (iii).**[2 marks]****Question 3c**

c)

Explain one limitation of the model.

[1 mark]

Question 4a

Two IB students, Naomi and Harry, are investigating the relationship between performance in Mathematics and Biology exams for students at their school. All students at their school study both Mathematics and Biology. They have recorded the most recent exam percentage scores for a sample of students. The results are shown in the table below.

Student	A	B	C	D	E	F	G	H
Mathematics	80	63	45	77	72	68	87	73
Biology	82	70	56	86	69	75	86	77

Harry decides to perform a pooled two sample t-test at the 5% significance level to determine if the mean score is higher in Biology than Mathematics.

- a)
- (i) State any assumptions that Harry needs to make.
 - (ii) State the null and alternative hypotheses for Harry's test.
 - (iii) State the conclusion of Harry's test. Justify your answer.

[6 marks]

Question 4b

Naomi decides to perform a paired t-test at the 5% significance level to determine if a student scores higher in Biology than Mathematics.

b)

- (i) State any assumptions that Naomi needs to make.

- (ii) State the null and alternative hypotheses for Naomi's test.

- (iii) State the conclusion of Naomi's test. Justify your answer.

[6 marks]

Question 4c

c)

Explain why the two tests seem to suggest different results.

[2 marks]

Question 5a

The amount of sleep, in hours, that Max gets in a night can be modelled using a normal distribution with standard deviation 0.94 hours. Max claims that the mean amount of sleep he gets each night, μ , is 7.24 hours. Max's three children, Wanda, Pietro, and Lorna, disagree with his claim. They each conduct a hypothesis test, with a 5% level of significance, using the same null hypothesis $H_0 : \mu = 7.24$.

Wanda uses the alternative hypothesis $H_1 : \mu < 7.24$.

Pietro uses the alternative hypothesis $H_1 : \mu \neq 7.24$.

Lorna uses the alternative hypothesis $H_1 : \mu > 7.24$.

To perform their tests, the three children take a random sample of 12 nights and calculate the mean amount of sleep, \bar{x} hours, that Max gets per night.

a)

In the case where the test provides sufficient evidence for Pietro to reject the null hypothesis, explain why there is sufficient evidence for exactly one of his sisters to also reject the null hypothesis.

[2 marks]

Question 5b

b)

In the case where there is sufficient evidence for only Wanda to reject the null hypothesis, find the range of values for \bar{x} . Give the endpoints of the interval to 4 significant figures.

[4 marks]

Question 5c

c)

In the case where the actual value of μ is 8.05, find the probability that Lorna will make a Type II error in her test conclusion.

[4 marks]

Question 6a

In the town of Pennyslavia, there is an increase in the number of fake coins being used. It is known that a real coin has a 50% chance of landing on tails when flipped and a fake coin has a 35% chance of landing on tails. Nick, a shopkeeper, suspects one of his coins is fake so he decides to conduct a test using the null hypothesis $H_0 : p = 0.5$ against the alternative hypothesis $H_1 : p = 0.35$. Nick will flip the coin 60 times and he will reject the null hypothesis if the coin lands on tails less than 25 times.

a)

(i) Explain what a Type I error is in the context of the question.

(ii) Find the probability that Nick makes a Type I error in his test conclusion.

[4 marks]

Question 6b

b)

(i)

Explain what a Type II error is in the context of the question.

(ii)

Find the probability that Nick makes a Type II error in his test conclusion.

[4 marks]**Question 6c**

40% of coins in Pennyslavia are fake.

c)

Find the probability that Nick makes an error in his test conclusion.

[2 marks]

Question 7a

In a fast-food restaurant there is a machine which cooks specialist fries. The number of times the machine breaks down can be modelled by a Poisson distribution with a mean rate of 1.2 times per hour. The manager wants to decrease the rate of breakdowns, so she purchases the newest model of the machine. She conducts a hypothesis test using a specified level of significance by monitoring the new machine for a six-hour period and recording the number of breakdowns.

a)

(i) Write down the null and alternative hypotheses for the manager's test.

(ii) Explain why the probability of a Type I error will be less than the specified level of significance.

(iii) Given that the probability of the manager making a Type I error is 0.07192 correct to four significant figures, find the critical value for the test.

[5 marks]

Question 7b

During the six-hour period, the new machine breaks down 4 times.

b)

State, with a reason, the conclusion to the manager's test.

[1 mark]

Question 7c

c)

Suggest one way in which the manager could improve the reliability of her test.

[1 mark]

Question 8a

The mass of an adult pig in England can be modelled using a normal distribution with a mean of 255 kg and a variance of 2000 kg². A pig is labelled as *supersized* if it weighs more than 350 kg.

- a)
Using the model, find the probability that a randomly selected adult pig in England is labelled as *supersized*.

[1 mark]

Question 8b

Ramon, a farmer, raises adult pigs in Scotland. He believes that the probability that one of his pigs is labelled *supersized* is higher than the probability given by the model for adult pigs in England. To test his belief Ramon randomly selects 12 pigs that he owns and finds that two of them are classed as *supersized*.

- b)
Stating the hypotheses clearly, test Ramon's belief using a 5% significance level.

[4 marks]

Question 8c

The mass of an adult pig in Ramon's farm can also be modelled by a normal distribution. Ramon believes that the mean mass is higher than 255 kg. Using the 12 pigs in his sample, Ramon calculates that the mean of the sample is 273 kg and the standard deviation of the sample is 52.1 kg.

c)

- (i) State the name of the test that Ramon should use. Justify your answer.
- (ii) Write down the null and alternative hypotheses for the test.
- (iii) State, with a reason, the conclusion of the test using a 5% significance level.

[5 marks]**Question 8d**

d)

Ramon claims that the two conclusions contradict each other. Explain why it is possible for both conclusions to be correct.

[2 marks]

Question 9a

There is a large cohort of students studying the statistics course at the University of Bernoulli. Students have one of two teachers: Jacob or Daniel. At the end of the term all students sit the same exam and receive a score. The scores of students in Jacob's class are known to be normally distributed with mean 72 and standard deviation 9. The scores of students in Daniel's class are normally distributed with mean 51 and standard deviation 5.

After the exams have been marked, they are anonymised so that the student or their teacher is not known. To determine whose class each student is in, Jacob and Daniel devise a hypothesis test. They use the null hypothesis that the student is in Jacob's class and will reject the null hypothesis if the score of the student is less than k .

They want to choose the value of k so that it satisfies the conditions:

- k is an integer,
- the probability of a Type I error is less than 5%,
- the probability of a Type II error is less than 15%.

a)

Show that there is only one value of k that satisfies the three conditions. Find the value of k .

[5 marks]

Question 9b

There are 35 students in Jacob's class and 45 students in Daniel's class.

b)

Estimate the number of students whose teacher will be correctly identified using Jacob and Daniel's test.

[3 marks]

