Statistical Hypothesis Testing

This topic covers hypothesis tests for a proportion in a binomial distribution — you saw the binomial distribution in the previous topic, so it should feel like an old friend. You also get to criticise people's sampling methods...

1 Splash Electronics Ltd. sell waterproof cameras. Before the company sends a batch of cameras out, they test a sample of cameras to see how long they can stay underwater for before water leaks in.

a) Give a reason why the company would test a sample of their cameras rather than the whole population.

.....

.....

(1 mark)

The company decides to test a random sample of 15 cameras from a batch of 300 cameras.
Every camera in the batch has a unique product code on the inside cover.

b) Describe how the company could obtain a simple random sample of size 15 from the batch of cameras.

.....

.....

.....

(3 marks)

2 Josie wants to find out what the most popular type of music is amongst the pupils in her school. She decides to ask a sample of the pupils.

a) Identify the population that Josie is interested in.

.....

(1 mark)

Mike says that it would be better to carry out a census, rather than take a sample.

b) Give one reason why he might have said this.

.....

.....

(1 mark)

Josie decides to select all the students in her GCSE music class as her sample.

c) Name the sampling method Josie is using.

.....

(1 mark)

d) Explain whether or not Josie's sample is likely to be representative of the population.

.....

.....

(1 mark)

Statistical Hypothesis Testing

3 Jamila is investigating the pay rises given to working adults in her town last year. The table below shows the number of working adults in Jamila's town.

Age (in years)	18-27	28-37	38-47	48-57	Over 57
No. of working adults	1200	2100	3500	3200	1500

a) Jamila plans to use stratified sampling to select a sample of 50 working adults from her town. Calculate how many people from each age group should be in the sample.

Remember to round decimals to the nearest whole number, and check your total is 50.

18-27 = ..., 28-37 = ..., 38-47 = ..., 48-57 = ..., Over 57 = ...

(4 marks)

b) Jamila wants to investigate the pay rises given to working adults across the UK last year. Can she use her sample data to draw conclusions about the whole population? Explain your answer.

.....

.....

(1 mark)

4 The residents of a town are being asked about a plan to build a wind farm in the area. Past records show that 10% of residents were against the plan. Campaigners claim that the proportion of residents who are against the plan has increased. A random sample of 50 residents is surveyed.

a) Write down suitable hypotheses to test the campaigners' claim.

.....

.....

(1 mark)

b) Of the sampled residents, 6 are against the plan. Using a 1% significance level, test whether there is evidence to suggest that the proportion of residents opposed to the plan has increased.

.....

.....

(5 marks)

Statistical Hypothesis Testing

5 Nate teaches judo classes at 'basic' and 'advanced' levels. Last year, the proportion, p, of his students who had done judo for at least two years was 0.2. Nate moves to a different judo club and claims that at this club, the proportion of his students who have done judo for at least two years has changed. To test this, he plans to survey a sample of 20 of his new students.

a) State suitable hypotheses that could be used to test Nate's claim.

.....

.....

(1 mark)

b) Find the critical region for a two-tailed test of Nate's claim at the 5% level of significance. The probability of rejection in each tail should be less than 0.025.

Use the binomial tables (see p97-101) or a calculator to find the values that give the correct probability in each tail.

.....

(3 marks)

c) Find the actual significance level of a test based on your critical region from part b).

.....

(1 mark)

d) Given that 7 of the sampled students have done judo for at least two years, carry out the test of Nate's claim.

.....

.....

.....

(2 marks)

e) To select his sample, Nate chose 20 students from the same judo class. Comment on the validity of the model used to carry out the test of Nate's claim. Explain your answer.

.....

.....

.....

.....

(2 marks)

6 For Heathrow during the period May to October 1987, it is known that 50% of the days had a maximum temperature of less than 19 °C. Adil wants to carry out a hypothesis test to investigate whether the proportion of days with a maximum temperature of less than 19 °C was different for the same period in 2015.

a) Adil uses the large data set to randomly sample 30 days from the period May to October 2015. 14 of the days had a maximum temperature of less than 19 °C.
Use this information to carry out Adil's test at the 10% level of significance.

You can either work out the p-value or find the critical region.

.....

.....

(6 marks)

b) Adil also plans to investigate the daily maximum temperature for Hurn during the period May to October 2015. He takes a sample from the large data set using systematic sampling. Suggest one reason why systematic sampling might give a more representative sample than simple random sampling.

.....

.....

(1 mark)

7 The records for 2016 suggest that 45% of the members of a gym use the swimming pool.
The gym's manager thinks that the popularity of the swimming pool has decreased over recent months.

a) Out of a random sample of 16 gym members, 3 of them use the pool. Using a 5% level of significance, test whether there is evidence to suggest that the popularity of the pool has decreased.

.....

.....

(6 marks)

b) The manager decides that the same test should be done again, but this time using a larger sample. He surveys a random sample of 50 members and carries out the test. He concludes that at the 5% level of significance there is evidence to suggest that the popularity of the pool has decreased.
Find the maximum possible number of gym members in the sample of 50 who use the pool.

.....

(4 marks)

EXAM TIP

With hypothesis tests, think carefully about whether it's easier to find the p-value or the critical region — and whether it's easier to use the binomial tables or your calculator functions. And make sure you always write a proper conclusion. For example, don't just say 'reject H_0', you also need to explain what rejecting H_0 means in the context of the question.

Score