Edexcel A Level Statistics: Hypothesis Testing Practice

Focus: Choosing the correct test and setting up hypotheses.

Section A: Questions

Question 1 A local council claims that 40% of residents cycle to work. A researcher believes this figure has decreased. She surveys a random sample of 30 residents and finds that 8 of them cycle to work.
a) State the distribution that should be used to model the test statistic in this trial.
b) Write down the Null and Alternative Hypotheses for this test.
c) Define the parameter p in context.

Question 2 A machine is set to produce metal rods with a mean length of 150mm and a standard deviation of 4mm. A quality control manager takes a random sample of 50 rods to check if the mean length has changed. The mean length of the sample is 148.8mm.
a) State the suitable test statistic and calculate its value.
b) Explain why a Normal distribution can be used for the sample mean, even if the distribution of individual rod lengths is unknown.
c) Write down the hypotheses for this test.

Question 3 A scientist is investigating the relationship between daily temperature (t) and the number of visitors to a park (v). He calculates the Product Moment Correlation Coefficient (PMCC) for 10 days of data and finds $r = 0.65$. He wants to test for a positive linear correlation at the 5% level of significance. (Critical value for $n = 10$ at 5% is 0.5494).
a) State the hypotheses in terms of ρ.
b) State the conclusion of the test in context.

Question 4 A six-sided die is suspected of being biased towards the number 6. The die is rolled 20 times.
a) Determine the critical region for the number of 6s rolled at the 5% level of significance.
b) If the number 6 appears 8 times, what is the conclusion of the test?

Question 5 Identify which hypothesis test (Binomial, Normal, or Correlation) is appropriate for the following scenarios:
a) Testing if the average weight of apples in a crate is less than 100g.
b) Testing if there is a link between a student's height and their score in a math test.
c) Testing if a coin lands on Heads more than 50% of the time.

Section B: Mark Scheme & Explanations

Q1 Mark Scheme

- **a) Binomial Distribution** $(B(30, 0.4))$
  - *Reasoning:* We are dealing with a number of successes (cyclists) in a fixed number of

trials.

- **b)** $H_0 : p = 0.4, \quad H_1 : p < 0.4$

  - *Note:* H_1 is less than ($<$) because the researcher believes the figure has *decreased*.

- **c)** p **is the probability that a randomly selected resident cycles to work.**

  - *Tip:* Always define parameters in the context of the question to ensure you get the mark.

Q2 Mark Scheme

- **a)** $Z = -2.12$

  - Formula: $Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

  - Calculation: $\frac{148.8 - 150}{4/\sqrt{50}} = \frac{-1.2}{0.5657} \approx -2.12$

- **b) The Central Limit Theorem (CLT)**

  - *Reasoning:* Since the sample size ($n = 50$) is large (> 30), the distribution of the sample mean approximates a Normal distribution regardless of the population's distribution.

- **c)** $H_0 : \mu = 150, \quad H_1 : \mu \neq 150$

  - *Reasoning:* The manager checks if the length has *changed* (could be up or down), so use $\neq$ (two-tailed).

Q3 Mark Scheme

- **a)** $H_0 : \rho = 0, \quad H_1 : \rho > 0$

  - *Reasoning:* Use ρ (rho) for population correlation. Use $>$ because he is testing for a *positive* correlation.

- **b) Reject** H_0**. There is evidence of a positive correlation.**

  - *Reasoning:* The test statistic $r = 0.65$ is greater than the critical value 0.5494. It falls in the critical region.

  - *Context:* There is evidence to suggest that as temperature increases, visitor numbers increase.

Q4 Mark Scheme

- **a) Critical Region:** $X \geq 7$

  - Method: We look for $P(X \geq x) < 0.05$ where $X \sim B(20, 1/6)$.

  - Calculation:

    - $P(X \geq 6) = 1 - P(X \leq 5) = 1 - 0.8982 = 0.1018$ (Not significant)

    - $P(X \geq 7) = 1 - P(X \leq 6) = 1 - 0.9629 = 0.0371$ (Significant)

  - Therefore, the critical region is $\{7, 8, \ldots, 20\}$.

- **b) Reject** H_0**. The die is biased.**

  - *Reasoning:* The observed value (8) is in the critical region ($8 \geq 7$).

Q5 Mark Scheme

- **a) Normal Test (Test for Mean)**

- *Keyword:* "Average weight" implies testing a mean (μ).

- **b) Correlation Test (PMCC)**

  - *Keyword:* "Link between... height and score" implies a relationship between two variables.

- **c) Binomial Test**

  - *Keyword:* "Coin lands on Heads" implies a count of successes/proportion.