Test answers - Hypothesis tests: correlation and Normal mean

Year: Year 13 · Duration: 40 minutes · /20

Answer to Question 1 (4 marks)

a. $H_0 : \rho = 0$, $H_1 : \rho \neq 0$ where ρ is the population correlation coefficient between temperature and ice lolly sales (2).
b. $0.58 > 0.5760$, so reject H_0 (1): there is sufficient evidence at the 5% level of correlation between temperature and ice lolly sales (1).

The test is two-tailed because the owner only asks whether there is correlation, not its direction. The value 0.58 only just exceeds 0.5760, so the evidence is weak: with 12 days of data, a slightly lower r would not be significant. Hypotheses must use ρ, not r, to score in (a).

Answer to Question 2 (4 marks)

Method: if $X \sim N(\mu, \sigma^2)$, the mean of a random sample of size n satisfies $\bar{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$. The standard error is $\dfrac{\sigma}{\sqrt{n}}$; standardise with $Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}}$. Mark scheme: 1 mark for part a), 1 mark for part b), 2 marks for part c).

a. The standard error is $\dfrac{15}{\sqrt{25}} = 3$. $\bar{X} \sim N(75, 3^2)$ (1).
b. Standardise: $z = \dfrac{72 - 75}{3} = -1$. $P(Z < -1) = 0.1587$ (1). So $P(\bar{X} < 72) \approx 0.159$.
c. Standardise: $z = \dfrac{80 - 75}{3} = 1.667$, then take the upper tail.
$P(\bar{X} > 80) = P(Z > 1.667) = 0.0478$ (2). So $P(\bar{X} > 80) \approx 0.0478$.

Answer to Question 3 (5 marks)

Method: state the hypotheses about μ, find the distribution of $\bar{X}$ under H_0, calculate the test statistic or p-value, compare with the significance level, then conclude in context. Mark scheme as shown in brackets.

a. $H_0 : \mu = 330$, $H_1 : \mu < 330$ (1). $\bar{X} \sim N\left(330, \dfrac{9}{40}\right)$, standard error 0.4743;
$z = \dfrac{-1.1}{0.4743} = -2.319$ (1). $p = 0.0102 < 0.05$ (1), so reject H_0: sufficient evidence that the mean volume is less than 330 ml (1).
b. $0.0102 > 0.01$, so at the 1% level we do not reject H_0: the conclusion changes (1).

So there is evidence at the 5% level, but not at the 1% level, that the bottles contain less than 330 ml on average.

Answer to Question 4 (4 marks)

Method: the critical region is the set of values of $\bar{x}$ that lead to rejecting H_0. For a two-tailed test at 5%, put 2.5% in each tail and use $z = \pm 1.96$ with the standard error $\frac{6}{\sqrt{16}} = 1.5$.

a. $\bar{X} \sim N(80, 1.5^2)$ (1). Critical values $80 \pm 1.96 \times 1.5 = 77.06$ and 82.94 (1). Critical region: $\bar{x} \leq 77.06$ or $\bar{x} \geq 82.94$ (1).
b. $83.5 > 82.94$, in the critical region: reject H_0 (1).

Conclusion in context: there is evidence at the 5% level that the population mean is not 80.

Answer to Question 5 (3 marks)

a. One-tailed: $0.41 > 0.3783$, reject H_0 (1). Two-tailed: $0.41 < 0.4438$, do not reject H_0 (1).
b. The one-tailed test puts the whole 5% in one tail, so its critical value is smaller; the direction must be decided before looking at the data (1).

So the same sample gives evidence of positive correlation with a one-tailed test but no evidence of correlation with a two-tailed test. Choosing the one-tailed version after seeing $r = 0.41$ would be cheating: the true significance level would no longer be 5%. The hypotheses for (a) are $H_0 : \rho = 0$ and $H_1 : \rho > 0$ or $H_1 : \rho \neq 0$.