2.9 Statistics: Hypothesis Testing Depth: 1-hour Pack

EDEXCEL 9MA0 · Calculator allowed · about 60 minutes

Total Marks

Name: _____ Date: ____ / ____ / _____

Answer ALL questions. Show all your working.

Hypothesis Testing Toolkit: Formulae and Standard Normal Critical Values

Original material written for Revision Library, styled on Edexcel A-level Mathematics (9MA0) Statistics content.

For a binomial test of H0: p = p0, use X ~ B(n, p0) and compute the relevant cumulative probability (a tail probability, or a p-value) directly, either from cumulative binomial tables or a calculator's binomial cdf function. For a normal test of H0: mu = mu0 with known sigma, standardise using z = (xbar - mu0)/(sigma/sqrt(n)) and compare z with a critical value, or convert to a critical region for xbar directly. For a correlation test of H0: rho = 0, compare the sample product moment correlation coefficient, r, with a critical value from a table of critical values for the sample size n (such a critical value will always be given where needed). Standard normal critical z-values used throughout this pack: one-tailed 10% (or two-tailed 20%) = 1.2816; one-tailed 5% (or two-tailed 10%, i.e. 5% in each tail) = 1.6449; one-tailed 1% (or two-tailed 2%) = 2.3263; two-tailed 5% (2.5% in each tail) = 1.9600; two-tailed 1% (0.5% in each tail) = 2.5758. When finding a binomial critical region, choose the region whose cumulative probability is as close as possible to the stated significance level, unless a question specifically requires the probability to remain below that level; the actual (achieved) significance level is almost never exactly equal to the nominal one, because the binomial distribution is discrete.

1 A machine at a bakery in Bristol is set to produce loaves such that 25% have a well-fired crust, according to specification. After maintenance work, the quality control manager, Tomasz, samples 18 loaves at random and finds that only 1 has a well-fired crust. Let X be the number of loaves, out of 18, with a well-fired crust, with X ~ B(18, p) under H0.

(a) State suitable hypotheses for a two-tailed test of whether the proportion of well-fired loaves has changed. **(2)**

(b) Assuming X ~ B(18, 0.25), find P(X ≤ 1). **(3)**

(c) Using a 10% significance level, test whether there is evidence that the proportion of well-fired loaves has changed. State your conclusion in context. **(2)**

(Total for Question 1 is 7 marks)

© Revision Library | revisionlibrary.co.uk

2 The proportion of a population with a certain characteristic was 0.3. A researcher believes that this proportion has decreased. In a random sample of 15 people, 1 have the characteristic.

(a) State the hypotheses for a one-tailed test. **(2)**

(b) Assuming H0 is true, find the probability of obtaining a value of X at least as extreme as 1, where X ~ B(15, 0.3). **(2)**

(c) Test the researcher's belief at the 10% level of significance, stating your conclusion in context. **(3)**

(Total for Question 2 is 7 marks)

---

3 A student records 9 pairs of values for the length of time spent revising and the test score. The product moment correlation coefficient is r = -0.585. The critical value of r for a one-tailed test with a 5% significance level and n = 9 is 0.5822.

(a) Write down suitable hypotheses to test for negative correlation. **(2)**

(b) Carry out the test and state your conclusion in context. **(3)**

(Total for Question 3 is 5 marks)

---

4 A manufacturer claims that 15% of its light bulbs last more than 1000 hours. A retailer believes the proportion is higher than claimed and tests a random sample of 40 bulbs. Let p be the true probability, and X the number in the sample, so under H0, X ~ B(40, 0.15).

(a) State suitable hypotheses for the test, in terms of p. **(2)**

(b) Using a 1% significance level, find the critical region for the test. **(3)**

(c) Write down the actual significance level of the test, and state the probability of a Type I error. **(2)**

(Total for Question 4 is 7 marks)

---

5 The mass of bags of flour produced at a factory in Preston is modelled as X ~ N(1000, 64) grams (mean 1000 g, standard deviation 8 g), as stated on the packaging. Following a change of flour supplier, quality inspector Olusegun wants to test whether the mean mass has changed, using a random sample of 25 bags.

(a) State suitable hypotheses, and find, in terms of xbar (the sample mean mass), the critical region for a two-tailed test at the 1% significance level. **(5)**

(b) Given that the mean mass of the sample of 25 bags is found to be 1003.2 g, comment on whether there is evidence that the mean mass has changed. **(2)**

(Total for Question 5 is 7 marks)

© Revision Library | revisionlibrary.co.uk

6 A researcher, Chen, investigates whether there is a linear relationship between the number of hours of daylight (x) and the number of visitors (y) to a National Trust garden in Kent, using data from a random sample of 12 days spread throughout the year. The product moment correlation coefficient for the sample is calculated as $r = 0.62$.

(a) State suitable hypotheses to test whether there is positive linear correlation between hours of daylight and number of visitors. **(2)**

(b) Given that the critical value for a one-tailed test at the 5% significance level, for a sample of size 12, is 0.4973, determine whether there is significant evidence of positive correlation. **(3)**

(c) Chen argues that the value of r shows that longer daylight hours directly cause more visitors. Comment on this claim. **(2)**

(Total for Question 6 is 7 marks)

---

7 An orchard in Herefordshire has historically produced apples with weights normally distributed with mean 180 g and standard deviation 22 g. After introducing a new irrigation system, the orchard manager, Liam, samples 50 apples at random and records a sample mean weight of 186.4 g. Assume the standard deviation is unchanged.

(a) State suitable hypotheses to test whether the irrigation system has changed the mean weight of the apples. **(2)**

(b) Show that the test statistic is $z = 2.06$ (awrt), and hence carry out the test at the 5% significance level, stating your conclusion clearly. **(5)**

(c) Liam decides to repeat the test using a 1% significance level instead. State, with a reason, whether his conclusion would change. **(2)**

(d) Explain what is meant by a Type I error in this context, and state the probability of a Type I error for the test in part (c). **(2)**

(Total for Question 7 is 11 marks)

© Revision Library | revisionlibrary.co.uk

Mark scheme · 2.9 Statistics: Hypothesis Testing Depth: 1-hour Pack

Question 1

(a) H0: p = 0.25, where p is the population proportion of well-fired loaves **B1**

(a) H1: p ≠ 0.25 (oe) **B1**

(a) Answer: H0: p = 0.25, H1: p ≠ 0.25

(b) attempt P(X=0) + P(X=1) using X ~ B(18, 0.25) **M1**

(b) $P(X = 0) = 0.75^{18} = 0.00564$ (awrt) **A1**

(b) P(X ≤ 1) = 0.0395 (awrt 3sf) **A1**

(b) Answer: P(X ≤ 1) = 0.0395 (3sf)

(c) compare 0.0395 with 0.05 (half of the 10% two-tailed significance level) **M1**

(c) 0.0395 < 0.05, so reject H0; evidence, at the 10% level, that the proportion of well-fired loaves has changed **A1ft**

(c) Answer: 0.0395 < 0.05, so reject H0: evidence the proportion of well-fired loaves has changed.

Question 2

(a) H0: p = 0.3 **B1**

(a) H1: p < 0.3 **B1**

(a) Answer: H0: p = 0.3, H1: p < 0.3

(b) P(X ≤ 1) **M1**

(b) 0.0353 **A1**

(b) Answer: 0.0353

(c) compares 0.0353 with 0.1 **M1**

(c) reject H0 **A1**

(c) there is sufficient evidence that the proportion has decreased **B1**

(c) Answer: reject H0

Question 3

(a) H0: ρ = 0 **B1**

(a) H1: ρ < 0 **B1**

(a) Answer: H0: ρ = 0, H1: ρ < 0

(b) compares |r| = 0.585 with the critical value 0.5822 in the correct tail **M1**

(b) reject H0 **A1**

(b) there is evidence of negative correlation **B1**

(b) Answer: reject H0

Question 4

(a) H0: p = 0.15 **B1**

(a) H1: p > 0.15 **B1**

(a) Answer: H0: p = 0.15, H1: p > 0.15

(b) uses the binomial distribution B(40, 0.15) to find cumulative probabilities near the tail **M1**

(b) P(X ≥ 13) = 0.004313 ≤ 0.01 **A1**

(b) critical region: X ≥ 13 **A1**

(b) Answer: critical region X ≥ 13

(c) actual significance level = 0.00431 **B1**

(c) this is the probability of a Type I error (rejecting H0 when it is true) **B1**

(c) Answer: 0.00431

Question 5

(a) H0: μ = 1000, H1: μ ≠ 1000 **B1**

(a) identify the two-tailed 1% critical value z = 2.5758 **M1**

(a) use xbar = $\mu \pm z^*\sigma/\sqrt{n}$ **M1**

(a) lower boundary = 995.9 (awrt) **A1**

(a) upper boundary = 1004.1 (awrt) **A1**

(a) Answer: Critical region: xbar < 995.9 or xbar > 1004.1 (g).

(b) 1003.2 does not lie in the critical region, since 1003.2 < 1004.1 **B1**

(b) insufficient evidence, at the 1% level, that the mean mass of the bags has changed (do not reject H0) **B1ft**

(b) Answer: 1003.2 is not in the critical region, so insufficient evidence the mean mass has changed.

Question 6

(a) H0: ρ = 0, where ρ is the population product moment correlation coefficient **B1**

(a) H1: ρ > 0 (oe) **B1**

(a) Answer: H0: ρ = 0, H1: ρ > 0

(b) compares r = 0.62 with the critical value 0.4973 **B1**

(b) 0.62 > 0.4973, so reject H0 **B1**

(b) conclusion in context: significant evidence, at the 5% level, of positive correlation between hours of daylight and the number of visitors **B1**

(b) Answer: 0.62 > 0.4973, so reject H0: significant evidence of positive correlation.

(c) correlation does not imply causation **B1**

(c) valid alternative explanation, e.g. a confounding/lurking variable such as warmer weather or school holidays could independently affect both daylight hours (via season) and visitor numbers, so a direct causal link cannot be concluded from correlation alone **B1**

(c) Answer: Correlation does not imply causation; a confounding variable (e.g. season or weather) may explain the association without daylight hours directly causing higher visitor numbers.

© Revision Library | revisionlibrary.co.uk

Question 7

(a) H0: $\mu = 180$, where μ is the population mean weight in grams **B1**

(a) H1: $\mu \neq 180$ (oe) **B1**

(a) Answer: H0: $\mu = 180$, H1: $\mu \neq 180$

(b) standardise, $z = (\text{xbar} - \mu)/(\sigma/\sqrt{n})$ **M1**

(b) $z = 2.06$ (2.057 to 3dp) cso **A1**

(b) critical value ± 1.96 (two-tailed, 5%) stated for comparison **B1**

(b) valid comparison, $|z| > 1.96$ (2.057 > 1.96) **M1**

(b) reject H0; evidence, at the 5% level, that the mean weight of the apples has changed **A1ft**

(b) Answer: z = 2.06 (2.057, 3dp); 2.057 > 1.96, so reject H0: evidence the mean weight has changed.

(c) new critical value ± 2.5758 (two-tailed, 1%) identified, and $2.057 < 2.5758$ **B1**

(c) conclusion changes: insufficient evidence, at the 1% level, that the mean weight of the apples has changed (do not reject H0) **B1**

(c) Answer: 2.057 < 2.5758, so at the 1% level there is insufficient evidence of a change (the conclusion changes).

(d) a Type I error is rejecting H0 (concluding the mean weight has changed) when H0 is actually true (the mean weight of the apples is still 180 g) **B1**

(d) P(Type I error) = 0.01 **B1**

(d) Answer: A Type I error is rejecting H0 when it is true; P(Type I error) = 0.01 for the 1% level test.

© Revision Library | revisionlibrary.co.uk