Worksheet - Hypothesis tests: correlation and Normal mean

Year: Year 13 · 22

Show your working.

Question 1

Question 1 – Writing hypotheses *(Section A · Skills practice)*

Write down suitable null and alternative hypotheses for each situation, defining any symbols you use.

a. A manufacturer claims that the mean mass of a bag of flour is 1.5 kg. An inspector thinks the mean is lower. [2 marks]
b. A teacher believes there is a positive correlation between hours of revision and test mark. [2 marks]
c. A café owner wants to check whether the mean waiting time has changed from 4 minutes. [2 marks]

Question 2

Question 2 – Testing for correlation *(Section A · Skills practice)*

A sample of 10 pairs of values gives a product moment correlation coefficient $r = 0.58$. The critical values for $n = 10$ are 0.5494 (5%, one-tailed) and 0.6319 (5%, two-tailed).

a. Test at the 5% level for positive correlation. [3 marks]
b. Test at the 5% level whether there is any correlation (two-tailed). [3 marks]

Question 3

Question 3 – A negative correlation *(Section A · Skills practice)*

For $n = 12$ the 5% critical values are 0.4973 (one-tailed) and 0.5760 (two-tailed). A sample gives $r = -0.55$.

a. Test, at the 5% level, for negative correlation. [3 marks]
b. Test, at the 5% level, for correlation in either direction. [3 marks]

Question 4

Question 4 – The sample mean *(Section A · Skills practice)*

$X \sim N(50, 12^2)$ and a random sample of 16 observations is taken. $\bar{X}$ is the sample mean.

a. State the distribution of $\bar{X}$. [2 marks]
b. Find $P(\bar{X} > 53)$. [2 marks]
c. Find $P(\bar{X}$. [2 marks]

Question 5

Question 5 – A critical region *(Section A · Skills practice)*

A population is $N(\mu, 9^2)$. A sample of 36 is taken to test $H_0 : \mu = 120$ against $H_1 : \mu$ at the 1% significance level.

a. State the distribution of $\bar{X}$ under H_0. [1 mark]
b. Find the critical region for $\bar{X}$. [3 marks]
c. A sample has mean 117.2. State the conclusion. [2 marks]

Question 6

Question 6 – A p-value for a two-tailed test *(Section A · Skills practice)*

$X \sim N(\mu, 6^2)$. A sample of 20 has mean 42.5. Test $H_0 : \mu = 40$ against $H_1 : \mu \neq 40$ at the 5% significance level.

a. State the distribution of $\bar{X}$ under H_0 and calculate the test statistic. [3 marks]
b. Find the p-value. [2 marks]
c. State your conclusion in context. [2 marks]

Question 7

Question 7 – Revision and results *(Section B · Real-life practice)*

Eight Year 13 students recorded the number of hours h they spent revising for a mock exam and their percentage mark m.

Hours h	2	4	5	7	8	10	12	13
Mark m	41	50	48	62	60	71	78	80

a. Use your calculator to find r. [1 mark]
b. The 5% one-tailed critical value for $n = 8$ is 0.6215. Test whether there is positive correlation. [3 marks]
c. A student says "this proves that revising causes higher marks". Comment. [2 marks]

Question 8

Question 8 – Hot chocolate sales *(Section B · Real-life practice)*

A café records the outside temperature T (degrees Celsius) at noon and the number of

hot chocolates H sold in the afternoon on ten days.

T	2	5	8	11	14	17	20	23	26	29
H	94	90	81	75	70	58	52	41	38	30

a. Calculate r. [1 mark]
b. The 1% one-tailed critical value for $n = 10$ is 0.7155. Test for negative correlation at the 1% level. [3 marks]
c. The owner uses the line of best fit to predict sales when $T = 40$. Explain why this is unreliable. [1 mark]

Question 9

Question 9 – Bags of sweets *(Section B · Real-life practice)*

A sweet factory states that the mass of sweets in a bag is $N(200, 4^2)$ grams. A consumer group weighs 20 bags and finds a mean of 198.6 g. Assume the standard deviation is unchanged.

a. Write down suitable hypotheses. [2 marks]
b. Find the p-value. [3 marks]
c. Test at the 5% level and then at the 10% level. [3 marks]

Question 10

Question 10 – Journey times to school *(Section B · Real-life practice)*

The time for a school bus journey is normally distributed with mean 42 minutes and standard deviation 6 minutes. After roadworks begin, the mean of 30 randomly chosen journeys is 44.1 minutes. Assume the standard deviation is unchanged.

a. Test, at the 5% level, whether the mean journey time has increased. [5 marks]
b. State the probability of a Type I error. [1 mark]
c. Explain why the test uses the distribution of $\bar{X}$ rather than X. [1 mark]

Question 11

Question 11 – Battery life *(Section B · Real-life practice)*

A phone manufacturer claims that the battery lasts a mean of 120 hours between charges, with standard deviation 15 hours. A review site tests 25 phones and finds a mean of 114.2 hours.

a. Test at the 5% level whether the claim is too high. [5 marks]
b. Would the conclusion be the same at the 1% level? [2 marks]
c. State one assumption you have made. [1 mark]

Question 12

Question 12 – A reading course *(Section B · Real-life practice)*

Before a reading course, the reading speed of adults is normally distributed with mean 230 words per minute and standard deviation 30. After the course, the mean speed of a random sample of 36 adults is 241 words per minute. Assume the standard deviation is unchanged.

a. Test at the 5% level whether the course has changed the mean reading speed. [6 marks]
b. Find the critical region for $\bar{X}$. [3 marks]
c. Say, with a reason, whether the course has been shown to cause the change. [1 mark]

Question 13

Question 13 – Swimming times *(Section B · Real-life practice)*

The table shows the age a (years) and the time t (seconds) taken by ten club swimmers to swim 100 metres.

a	10	11	12	13	14	15	16	17	18	19
t	88	85	84	79	80	75	76	72	73	70

a. Find r. [1 mark]
b. Test at the 5% level (two-tailed, critical value 0.6319) for correlation. [3 marks]
c. Explain why the data do not suggest that a swimmer will keep improving after age 19. [2 marks]

Question 14

Question 14 – Fertiliser trial *(Section B · Real-life practice)*

Without fertiliser a tomato plant grows by a mean of 12.0 cm in a week with standard deviation 2.4 cm. A gardener applies a new fertiliser to 16 plants and measures a mean growth of 13.1 cm.

a. Test at the 5% level whether the fertiliser increases growth. [5 marks]
b. Repeat the test at the 1% level. [2 marks]
c. State two improvements to the design of the trial. [2 marks]

Question 15

Question 15 – Correlation from summary statistics *(Section C · Problem solving)*

A researcher records the monthly rainfall x (mm) and the number of visitors y (hundreds) to a museum for 12 months. The summary statistics are $S_{xx} = 250$, $S_{yy} = 480$ and $S_{xy} = -190$.

a. Calculate r. [2 marks]
b. The 5% critical values for $n = 12$ are 0.4973 (one-tailed) and 0.5760 (two-tailed). Test for negative correlation at 5%, then test for correlation of either sign at 5%. [5 marks]
c. The visitor numbers are recoded as $z = \dfrac{y - 20}{5}$. State the value of r between x and z, with a reason. [2 marks]
d. Discuss which test the researcher should have used if no direction had been predicted before the data was collected. [2 marks]

Question 16

Question 16 – Critical region and error probability *(Section C · Problem solving)*

A wire manufacturer states that the tensile strength of its wire is normally distributed with mean 50 and standard deviation 10 (units of 100 newtons). The quality team takes samples of 25 and tests $H_0 : \mu = 50$ against $H_1 : \mu > 50$ at the 5% level.

a. Find the critical region for $\bar{X}$. [4 marks]
b. A sample has mean 54. State the conclusion. [2 marks]
c. State the probability of rejecting H_0 when $\mu = 50$. [1 mark]
d. In fact the true mean has risen to 55. Find the probability that the test correctly rejects H_0. [3 marks]

Question 17

Question 17 – How large a sample? *(Section C · Problem solving)*

The lengths of bolts from a machine are normally distributed with standard deviation 8 mm. An engineer wants the sample mean of n bolts to be within 1 mm of the true mean with probability at least 0.95.

a. Show that $P(|\bar{X} - \mu|$. [3 marks]
b. Find the smallest n. [4 marks]
c. The engineer instead takes 100 bolts. Find the probability that $\bar{X}$ is within 1 mm of the true mean. [3 marks]

Question 18

Question 18 – The effect of an outlier *(Section C · Problem solving)*

A scientist plots the dose x of a drug against the response y for eight patients.

x	1	2	3	4	5	6	7	8
y	3	5	4	7	8	9	11	30

a. Show that $r = 0.792$ for all eight points. [1 mark]
b. The 5% one-tailed critical value for $n = 8$ is 0.6215. Test for positive correlation.

[3 marks]

c. The point $(8, 30)$ is an outlier. Without it, $r = 0.968$ for $n = 7$ (critical value 0.6694). Test again. [2 marks]
d. Comment on the effect of the outlier on r and on the conclusions. [2 marks]

Question 19

Question 19 – Interpreting three tests *(Section C · Problem solving)*

A school science department runs three hypothesis tests on a normal mean, each with $H_0 : \mu = \mu_0$.

Test	Context	Alternative	p-value
A	Plant height	$\mu > \mu_0$	0.032
B	Reaction time	$\mu \neq \mu_0$	0.081
C	Pulse rate	μ	0.0049

a. For each test, state the conclusion at the 5% level. [3 marks]
b. For each test, state the conclusion at the 1% level. [3 marks]
c. Test B is two-tailed. State the value that should be compared with 5% if the p-value shown is for one tail only, and give the conclusion. [2 marks]
d. The teacher says "Test C proves that the mean pulse rate is lower". Criticise this statement. [2 marks]

Question 20

Question 20 – Running several tests *(Section C · Problem solving)*

A school runs 5 independent hypothesis tests on different classes, each at the 5% significance level. In every case the null hypothesis is actually true.

a. Find the probability that none of the tests rejects H_0. [2 marks]
b. Find the probability that at least one test wrongly rejects H_0. [2 marks]
c. Find the probability that exactly one test wrongly rejects H_0. [3 marks]
d. Explain why running many tests increases the risk of a false conclusion, and suggest a remedy. [2 marks]

Question 21

Question 21 – Challenge: how big must the sample be? *(Section D · Challenge)*

A population is $N(\mu, 10^2)$. A researcher tests $H_0 : \mu = 50$ against $H_1 : \mu > 50$ at the 5% level using a sample of size n. The sample mean turns out to be 52.

a. Show that H_0 is rejected only if $n \geq 68$. [5 marks]
b. Explain what this shows about the link between sample size and significance. [2 marks]

Question 22

Question 22 – Challenge: correlation and decision *(Section D · Challenge)*

A sample of 15 pairs gives $r = 0.45$. For $n = 15$ the 5% critical values are 0.4409 (one-tailed) and 0.5140 (two-tailed); the 1% one-tailed value is 0.5923.

a. State the conclusion of a one-tailed test for positive correlation at the 5% level and at the 1% level. [3 marks]
b. State the conclusion of a two-tailed test at the 5% level. [2 marks]
c. A student says "the result shows weak correlation, so there is no correlation". Explain the error. [2 marks]