Past-paper walk-through

Hypothesis tests ▪ 6.5

eXercise

Questions in this set

- 9709/61 N25 Q2 ▪ 8 marks
- 9709/61 N25 Q5 ▪ 6 marks
- 9709/62 N25 Q4 ▪ 7 marks
- 9709/62 N25 Q6 ▪ 12 marks
- 9709/61 J25 Q3 ▪ 8 marks
- 9709/61 J25 Q6 ▪ 7 marks

Reading the mark scheme

- **B1** a mark that stands on its own – the point is either made or not.
- **M1** a *method* mark: the right approach, even if the number is wrong.
- **A1** an *answer* mark, and it usually needs its M mark first.
- **C1** a working mark on the way to the answer.
- // separates answers the examiner accepts equally.
- **ecf** = error carried forward: a later part is marked on your own earlier answer.
- **owtte** = "or words to that effect" – the wording need not match.

Question 2

9709/61 N25 ▪ Q2 ▪ 8 marks

The mean mass of packets of Trueleaf tea is supposed to be 500 grams. An inspector weighs 60 randomly chosen packets, giving $n = 60$, $\sum x = 29970$, $\sum x^2 = 14970300$. Test, at the 5% significance level, whether the population mean mass is 500 grams. **[8]**

Solution 2

Worked steps

Four stages, and every one of them is marked. Set them out as four labelled steps.

1. Unbiased estimates from the summary statistics.

- $\bar{x} = \dfrac{29970}{60} = 499.5$
- $s^2 = \dfrac{1}{n-1}\left(\sum x^2 - \dfrac{(\sum x)^2}{n}\right) = \dfrac{1}{59}\left(14\,970\,300 - \dfrac{29970^2}{60}\right) = 4.831$

Solution 2

2. **The hypotheses.** "Whether the mean is 500" gives no direction, so the test is **two-tailed**:
   - $H_0 : \mu = 500$
   - $H_1 : \mu \neq 500$
3. **The test statistic.**
   - $z = \dfrac{\bar{x} - \mu}{s/\sqrt{n}} = \dfrac{499.5 - 500}{\sqrt{4.831/60}} = -1.76$

Solution 2

4. **Compare and conclude.** At 5% two-tailed the critical values are ± 1.96:
   - $|-1.76| < 1.96$, so **do not reject** H_0.
   - **In context**: there is insufficient evidence, at the 5% level, that the population mean mass differs from 500 grams.

Three things separate full marks from most of them:

- **The divisor is** $n - 1 = 59$ for the unbiased variance estimate, not 60.
- **Two-tailed** because the question asks *whether the mean is* 500, not whether it is less. A one-tailed critical value of 1.645 would reject, which is the wrong conclusion — the tail choice changes the answer here.
- **Conclude in context and without overclaiming.** "Do not reject H_0" is not "the mean is 500"; the honest wording is *insufficient evidence that it differs*.

Solution 2

Mark scheme

- $\bar{x} = 499.5$, $s^2 = \frac{1}{59}\left(14970300 - \frac{29970^2}{60}\right) = 4.831$. $H_0 : \mu = 500$, $H_1 : \mu \neq 500$. $z = \frac{499.5-500}{\sqrt{4.831/60}} = -1.76$; $|z| < 1.96$, so do not reject H_0 — no evidence the mean differs from 500 g.

Question 5

9709/61 N25 ▪ Q5 ▪ 6 marks

It is known that 20% of households in a certain country contain more than 4 people. Laxmi believes that in her town the percentage is lower than 20%. She takes a random sample of 40 households and notes the number containing more than 4 people, then carries out a test at the 2.5% significance level using a binomial distribution.

(a) Find the probability of a Type I error. **[4]**

(b) State the rejection region for the test. **[1]**

(c) Laxmi finds that exactly 2 households in her sample contain more than 4 people. Explain why it is impossible for Laxmi to make a Type II error. **[1]**

Solution 5

(a)

Mark scheme

- $X \sim B(40, 0.2)$, lower-tail. $P(X \leqslant 2) = 0.00794 < 0.025$ but $P(X \leqslant 3) = 0.0285 > 0.025$, so the rejection region is $X \leqslant 2$ and $P(\text{Type I error}) = P(X \leqslant 2) = 0.00794$.

Solution 5

(b)

Mark scheme

- Reject H_0 if $X \leqslant 2$.

Solution 5

(c)

Mark scheme

- $X = 2$ lies in the rejection region, so H_0 is rejected. A Type II error means failing to reject a false H_0; since H_0 is rejected, a Type II error is impossible.

Question 4

9709/62 N25 ▪ Q4 ▪ 7 marks

An inspector believes that 18% of cups made at a certain factory contain flaws. The factory owner claims that the true percentage is less than 18%. The inspector examines a random sample of 40 cups and finds that 3 of them contain flaws.

(a) Stating a necessary assumption, use a binomial distribution to test the factory owner's claim at the 5% significance level. **[6]**

(b) Explain why it would **not** be appropriate to use the Poisson approximation to the binomial distribution to carry out the test in part (a). **[1]**

Solution 4

(a)

Mark scheme

- Assume cups are flawed independently. $H_0 : p = 0.18$, $H_1 : p < 0.18$, $X \sim B(40, 0.18)$. $P(X \leqslant 3) = 0.0542 > 0.05$, so do not reject H_0 — insufficient evidence that the percentage is less than 18

Solution 4

(b)

Mark scheme

- The Poisson approximation requires p to be small (with n large); here $p = 0.18$ is not small, so it is not appropriate.

Question 6

9709/62 N25 ▪ Q6 ▪ 12 marks

6 The weekly profit, in dollars, made by a certain firm has a normal distribution. In the past, the weekly profit had the distribution $N(736, 26^2)$. Following a change in management, the mean weekly profit for 35 randomly chosen weeks is \$725.

(a) Stating a necessary assumption, test at the 2% significance level whether the mean weekly profit has decreased. [6]

Question 6

The mean weekly profit for another random sample of 35 weeks is found and a similar test is carried out at the 2% significance level.

(b) State the probability of a Type I error. [1]

Question 6

(c) Given that the mean weekly profit is now in fact \$718, find the probability of a Type II error. [5]

Additional page

If you use the following page to complete the answer to any question, the question number must be clearly shown.

Solution 6

(a) Worked steps

The question asks for an **assumption**, so state it before testing.

- Assume the **population standard deviation is unchanged** at 26, and that the distribution is still **normal**.

Then the four stages:

- $H_0 : \mu = 736$; $H_1 : \mu < 736$ — **one-tailed**, because the question asks whether the profit has *decreased*.
- $z = \dfrac{725 - 736}{26/\sqrt{35}} = \dfrac{-11}{4.395} = -2.50$
- At the 2% level, one-tailed, the critical value is -2.054.
- $-2.50 < -2.054$, so **reject** H_0: there is evidence, at the 2% level, that the mean weekly profit has decreased.

Solution 6

Two marks that are easy to drop. The **assumption** is worth stating explicitly — it is why using σ rather than an estimate from the sample is legitimate. And the **direction**: "decreased" is one-tailed, so the critical value is -2.054 and not -2.326.

Mark scheme

- Assume the standard deviation is unchanged (26) and the distribution still normal. $H_0 : \mu = 736$, $H_1 : \mu < 736$. $z = \dfrac{725 - 736}{26/\sqrt{35}} = -2.50 < -2.054$, so reject H_0: there is evidence the mean weekly profit has decreased.

Solution 6

(b)

Mark scheme

- Probability of a Type I error $= 0.02$ (the significance level).

Solution 6

(c) Worked steps

A **Type II error** is failing to reject H_0 **when it is false**. So find the region where the test does not reject, then find its probability under the **true** mean.

1. The critical mean — the boundary of the rejection region under H_0:

- $\bar{x}_{\text{crit}} = 736 - 2.054 \times \dfrac{26}{\sqrt{35}} = 736 - 9.03 = 726.97$
- The test does **not** reject when $\bar{X} > 726.97$.

Solution 6

2. That probability, using the true mean of 718:

- $P(\text{Type II}) = P\big(\bar{X} > 726.97 \mid \mu = 718\big) = P\left(Z > \dfrac{726.97 - 718}{26/\sqrt{35}}\right)$
- $= P(Z > 2.04) = \mathbf{0.0206}$

The structure is what to remember: **the critical value comes from H_0, the probability comes from the true mean.** Using one mean for both is the error the question is built around, and it gives an answer that looks reasonable.

Solution 6

Keep the **direction** consistent: since the test rejects for *small* means, not rejecting means $\bar{X}$ is **above** the critical value.

Mark scheme

- Critical mean $\bar{x} = 736 - 2.054 \cdot \dfrac{26}{\sqrt{35}} = 726.97$. With true mean 718:
  $P(\text{Type II}) = P(\bar{X} > 726.97) = P\left(Z > \dfrac{726.97 - 718}{26/\sqrt{35}}\right) = P(Z > 2.04) = 0.0206$.

Question 3

9709/61 J25 ▪ Q3 ▪ 8 marks

The time, T minutes, for a certain daily bus journey is normally distributed. The bus company claims that the mean of T is 45. A passenger believes that the mean of T is actually greater than 45. She notes the times taken for this journey on a random sample of 60 days. The results are summarised below: $n = 60$, $\sum t = 2750$, $\sum t^2 = 127000$.

(a) Calculate unbiased estimates of the population mean and variance. **[3]**

(b) Test the passenger's belief at the 5% significance level. **[5]**

Solution 3

(a) Worked steps

Two formulae, and the divisor is what makes each estimate **unbiased**.

- $\widehat{\mu} = \bar{x} = \dfrac{2750}{60} = \dfrac{275}{6} = \mathbf{45.8}$
- $\widehat{\sigma}^2 = \dfrac{n}{n-1}\left(\dfrac{\sum x^2}{n} - \bar{x}^2\right) = \dfrac{60}{59}\left(\dfrac{127\,000}{60} - \left(\dfrac{275}{6}\right)^2\right) = \mathbf{16.2}$

The $\dfrac{n}{n-1}$ factor is the whole point of the word **unbiased**, and it is the mark: the sample variance systematically understates the population's, and multiplying by $\dfrac{n}{n-1}$ corrects it.

Solution 3

Keep $\bar{x}$ as the exact fraction $\frac{275}{6}$ while squaring it. Rounding to 45.8 first and then squaring loses accuracy that shows up in the third significant figure of the variance — and the test statistic in part (b) is sensitive to it. **Mark scheme**

- $\widehat{\mu} = \dfrac{2750}{60} = \dfrac{275}{6} = 45.8$; $\widehat{\sigma}^2 = \dfrac{60}{59}\left(\dfrac{127000}{60} - \left(\dfrac{275}{6}\right)^2\right) = 16.2$ (i.e. $\frac{2875}{177}$).

Solution 3

(b) Worked steps

Four stages, all marked.

- H_0: mean $= 45$; H_1: mean > 45 — **one-tailed**, since the passenger believes the journey takes *longer*.
- $z = \dfrac{45.83 - 45}{\sqrt{16.24/60}} = \dfrac{0.833}{0.520} = 1.60$
- Critical value at the 5% level, one-tailed: 1.645.
- $1.60 < 1.645$, so **do not reject** H_0: there is insufficient evidence, at the 5% level, that the mean journey time exceeds 45 minutes.

Solution 3

Three points:

- **Use the estimates from part (a)**, unrounded — with 16.2 rounded down the statistic drifts, and this test is close enough to the boundary for that to matter.
- **Divide by $\sqrt{\widehat{\sigma}^2/n}$**, the standard error, not by the standard deviation.
- **Conclude in context, and do not overclaim.** Failing to reject is not proof that the mean is 45; it is a failure to find evidence against it. With $z = 1.60$ against a boundary of 1.645 the result is close, and saying so is a fair observation.

Solution 3

Mark scheme

- H_0: mean $= 45$, H_1: mean > 45. $z = \dfrac{45.83 - 45}{\sqrt{16.24/60}} = 1.60$. Since $1.60 < 1.645$, do not reject H_0: insufficient evidence that the mean journey time exceeds 45 minutes.

Question 6

9709/61 J25 ▪ Q6 ▪ 7 marks

A manufacturer of cell phones claims that 25% of students own a Pumpkin phone. Jeyeraj thinks that the proportion of students at his large college who own a Pumpkin phone is less than 25%. He plans to test the manufacturer's claim. He chooses a random sample of 30 students at his college. If the number of students who own a Pumpkin phone is less than 5, Jeyeraj will reject the manufacturer's claim.

(a) State suitable hypotheses for the test. **[1]**

(b) Given that the true proportion of students at the college who own a Pumpkin phone is 10%, use a binomial distribution to find the probability of a Type II error. **[3]**

(c) At Florence's college, in a random sample of 40 students, it was found that 5 own a Pumpkin phone. Calculate an approximate 95% confidence interval for the proportion of students at Florence's college who own a Pumpkin phone. **[3]**

Solution 6

(a)

Mark scheme

- H_0: $p = 0.25$; H_1: $p < 0.25$ (proportion of students owning a Pumpkin phone).

Solution 6

(b)

Mark scheme

- True $p = 0.1$, $X \sim B(30, 0.1)$. Rejection region is $X < 5$, so Type II error $= \mathrm{P}(X \geqslant 5) = 1 - \mathrm{P}(X \leqslant 4) = 0.175$.

Solution 6

(c)

Mark scheme

- $\hat{p} = \frac{5}{40} = 0.125.$ 95