Question 1

Worked Solution

A random sample of 8 chocolates: 5 had hard centres. Test at 5% whether proportion with hard centres is not 30%.

Step 1: State hypotheses.

$$H_0 : p = 0.3 \qquad H_1 : p \neq 0.3$$

where p is the population proportion of chocolates with hard centres.

Step 2: Define test statistic and distribution under H_0.

Let X = number of chocolates (out of 8) with hard centres. Under H_0: $X \sim \mathrm{B}(8, 0.3)$.

Step 3: Find the relevant tail probability.

The observed value is $x = 5$. Since this is above the mean ($np = 2.4$), we test the upper tail:

$$\mathrm{P}(X \geq 5) = 1 - \mathrm{P}(X \leq 4) = 1 - 0.9420 = 0.0580$$

(Also check lower critical region: $\mathrm{P}(X \leq 0) = (0.7)^8 = 0.0576$; compare with 0.025.)

Step 4: Compare with critical value.

For a two-tailed test at 5%, compare each tail probability with 0.025.

Upper tail: $\mathrm{P}(X \geq 5) = 0.0580 > 0.025$.

Also: $\mathrm{P}(X \geq 6) = 1 - \mathrm{P}(X \leq 5) = 1 - 0.9887 = 0.0113 < 0.025$.

So the upper critical region is $X \geq 6$. The observed value $x = 5$ does not lie in the critical region.

Step 5: Conclusion.

Do not reject H_0. There is insufficient evidence at the 5% significance level that the proportion of chocolates with hard centres is different from 30%.

Question 2

Worked Solution

Part (i): $R \sim \mathrm{B}(6, p)$, observed $r = 6$, test $H_0 : p = 0.45$ vs $H_1 : p \neq 0.45$ at 5%

Step 1: Under H_0: $R \sim \mathrm{B}(6, 0.45)$. Observed value $r = 6$ is in the upper tail.

$$\mathrm{P}(R \geq 6) = (0.45)^6 = 0.00830\ldots \approx 0.0083$$

Step 2: Compare with 0.025 (two-tailed at 5%):

$0.0083 < 0.025$, so $r = 6$ lies in the upper critical region.

Step 3: Conclusion.

$\mathrm{P}(R = 6) = 0.0083 < 0.025$. Reject H_0. There is sufficient evidence at the 5% level that $p \neq 0.45$.

Part (ii): $S \sim \mathrm{B}(n, 0.45)$, observed $s = 1$. Find largest n for which H_0 is not rejected.

We need $\mathrm{P}(S \leq 1) > 0.025$ (so that $s = 1$ does not lie in the lower critical region).

$\mathrm{P}(S \leq 1) = (0.55)^n + n(0.45)(0.55)^{n-1}$

Test values:

$$n = 9 : \quad \mathrm{P}(S \leq 1) = (0.55)^9 + 9(0.45)(0.55)^8 = 0.00751 + 0.0555 = 0.0385 > 0.025$$
$$n = 10 : \quad \mathrm{P}(S \leq 1) = (0.55)^{10} + 10(0.45)(0.55)^9 = 0.00253 + 0.02330 = 0.0233 < 0.025$$

At $n = 9$, $\mathrm{P}(S \leq 1) = 0.0385 > 0.025$ so H_0 is not rejected. At $n = 10$, $\mathrm{P}(S \leq 1) = 0.0233 < 0.025$ so H_0 would be rejected.

The largest value of n for which H_0 is not rejected is $n = \mathbf{9}$.

Question 3

Worked Solution

Part (i): Test $H_0 : p = 0.35$ vs $H_1 : p < 0.35$; $n = 14$, $x = 2$ at 2.5%

Step 1: Hypotheses.

$$H_0 : p = 0.35 \qquad H_1 : p < 0.35$$

where p is the proportion of households that can receive Channel C.

Step 2: Under H_0: $X \sim \text{B}(14, 0.35)$. Observed $x = 2$ (lower tail).

$$\text{P}(X \leq 2) = \text{P}(X = 0) + \text{P}(X = 1) + \text{P}(X = 2)$$

$$\text{P}(X = 0) = (0.65)^{14} = 0.00178\ldots$$
$$\text{P}(X = 1) = 14(0.35)(0.65)^{13} = 14 \times 0.35 \times 0.00274 = 0.01343\ldots$$
$$\text{P}(X = 2) = \binom{14}{2}(0.35)^2(0.65)^{12} = 91 \times 0.1225 \times 0.00421 = 0.04696\ldots$$

$$\text{P}(X \leq 2) = 0.00178 + 0.01343 + 0.04696 = 0.06217\ldots$$

Check $\text{P}(X \leq 1) = 0.00178 + 0.01343 = 0.01521$. Since $0.0152 < 0.025$, the critical region is $X \leq 1$.

The observed $x = 2$ does not lie in the critical region ($\text{P}(X \leq 2) = 0.0622 > 0.025$).

Step 3: Conclusion.

Do not reject H_0. There is insufficient evidence at the 2.5% significance level that the proportion of households that can receive Channel C is less than 0.35.

Part (ii): Find largest n such that 0 receiving does not lead to rejection of H_0

Same hypotheses, same significance level 2.5%, same $p = 0.35$. Observed $x = 0$.

We need $\text{P}(X = 0) > 0.025$:

$$\text{P}(X = 0) = (0.65)^n > 0.025$$

$$n \ln(0.65) > \ln(0.025) \Rightarrow n < \frac{\ln(0.025)}{\ln(0.65)} = \frac{-3.6889}{-0.4308} = 8.563\ldots$$

So largest integer n is $n = 8$. Check: $(0.65)^8 = 0.0319 > 0.025$ ✓; $(0.65)^9 = 0.0207 < 0.025$.

Largest value of n for which H_0 is not rejected is $n = \mathbf{8}$.

Question 4

Worked Solution

Part (i): Test at 10% whether proportion of letter e is less than 19%

Message of 20 letters: 1 letter e.

Step 1: Hypotheses.

$$H_0 : p = 0.19 \qquad H_1 : p < 0.19$$

where p is the proportion of letter e in the language.

Step 2: Under H_0: $X \sim \mathrm{B}(20, 0.19)$. Observed $x = 1$.

$$\mathrm{P}(X \leq 1) = \mathrm{P}(X = 0) + \mathrm{P}(X = 1)$$

$$\mathrm{P}(X = 0) = (0.81)^{20} = 0.01480\ldots$$
$$\mathrm{P}(X = 1) = 20(0.19)(0.81)^{19} = 20 \times 0.19 \times 0.01827 = 0.06943\ldots$$

$$\mathrm{P}(X \leq 1) = 0.01480 + 0.06943 = 0.08423\ldots$$

Step 3: Compare with 0.10: $0.0842 < 0.10$.

Critical region: $\mathrm{P}(X \leq 0) = 0.0148 < 0.10$ and $\mathrm{P}(X \leq 1) = 0.0842 < 0.10$, so the critical region is $X \leq 1$.

The observed $x = 1$ lies in the critical region.

Reject H_0. There is significant evidence at the 10% level that the proportion of letter e in the language from which the message is a sample is less than 19% (less than in German).

Part (ii): Why binomial might not be appropriate

The letters in the message are not independent — consecutive letters in a word/language are correlated (the occurrence of one letter depends on which letters surround it).

Question 5

Worked Solution

Agency claims at least 65% favour development. Sample of 12: 6 favour.

Part (i): Test at 10% whether result is consistent with agency's claim

Step 1: Hypotheses.

$$H_0 : p = 0.65 \qquad H_1 : p < 0.65$$

(Testing whether the true proportion is less than 65%, which would contradict the agency's claim.)

Step 2: Under H_0: $X \sim \mathrm{B}(12, 0.65)$. Observed $x = 6$ (lower tail).

$$\mathrm{P}(X \leq 6) = \sum_{k=0}^{6} \binom{12}{k} (0.65)^k (0.35)^{12-k}$$

From tables: $\mathrm{P}(X \leq 6) = 0.2127$.

Step 3: Compare with 0.10: $0.2127 > 0.10$.

Critical region: $\mathrm{P}(X \leq 5) = 0.0846 < 0.10$. So critical region is $X \leq 5$. Observed $x = 6$ does not lie in the critical region.

Do not reject H_0. There is insufficient evidence at the 10% level to conclude that the proportion in favour is less than 65%. The result is consistent with the agency's claim.

Part (ii): Residents' group claims no more than 35% are in favour

The residents claim $p \leq 0.35$. Since we failed to reject $H_0 : p = 0.65$ even at 10%, there is certainly no evidence to support $p \leq 0.35$ (which is even further from 0.65). The same conclusion applies.

Insufficient evidence to support the residents' claim: since we could not reject $p = 0.65$ against $p < 0.65$, there is even less evidence for $p \leq 0.35$.

Part (iii): Test at 15%, sample $2n$, exactly n in favour; find smallest n to reject agency's claim

Under $H_0 : p = 0.65$, $X \sim \mathrm{B}(2n, 0.65)$. Observed $x = n$ (half the sample). We reject if $\mathrm{P}(X \leq n) < 0.15$.

Test values:

$$\begin{aligned}
n = 8 &: \mathrm{P}(X \leq 8 \mid \mathrm{B}(16, 0.65)) = 0.0563 < 0.15 \quad \checkmark \text{ (reject)} \\
n = 7 &: \mathrm{P}(X \leq 7 \mid \mathrm{B}(14, 0.65)) = 0.0898 < 0.15 \quad \checkmark \\
n = 6 &: \mathrm{P}(X \leq 6 \mid \mathrm{B}(12, 0.65)) = 0.2127 > 0.15 \quad \text{(do not reject)}
\end{aligned}$$

Wait — we need the *smallest* n for which we reject. Checking systematically: as n increases, the sample provides more evidence against H_0 if $x/2n = 0.5 < 0.65$.

From the tables: smallest n where $P(X \le n \mid B(2n, 0.65)) < 0.15$ is $n = 7$ (giving $P = 0.0898$, consistent); checking smaller n: for $n = 6$, $P(X \le 6 \mid B(12, 0.65)) = 0.2127 > 0.15$, so do not reject.

Smallest $n =$ **7**: for $n = 7$, $P(X \le 7 \mid B(14, 0.65)) = 0.0898 < 0.15$, so reject H_0 at 15%.

Question 6

Worked Solution

Drug development: company continues if $p > 0.7$. Test on 14 patients at 5%.

$$H_0 : p = 0.7 \qquad H_1 : p > 0.7$$

$X \sim \mathrm{B}(14, 0.7)$ under H_0.

Part (i): Find critical region

We need the smallest c such that $\mathrm{P}(X \geq c) < 0.05$, i.e. $1 - \mathrm{P}(X \leq c - 1) < 0.05$.

$$\begin{aligned}\mathrm{P}(X \geq 13) &= \mathrm{P}(X = 13) + \mathrm{P}(X = 14) = 14(0.7)^{13}(0.3) + (0.7)^{14} \\ &= 14 \times 0.02542 \times 0.3 + 0.01178 = 0.10677 + 0.01178 = 0.04754\ldots < 0.05\end{aligned}$$

Check: $\mathrm{P}(X \geq 12) = 1 - \mathrm{P}(X \leq 11) = 1 - 0.9525 = 0.1608 > 0.05$.

So the critical region is $X \geq 13$, with probability 0.0475.

Critical region: $X \geq 13$, with probability 0.0475 in the tail.

Part (ii): 12 of 14 patients show improvement; carry out test

Observed $x = 12$. The critical region is $\{X \geq 13\}$.

Since $x = 12$ does not lie in the critical region:

Do not reject H_0. There is insufficient evidence at the 5% significance level that the proportion showing improvement is greater than 0.7.

Question 7

Worked Solution

City: 40% from ethnic group Z. Company sample of 12: only 2 from group Z. Test at 5% whether proportion of employees from group Z is less.

Part (i): Assumption for the test to be valid

The sample must be **random** (each employee equally likely to be selected).

Part (ii): Appropriate way to obtain the sample

List all employees and number them sequentially. Use random numbers to select 12 employees from the list.

Part (iii): Carry out the test

Step 1: Hypotheses.

$$H_0 : p = 0.4 \qquad H_1 : p < 0.4$$

where p is the proportion of company employees from ethnic group Z.

Step 2: Under H_0: $X \sim \mathrm{B}(12, 0.4)$. Observed $x = 2$.

$$\mathrm{P}(X \leq 2) = \mathrm{P}(X = 0) + \mathrm{P}(X = 1) + \mathrm{P}(X = 2)$$

$$\begin{aligned} \mathrm{P}(X = 0) &= (0.6)^{12} = 0.00217 \\ \mathrm{P}(X = 1) &= 12(0.4)(0.6)^{11} = 0.01741 \\ \mathrm{P}(X = 2) &= 66(0.4)^2(0.6)^{10} = 0.06385 \end{aligned}$$

$$\mathrm{P}(X \leq 2) = 0.08344\ldots$$

Step 3: Compare with 0.05: $0.0834 > 0.05$. Also check: $\mathrm{P}(X \leq 1) = 0.01958 < 0.05$.

Critical region: $X \leq 1$. Observed $x = 2$ does not lie in the critical region.

$\mathrm{P}(X \leq 2) = 0.0834 > 0.05$. Do not reject H_0. There is insufficient evidence at the 5% level that the proportion of company employees from ethnic group Z is less than in the city. [Critical region: $X \leq 1$, probability 0.0196]

Part (iv): Would a 10% test be more or less supportive of the manager's belief?

At 10%: critical region would be $X \leq 2$ (since $\mathrm{P}(X \leq 2) = 0.0834 < 0.10$). The observed $x = 2$ would then lie in the critical region and H_0 would be rejected.

More supportive: at 10%, the critical region is $X \leq 2$, and since $x = 2$ falls in the critical region, H_0 would be rejected — providing evidence to support the manager's belief.

Question 8

Worked Solution

55% of pupils are girls. Claim: $p > 0.55$ for Head Student. Evidence: 6 of last 8 Head Students are girls.

Part (i): Test at 10%

Step 1: Hypotheses.

$$H_0 : p = 0.55 \qquad H_1 : p > 0.55$$

where p is the probability a Head Student is a girl.

Step 2: Under H_0: $R \sim \mathrm{B}(8, 0.55)$. Observed $r = 6$.

$$\mathrm{P}(R \geq 6) = 1 - \mathrm{P}(R \leq 5) = 1 - 0.7799 = 0.2201$$

Step 3: Compare with 0.10: $0.2201 > 0.10$.

Critical region: $\mathrm{P}(R \geq 7) = 1 - \mathrm{P}(R \leq 6) = 1 - 0.9368 = 0.0632 < 0.10$. So critical region is $R \geq 7$.

The observed $r = 6$ does not lie in the critical region.

Do not reject H_0. There is insufficient evidence at the 10% level that the probability a girl becomes Head Student is greater than 0.55. The claim is not supported.

[Critical region: $\{R \geq 7\}$; probability in tail: 0.0632]

Part (ii): Assumption needed for the test to be valid

The last 8 years must be a **random sample** of years when a Head Student was chosen (i.e. the outcomes are independent and representative).

Question 9

Worked Solution

$X \sim \mathrm{B}(20, p)$. Test $H_0 : p = 0.3$ vs $H_1 : p \neq 0.3$ at 5%.

Part (a): Critical region (each tail as close as possible to 2.5%)

Under H_0: $X \sim \mathrm{B}(20, 0.3)$.

Lower tail: Find largest c_1 with $\mathrm{P}(X \leq c_1) \leq 0.025$:

$$\mathrm{P}(X \leq 2) = 0.0355 > 0.025$$
$$\mathrm{P}(X \leq 1) = 0.0076 < 0.025 \quad \checkmark$$

Wait — we want closest to 0.025: $\mathrm{P}(X \leq 2) = 0.0355$ and $\mathrm{P}(X \leq 1) = 0.0076$. Closest to 0.025 is $X \leq 2$ (0.0355 vs 0.0076). But the instruction is that probability in tail should be *as close as possible* to 0.025. Compare $|0.0355 - 0.025| = 0.0105$ vs $|0.0076 - 0.025| = 0.0174$: so $X \leq 2$ is closer.

Lower critical region: $X \leq 2$ with probability 0.0355.

Upper tail: Find smallest c_2 with $\mathrm{P}(X \geq c_2) \leq 0.025$:

$$\mathrm{P}(X \geq 11) = 1 - \mathrm{P}(X \leq 10) = 1 - 0.9829 = 0.0171 < 0.025 \quad \checkmark$$
$$\mathrm{P}(X \geq 10) = 1 - \mathrm{P}(X \leq 9) = 1 - 0.9520 = 0.0480 > 0.025$$

Closest to 0.025: $|0.0171 - 0.025| = 0.0079$ vs $|0.0480 - 0.025| = 0.023$. So $X \geq 11$ is closer.

Upper critical region: $X \geq 11$ with probability 0.0171.

Critical region: $\{X \leq 2\} \cup \{X \geq 11\}$
Lower tail probability: 0.0355; Upper tail probability: 0.0171

Part (b): Actual significance level

$$\text{Actual significance level} = 0.0355 + 0.0171 = 0.0526 \approx 5.26\%$$

Actual significance level = 5.26% (or 0.0526)

Part (c): Observed $x = 3$; conclusion

$x = 3$ does not lie in the critical region ($X \leq 2$ or $X \geq 11$).

$x = 3$ is not in the critical region, so do not reject H_0. There is insufficient evidence that $p \neq 0.3$.

Question 10

Worked Solution

Part (a): $X \sim \mathrm{B}(40, 0.27)$, **find** $\mathrm{P}(X \geq 16)$

$$\mathrm{P}(X \geq 16) = 1 - \mathrm{P}(X \leq 15) = 1 - 0.949077\ldots = 0.0509\ldots$$

$\mathrm{P}(X \geq 16) = 0.0509$ (awrt 0.0509)

Part (b): State hypotheses

$$H_0 : p = 0.3 \qquad H_1 : p \neq 0.3$$

where p is the proportion of customers who buy baked beans in single tins.

Part (c): Critical region for two-tailed test at 10%, each tail < 0.05

Under H_0: $Y \sim \mathrm{B}(20, 0.3)$.

Lower tail:

$$\mathrm{P}(Y \leq 2) = 0.0355 < 0.05 \quad \checkmark$$
$$\mathrm{P}(Y \leq 3) = 0.1071 > 0.05$$

Lower critical region: $Y \leq 2$.

Upper tail:

$$\mathrm{P}(Y \geq 10) = 1 - \mathrm{P}(Y \leq 9) = 1 - 0.9520 = 0.0480 < 0.05 \quad \checkmark$$
$$\mathrm{P}(Y \geq 9) = 1 - \mathrm{P}(Y \leq 8) = 1 - 0.8867 = 0.1133 > 0.05$$

Upper critical region: $Y \geq 10$.

Critical region: $\{Y \leq 2\} \cup \{Y \geq 10\}$
Lower tail probability: 0.0355; Upper tail probability: 0.0480

Part (d): Actual significance level

$$\text{Actual significance level} = 0.0355 + 0.0480 = 0.0835 \approx 8.35\%$$

Actual significance level = 8.35%

Part (e): Manager observes 12 of 20 buying single tins

$y = 12 \geq 10$, so $y = 12$ lies in the upper critical region.

12 lies in the critical region ($Y \geq 10$). Reject H_0. The manager's suspicion is supported — there is evidence of a change in the proportion buying single tins.

Part (f): Local scout group visited — validity of model

The model is not valid: the 20 customers do not form a random sample since the scout group (buying in bulk for camping) are not representative of typical customers. The independence assumption is violated.

Question 11

Worked Solution

Part (a): Critical region for two-tailed test at 10%, each tail < 0.05

Under $H_0 : p = 0.3$, $X \sim \mathrm{B}(20, 0.3)$ (same as Q10 part (c)).

Lower tail: $\mathrm{P}(X \leq 2) = 0.0355 < 0.05$. Lower critical region: $X \leq 2$.

Upper tail: $\mathrm{P}(X \geq 10) = 0.0480 < 0.05$. Upper critical region: $X \geq 10$.

Critical region: $\{X \leq 2\} \cup \{X \geq 10\}$
Lower tail: $\mathrm{P}(X \leq 2) = 0.0355$; Upper tail: $\mathrm{P}(X \geq 10) = 0.0480$

Part (b): Actual significance level

$$0.0355 + 0.0480 = 0.0835 \approx 8.35\%$$

Actual significance level = 8.35% (awrt 0.083 or 0.084)

Part (c): Manager finds 11 customers bought single tins; comment

$x = 11 \geq 10$, so $x = 11$ lies in the critical region.

11 is in the critical region ($X \geq 10$). Reject H_0. There is evidence of a change (increase) in the proportion of customers buying baked beans in single tins.

Question 12

Worked Solution

Part (a): Define the critical region of a test statistic

The critical region is the set of values of the test statistic for which the null hypothesis is **rejected** in a hypothesis test.

Part (b): $X \sim \mathrm{B}(30, 0.3)$**, test** $H_0 : p = 0.3$ **vs** $H_1 : p \neq 0.3$ **at 1%, each tail as close to 0.5% as possible**

Under H_0: $X \sim \mathrm{B}(30, 0.3)$.

Lower tail:

$$\mathrm{P}(X \leq 3) = 0.0093 > 0.005$$
$$\mathrm{P}(X \leq 2) = 0.0021 < 0.005$$

Closest to 0.005: $|0.0093 - 0.005| = 0.0043$ vs $|0.0021 - 0.005| = 0.0029$. $X \leq 2$ is closer. Lower critical region: $X \leq 2$, probability 0.0021.

Upper tail:

$$\mathrm{P}(X \geq 16) = 1 - \mathrm{P}(X \leq 15) = 1 - 0.9936 = 0.0064 > 0.005$$
$$\mathrm{P}(X \geq 17) = 1 - \mathrm{P}(X \leq 16) = 1 - 0.9979 = 0.0021 < 0.005$$

Closest to 0.005: $|0.0064 - 0.005| = 0.0014$ vs $|0.0021 - 0.005| = 0.0029$. $X \geq 16$ is closer. Upper critical region: $X \geq 16$, probability 0.0064.

Critical region: $\{X \leq 2\} \cup \{X \geq 16\}$
Lower tail: probability 0.0021; Upper tail: probability 0.0064

Part (c): Actual significance level

$$0.0021 + 0.0064 = 0.0085 \approx 0.85\%$$

Actual significance level $= 0.85\%$

Part (d): Observed $x = 15$**; comment**

$x = 15$: the critical region is $\{X \leq 2\} \cup \{X \geq 16\}$. Since 15 does not lie in either part of the critical region:

$x = 15$ does not lie in the critical region. Do not reject H_0. There is no significant evidence of a change in $p = 0.3$ at the 1% level.

End of Worked Solutions