Question 1

Worked Solution

Part (i)(A): $\mathrm{P}(X = 3)$ **where** $X \sim \mathrm{B}(16, 0.1)$

$$\mathrm{P}(X = 3) = \binom{16}{3}(0.1)^3(0.9)^{13} = 560 \times 0.001 \times 0.25420 = 0.1423$$

Alternatively, from tables: $\mathrm{P}(X \leq 3) - \mathrm{P}(X \leq 2) = 0.9316 - 0.7892 = 0.1424$.

> $\mathrm{P}(X = 3) = 0.142 \quad (3 \text{ s.f.})$

Part (i)(B): $\mathrm{P}(X \geq 3)$

$$\mathrm{P}(X \geq 3) = 1 - \mathrm{P}(X \leq 2) = 1 - 0.7892 = 0.2108$$

> $\mathrm{P}(X \geq 3) = 0.211 \quad (3 \text{ s.f.})$

Part (i)(C): Expected number using PIN 1234

$$\mathrm{E}(X) = np = 16 \times 0.1 = 1.6$$

> Expected number $= 1.6$

Part (ii): Null and alternative hypotheses

Let $p =$ the probability that a randomly selected cash machine user uses the PIN 1234.

$$H_0 : p = 0.1 \qquad H_1 : p < 0.1$$

The alternative hypothesis has this form because the advertising campaign aims to **reduce** the proportion of users who use PIN 1234, so we are testing for a decrease.

> $H_0 : p = 0.1, \quad H_1 : p < 0.1$
> H_1 has this form because the advertising campaign aims to reduce the proportion of people using 1234 as their PIN.

Part (iii)(A): Why the test cannot be carried out at the 10% significance level with $n = 20$

Under H_0: $X \sim \mathrm{B}(20, 0.1)$. For the test to be possible at 10%, we need the critical region to be non-empty — there must exist a value c such that $\mathrm{P}(X \leq c) < 0.10$.

$$\mathrm{P}(X \leq 0) = (0.9)^{20} = 0.1216 > 0.10$$

Since even the most extreme possible outcome ($X = 0$, nobody using PIN 1234) gives a probability of 0.1216, which exceeds 0.10, there is no critical region. The null hypothesis cannot be rejected at the 10% level whatever the observed value.

$P(X = 0) = (0.9)^{20} = 0.1216 > 0.10$, so even the most extreme result (no one using 1234) does not fall in the critical region. There is no critical region — H_0 cannot be rejected at the 10% significance level for any observed value.

Part (iii)(B): Lowest integer value of k for which H_0 could be rejected

We need $P(X = 0) < k/100$, i.e. $0.1216 < k/100$, so $k > 12.16$.

The lowest integer value is $k = 13$.

Lowest integer value of $k = \mathbf{13}$ (i.e. the 13% significance level)

Part (iv): New sample of 60, $X \sim B(60, 0.1)$, observed $x = 2$; test at 5%

Using the same hypotheses: $H_0 : p = 0.1$, $H_1 : p < 0.1$.

Under H_0: $X \sim B(60, 0.1)$. The relevant probability is the lower tail:

We are given $P(X \leq 2) = 0.0530$.

Compare with 5%: $0.0530 > 0.05$.

So the observed value $x = 2$ does not lie in the critical region ($P(X \leq 1) = 0.0138 < 0.05$, so the critical region is $X \leq 1$, and $x = 2$ is not in it).

$P(X \leq 2) = 0.0530 > 0.05$. Do not reject H_0.

There is insufficient evidence at the 5% significance level to suggest that the advertising campaign has been successful in reducing the proportion using PIN 1234.

Question 2

Worked Solution

$X \sim \mathrm{B}(20, 0.78)$: number of patients cured by the existing drug.

Part (i)(A): P(**exactly 19 cured**)

$$\mathrm{P}(X = 19) = \binom{20}{19}(0.78)^{19}(0.22)^{1} = 20 \times (0.78)^{19} \times 0.22$$
$$= 20 \times 0.17716\ldots \times 0.22 = 0.7795\ldots \times 0.22 = 0.03915\ldots \times 4 \approx 0.0392$$

$\mathrm{P}(X = 19) = 0.0392 \quad$ (3 s.f.)

Part (i)(B): P(**at most 18 cured**)

$\mathrm{P}(X \leq 18) = 1 - \mathrm{P}(X = 19) - \mathrm{P}(X = 20) = 1 - (0.0392 + 0.0069) = 1 - 0.0461 = 0.9539\ldots$

where $\mathrm{P}(X = 20) = (0.78)^{20} = 0.0069$.

$\mathrm{P}(X \leq 18) = 0.954 \quad$ (3 s.f.)

Part (i)(C): Expected number cured

$$\mathrm{E}(X) = np = 20 \times 0.78 = 15.6$$

$\mathrm{E}(X) = 15.6$

Part (ii): Hypothesis test at 1% — new drug, 19 of 20 cured

Let p = probability a patient is cured by the new drug.

Step 1: Hypotheses.

$$H_0 : p = 0.78 \qquad H_1 : p > 0.78$$

($H_1 : p > 0.78$ because the researchers hope the new drug is *more* effective.)

Step 2: Under H_0: $X \sim \mathrm{B}(20, 0.78)$. Observed $x = 19$ (upper tail).

$$\mathrm{P}(X \geq 19) = \mathrm{P}(X = 19) + \mathrm{P}(X = 20) = 0.0392 + 0.0069 = 0.0461$$

Step 3: Compare with 1%: $0.0461 > 0.01$.

$\mathrm{P}(X \geq 19) = 0.0461 > 0.01$. Do not reject H_0.
There is insufficient evidence at the 1% significance level to suggest that the new drug is more effective than the old one.

Part (iii): Would the conclusion differ at the 5% significance level?

$P(X \geq 19) = 0.0461 < 0.05$.

At the 5% significance level, $0.0461 < 0.05$, so we **would reject** H_0.

There would be sufficient evidence to suggest the new drug is more effective than the old one.

Question 3

Worked Solution

$X \sim \mathrm{B}(15, 0.85)$: number of seeds (out of 15) that germinate.

Part (i)(A): P(**exactly 12 germinate**)

$$P(X = 12) = \binom{15}{12}(0.85)^{12}(0.15)^3 = 455 \times (0.85)^{12} \times (0.15)^3$$
$$= 455 \times 0.14233 \times 0.003375 = 0.2184$$

Alternatively, from tables: $P(X \leq 12) - P(X \leq 11) = 0.3958 - 0.1773 = 0.2185$.

$P(X = 12) = 0.218$ (3 s.f.)

Part (i)(B): P(**fewer than 12 germinate**)

$$P(X < 12) = P(X \leq 11) = 0.1773$$

$P(X < 12) = 0.177$ (3 s.f.)

Part (ii): Hypotheses for test at 1%

Let p = probability that a one-year-old seed of this variety germinates.

$$H_0 : p = 0.85 \qquad H_1 : p < 0.85$$

H_1 has this form because Ramesh suspects the germination rate is **lower** for year-old seeds.

$H_0 : p = 0.85$, $H_1 : p < 0.85$
$H_1 : p < 0.85$ because the question investigates whether the germination rate has decreased.

Part (iii): $n = 20$**, 13 seeds germinate; carry out test at 1%**

Under H_0: $X \sim \mathrm{B}(20, 0.85)$. Observed $x = 13$ (lower tail).

$$P(X \leq 13) = 0.0219$$

Compare with 1%: $0.0219 > 0.01$.

$P(X \leq 13) = 0.0219 > 0.01$. Do not reject H_0.
There is insufficient evidence at the 1% significance level to conclude that the proportion of seeds germinating has decreased.

Part (iv): $n = 50$**, 33 germinate; critical value given as 35; complete the test at 1%**

Under H_0: $X \sim \text{B}(50, 0.85)$. Observed $x = 33$.

The critical value is 35, meaning the critical region is $X \leq 35$ — but we are told the critical value *for the test* is 35, so the critical region is $\{X \leq 35\}$.

Since $33 < 35$, the observed value **lies in the critical region**.

$33 < 35$, so $x = 33$ lies in the critical region. Reject H_0.

There is sufficient evidence at the 1% significance level to conclude that the proportion of seeds germinating has decreased.

Part (v): Find least n for which the critical region is non-empty at 1%

For the critical region to be non-empty, we need $\text{P}(X = 0) < 0.01$, where $X \sim \text{B}(n, 0.85)$:

$$(0.15)^n < 0.01$$

Test values:

$$n = 2: \quad (0.15)^2 = 0.0225 > 0.01$$
$$n = 3: \quad (0.15)^3 = 0.003375 < 0.01 \quad \checkmark$$

Alternative method using logarithms:

$$n \ln(0.15) < \ln(0.01) \implies n > \frac{\ln(0.01)}{\ln(0.15)} = \frac{-4.6052}{-1.8971} = 2.427\ldots$$

So least $n = 3$.

For $n = 2$: $\text{P}(X = 0) = (0.15)^2 = 0.0225 > 0.01$ — critical region empty.
For $n = 3$: $\text{P}(X = 0) = (0.15)^3 = 0.00338 < 0.01$ — critical region non-empty. ✓

Least value of $n =$ **3**.

Question 4

Worked Solution

$X \sim \mathrm{B}(10, 0.35)$: number of customers (out of 10) accessing the internet.

Part (i)(A): P(**exactly 5 accessing internet**)

$$\mathrm{P}(X = 5) = \binom{10}{5}(0.35)^5(0.65)^5 = 252 \times 0.005252 \times 0.116029 = 0.1536$$

Alternatively, from tables: $\mathrm{P}(X \le 5) - \mathrm{P}(X \le 4) = 0.9051 - 0.7515 = 0.1536$.

$\mathrm{P}(X = 5) = 0.154$ (3 s.f.)

Part (i)(B): $\mathrm{P}(X \ge 5)$

$$\mathrm{P}(X \ge 5) = 1 - \mathrm{P}(X \le 4) = 1 - 0.7515 = 0.2485$$

$\mathrm{P}(X \ge 5) = 0.249$ (3 s.f.)

Part (i)(C): Expected number accessing internet

$$\mathrm{E}(X) = 10 \times 0.35 = 3.5$$

$\mathrm{E}(X) = 3.5$

Part (ii): Two-tailed test at 5% — second coffee shop, $n = 20$, $x = 10$

Let p = probability a randomly selected customer at the second coffee shop is accessing the internet.

Step 1: Hypotheses.

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

$H_1 : p \ne 0.35$ because we are investigating whether the probability is *different* (could be higher or lower), not specifically in one direction.

Step 2: Under H_0: $X \sim \mathrm{B}(20, 0.35)$. Observed $x = 10$ (upper tail, since $10 > np = 7$).

$$\mathrm{P}(X \ge 10) = 1 - \mathrm{P}(X \le 9) = 1 - 0.8782 = 0.1218$$

Step 3: For a two-tailed test at 5%, compare with 2.5%: $0.1218 > 0.025$.

$\mathrm{P}(X \ge 10) = 0.1218 > 0.025$. Do not reject H_0.

There is insufficient evidence at the 5% significance level to conclude that the probability of a customer at this coffee shop accessing the internet is different from 0.35.

Part (iii): Larger sample $n = 200$, 90 accessing internet; given $\mathrm{P}(X \geq 90) = 0.0022$

Using same hypotheses: $H_0 : p = 0.35$, $H_1 : p \neq 0.35$ at 5%.

Observed $x = 90$ (upper tail).

$\mathrm{P}(X \geq 90) = 0.0022 < 0.025$.

$0.0022 < 0.025$. Reject H_0.

There is sufficient evidence at the 5% significance level to conclude that the probability of accessing the internet at this coffee shop is different from 0.35.

Question 5

Worked Solution

20 people selected at random; 13 make correct identification.

Part (i): Why null hypothesis should be $p = 0.5$

If people cannot identify bottled vs tap water, they are effectively guessing between two equally likely outcomes. By chance alone, the probability of a correct identification is $\frac{1}{2} = 0.5$, since there are two equally likely outcomes.

> If people are just guessing, each identification is like a 50:50 choice between two options. The probability of a correct guess by chance is 0.5, so $H_0 : p = 0.5$.

Part (ii): Why $H_1 : p > 0.5$

> $H_1 : p > 0.5$ because the researcher suspects people may do **better** than guessing — i.e. they may be able to correctly identify the water more than half the time.

Part (iii): Complete the test at 5%

Step 1: Hypotheses.

$$H_0 : p = 0.5 \qquad H_1 : p > 0.5$$

Step 2: Under H_0: $X \sim \mathrm{B}(20, 0.5)$. Observed $x = 13$.

$$\mathrm{P}(X \geq 13) = 1 - \mathrm{P}(X \leq 12) = 1 - 0.8684 = 0.1316$$

Step 3: Compare with 5%: $0.1316 > 0.05$.

> $\mathrm{P}(X \geq 13) = 0.1316 > 0.05$. Do not reject H_0.
>
> There is insufficient evidence at the 5% significance level to suggest that people can make a correct identification at better than chance.

Question 6

Worked Solution

Existing process: 5% of frames are faulty. New (cheaper) process: sample of 18, 4 found faulty. Test at 5% whether proportion has increased.

Step 1: Define p and state hypotheses.

Let p = probability that a randomly selected frame from the new process is faulty.

$$H_0 : p = 0.05 \qquad H_1 : p > 0.05$$

Step 2: Under H_0: $X \sim \mathrm{B}(18, 0.05)$. Observed $x = 4$ (upper tail).

$$\mathrm{P}(X \geq 4) = 1 - \mathrm{P}(X \leq 3) = 1 - 0.9891 = 0.0109$$

Step 3: Compare with 5%: $0.0109 < 0.05$.

$\mathrm{P}(X \geq 4) = 0.0109 < 0.05$. Reject H_0.

There is sufficient evidence at the 5% significance level to conclude that the proportion of faulty bicycle frames has increased.

End of Worked Solutions