Hypothesis testing Mixed exercise 4

1 **a** Let the random variable X denote the number of vehicles passing the point in a 10-minute period.

Use the Poisson distribution model $X \sim \text{Po}\left(\frac{39}{6}\right)$, i.e. $X \sim \text{Po}(6.5)$

$\text{P}(X = 6) = \frac{e^{-6.5}\, 6.5^6}{6!} = 0.1575$ (4 d.p.)

b Using the tables

$\text{P}(X \geqslant 8) = 1 - \text{P}(X \leqslant 7) = 1 - 0.6728 = 0.3272$

c $\text{H}_0: \lambda = 6.5 \quad \text{H}_1: \lambda < 6.5$

Assume H_0, so that $X \sim \text{Po}(6.5)$

Significance level 5%

From the tables $\text{P}(X \leqslant 2) = 0.0430$

$0.0430 < 0.05$

There is sufficient evidence at the 5% level to reject H_0, and conclude that the number of vehicles passing the point in a given period has decreased.

2 Let the random variable X denote the deformed red blood cells found in a 2.5 ml sample of Francesca's blood.

$\text{H}_0: \lambda = 3.2 \times 2.5 = 8 \quad \text{H}_1: \lambda < 8$

Assume H_0, so that $X \sim \text{Po}(8)$

Significance level 5%

From the tables $\text{P}(X \leqslant 4) = 0.0996$

$0.0996 > 0.05$

There is insufficient evidence at the 5% level to reject H_0 and to suggest that the mean number of deformed red blood cells has decreased.

3 Let the random variable X denote the number of days it takes Peter to complete the first crossword.

$\text{H}_0: p = 0.2 \quad \text{H}_1: p < 0.2$

Assume H_0, so that $X \sim \text{Geo}(0.2)$

Significance level 5%

$\text{P}(X \geqslant 7) = (1 - 0.2)^6 = 0.2621$ (4 d.p.)

$0.2621 > 0.05$

There is insufficient evidence at the 5% level to reject H_0 and to suggest that the crosswords are more difficult.

4 **a** Let the random variable X denote the number of days until Roisin sees the first fall of snow, so $X \sim \text{Geo}(0.45)$

$\text{P}(X \geqslant 3) = (1 - 0.45)^2 = 0.3025$

4 b $H_0 : p = 0.45 \quad H_1 : p < 0.45$

Assume H_0, so that $X \sim \text{Geo}(0.45)$

Significance level 5%

$P(X \geqslant 7) = (1 - 0.45)^6 = 0.0277$ (4 d.p.)

$0.0277 < 0.05$

There is sufficient evidence at the 5% level to reject H_0, and conclude that the meteorologist is correct.

5 a Let the random variable X denote the number of calls that Scoobie puts through to the wrong extension during a day in which he receives 150 calls, so $X \sim B(150, 0.03)$. Using a Poisson approximation $X \sim \text{Po}(150 \times 0.03)$, i.e. $X \sim \text{Po}(4.5)$

$$P(X = 5) = \frac{e^{-4.5}4.5^5}{5!} = 0.1708 \text{ (4 d.p.)}$$

b From the tables $P(X \leqslant 3) = 0.3423$

c Let the random variable Y denote the number of calls that Waldo puts through to the wrong extension during a day in which he receives 300 calls.

$H_0 : \lambda = 2 \times 4.5 = 9 \quad H_1 : \lambda < 9$

Assume H_0, so that $X \sim \text{Po}(9)$

Significance level 5%

From the tables $P(X \leqslant 4) = 0.0550$

$0.0550 > 0.05$

There is insufficient evidence at the 5% level to reject H_0 and there is therefore no evidence to suggest that the Waldo has decreased the rate at which calls are put through to the wrong extension.

6 a Let the random variable X denote the number of breakdowns in a one-month period, so $X \sim \text{Po}(1.75)$.

$$P(X = 3) = \frac{e^{-1.75}1.75^3}{3!} = 0.1552 \text{ (4 d.p.)}$$

b Let the random variable Y denote the number of breakdowns in a two-month period, so $Y \sim \text{Po}(3.5)$.

Using the tables

$P(Y \geqslant 6) = 1 - P(Y \leqslant 5) = 1 - 0.8576 = 0.1424$

c Let the random variable Z denote the number of months in a four-month period in which there are exactly 3 breakdowns. Use the probability from part **a** to 5 d.p., so $Z \sim B(4, 0.15522)$.

$$P(Z = 2) = \binom{4}{2} \times (0.15522)^2 \times (1 - 0.15522)^2 = 0.1032 \text{ (4 d.p.)}$$

6 d Let the random variable M denote the number of breakdowns in a four-month period, so $M \sim \text{Po}(7)$.

$H_0 : \lambda = 7 \quad H_1 : \lambda < 7$

Assume H_0, so that $M \sim \text{Po}(7)$

Significance level 5%, so require $P(X \leqslant c) < 0.05$

From the tables $P(M \leqslant 2) = 0.0296$ and $P(M \leqslant 3) = 0.0818$

$P(M \leqslant 2) < 0.05$ and $P(M \leqslant 3) > 0.05$ so the critical value is 2

Hence the critical region is $M \leqslant 2$

6 e Actual significance level $= \text{P}(M \leqslant 2) = 0.0296$

7 Let the random variable X denote the number of televisions sold in a two-day period.

$\text{H}_0: \lambda = 2\times3.5 = 7 \quad \text{H}_1: \lambda > 7$

Assume H_0, so that $X \sim \text{Po}(7)$

Significance level 5%

From the tables $\text{P}(X \geqslant 11) = 1 - \text{P}(X \leqslant 10) = 1 - 0.9015 = 0.0985$

$0.0985 > 0.05$

There is insufficient evidence at the 5% level to reject H_0 and therefore no evidence to suggest that advert increased the sales.

8 a Let the random variable X denote the rate of visits to the website on a Saturday.

$\text{H}_0: \lambda = 8.5 \quad \text{H}_1: \lambda > 8.5$

Assume H_0, so that $X \sim \text{Po}(8.5)$

Significance level 5%

From the tables $\text{P}(X \geqslant 12) = 1 - \text{P}(X \leqslant 11) = 1 - 0.8487 = 0.1513$

$0.1513 > 0.05$

There is insufficient evidence at the 5% level to reject H_0 and therefore no evidence to suggest that the rate of visits is greater on a Saturday.

b Requires finding the smallest positive integer c such that $\text{P}(X \geqslant c) < 0.05$

From the tables $\text{P}(X \geqslant 14) = 1 - \text{P}(X \leqslant 13) = 1 - 0.9486 = 0.0514$

and $\text{P}(X \geqslant 15) = 1 - \text{P}(X \leqslant 14) = 1 - 0.9726 = 0.0274$

$\text{P}(X \geqslant 14) > 0.05$ and $\text{P}(X \geqslant 15) < 0.05$, so $c = 15$

9 Let the random variable X denote the number of workers that are absent for at least one day in the last month, so $X \sim \text{B}(200, 0.05)$. Using a Poisson approximation $X \sim \text{Po}(200\times0.05)$, i.e. $X \sim \text{Po}(10)$

$\text{H}_0: \lambda = 10 \quad \text{H}_1: \lambda > 10$

Assume H_0, so that $X \sim \text{Po}(10)$

Significance level 5%

From the tables $\text{P}(X \geqslant 15) = 1 - \text{P}(X \leqslant 14) = 1 - 0.9165 = 0.0835$

$0.0835 > 0.05$

There is insufficient evidence at the 5% level to reject H_0 and no evidence to suggest that the percentage is higher than the manager thinks.

10 a Let the random variable X denote the number of products tested until the first defective one is found, so $X \sim \text{Geo}(0.15)$

$\text{P}(X = 5) = 0.15(1-0.15)^4 = 0.0783$ (4 d.p.)

b $\text{P}(X \geqslant 3) = (1-0.15)^2 = 0.7225$

10 c $H_0: p = 0.15 \quad H_1: p < 0.15$

Assume H_0, so that $X \sim \text{Geo}(0.15)$

Significance level 5%

Require $P(X \geqslant c) < 0.05$

So $(1-0.15)^{c-1} < 0.05$

$(c-1)\log 0.85 < \log 0.05$

$c-1 > \dfrac{\log 0.05}{\log 0.85}$

$c > 19.433$

So the critical region is $X \geqslant 20$

d Actual significance level $= P(X \geqslant 20) = (1-0.15)^{19} = 0.0456$ (4 d.p.)

11 a Let the random variable X denote the number of hurricanes in the area in August. Then use a Poisson model and test $H_0: \lambda = 4 \quad H_1: \lambda > 4$

b For the null hypothesis $(H_0: \lambda = 4)$ to be rejected at 5% level of significance find the smallest positive integer c such that $P(X \geqslant c) < 0.05$

Assume H_0, so that $X \sim \text{Po}(4)$

Significance level 5%, so require $P(X \geqslant c) < 0.05$

From the tables $P(X \geqslant 8) = 1 - P(X \leqslant 7) = 1 - 0.9489 = 0.0511$

and $P(X \geqslant 9) = 1 - P(X \leqslant 8) = 1 - 0.9786 = 0.0214$

$P(X \geqslant 8) > 0.05$ and $P(X \geqslant 9) < 0.05$ so the critical value is 9

The critical region is $X \geqslant 9$ and the number of hurricanes must increase to 9 for H_0 to be rejected.

c As $X = 8$ is not in the critical region there is insufficient evidence to reject H_0 and therefore the scientist's suggestion should be rejected.

12 a Let the random variable X denote the number of heads recorded in 30 spins of the coin, so $X \sim B(30, p)$. As p is not known to be small, the Poisson approximation cannot be used.

$H_0: p = 0.5 \quad H_1: p < 0.5$

Assume H_0, so that $X \sim B(30, 0.5)$

Significance level 2%, so require $P(X \leqslant c) < 0.02$

From the binomial cumulative distribution tables $P(X \leqslant 8) = 0.0081$ and $P(X \leqslant 9) = 0.0214$

$P(X \leqslant 8) < 0.02$ and $P(X \leqslant 9) > 0.02$ so the critical value is 8

Hence the critical region is $X \leqslant 8$

12 b Let the random variable Y denote the number of coin spins until it lands on heads for the first time, so $X \sim \text{Geo}(p)$

$H_0 : p = 0.5 \quad H_1 : p < 0.5$

Assume H_0, so that $Y \sim \text{Geo}(0.5)$

Significance level 2%

Require $P(Y \geqslant c) < 0.02$

So $(1-0.5)^{c-1} < 0.02$

$$c - 1 > \frac{\log 0.02}{\log 0.5}$$

$c > 6.644$

So the critical region is $Y \geqslant 7$

c The probability that Alison has incorrectly rejected H_0 is $P(X \leqslant 8) = 0.0081$

The probability that Paul has incorrectly rejected H_0 is $P(Y \geqslant 7) = 0.5^6 = 0.0156$

Challenge

a Let N denote the number of wells sunk until the oil company sinks before it strike oil for the third time in the new region, so $N \sim \text{Negative B}(3, p)$

This assumes that the probability of striking oil remains the same for each well drilled.

b $H_0 : p = 0.18 \quad H_1 : p > 0.18$

Assume H_0, so that $N \sim \text{Negative B}(3, 0.18)$

Significance level 5%

Require $P(N \leqslant c) < 0.05$

Need to calculate cumulative probability distributions

$$P(N \leqslant 3) = P(N = 3) = \binom{2}{2} 0.18^3 (1-0.18)^0 = 0.00583...$$

$$P(N \leqslant 4) = P(N \leqslant 3) + P(N = 4) = 0.00583 + \binom{3}{2} 0.18^3 (1-0.18)^1 = 0.00583 + 0.01435 = 0.02018$$

$$P(N \leqslant 5) = P(N \leqslant 4) + P(N = 5) = 0.02018 + \binom{4}{2} 0.18^3 (1-0.18)^2 = 0.02018 + 0.02353 = 0.04371$$

$$P(N \leqslant 6) = P(N \leqslant 5) + P(N = 6) = 0.04371 + \binom{5}{2} 0.18^3 (1-0.18)^3 = 0.04371 + 0.03216 = 0.07587$$

As $P(N \leqslant 5) < 0.05$ and $P(N \leqslant 6) > 0.05$, the critical region is $N \geqslant 5$

c Actual significance level $= P(N \leqslant 5) = 0.0437$ (4 d.p.)