# Quality of Tests

## Questions

**Q1.**

Information was collected about accidents on the *Seapron* bypass. It was found that the number of accidents per month could be modelled by a Poisson distribution with mean 2.5

Following some work on the bypass, the numbers of accidents during a series of 3-month periods were recorded. The data were used to test whether or not there was a change in the mean number of accidents per month.

(a) Stating your hypotheses clearly and using a 5% level of significance, find the critical region for this test. You should state the probability in each tail.

**(5)**

(b) State P(Type I error) using this test.

**(1)**

Data from the series of 3-month periods are recorded for 2 years.

(c) Find the probability that at least 2 of these 3-month periods give a significant result.

**(3)**

Given that the number of accidents per month on the bypass, after the work is completed, is actually 2.1 per month,

(d) find P(Type II error) for the test in part (a)

**(3)**

**(Total for question = 12 marks)**

**Q2.**

Sam and Tessa are testing a spinner to see if the probability, $p$, of it landing on red is less than $\frac{1}{5}$. They both use a 10% significance level.

Sam decides to spin the spinner 20 times and record the number of times it lands on red.

(a) Find the critical region for Sam's test.

**(2)**

(b) Write down the size of Sam's test.

**(1)**

Tessa decides to spin the spinner until it lands on red and she records the number of spins.

(c) Find the critical region for Tessa's test.

**(6)**

(d) Find the size of Tessa's test.

**(1)**

(e) (i) Show that the power function for Sam's test is given by

$$(1 - p)^{19}(1 + 19p)$$

(ii) Find the power function for Tessa's test.

**(4)**

(f) With reference to parts (b), (d) and (e), state, giving your reasons, whether you would recommend Sam's test or Tessa's test when $p = 0.15$

**(4)**

**(Total for question = 18 marks)**

**Q3.**

The number of customers entering Jeff's supermarket each morning follows a Poisson distribution.

Past information shows that customers enter at an average rate of 2 every 5 minutes.

Using this information,

(a) (i) find the probability that exactly 26 customers enter Jeff's supermarket during a randomly selected 1-hour period one morning,

**(2)**

(ii) find the probability that at least 21 customers enter Jeff's supermarket during a randomly selected 1-hour period one morning.

**(2)**

A rival supermarket is opened nearby. Following its opening, the number of customers entering Jeff's supermarket over a randomly selected 40-minute period is found to be 10

(b) Test, at the 5% significance level, whether or not there is evidence of a decrease in the rate of customers entering Jeff's supermarket. State your hypotheses clearly.

**(4)**

A further randomly selected 20-minute period is observed and the hypothesis test is repeated.
Given that the true rate of customers entering Jeff's supermarket is now 1 every 5 minutes,

(c) calculate the probability of a Type II error.

**(5)**

**(Total for question = 13 marks)**

**Q4.**

The number of accidents per year in *Daftstown* follows a Poisson distribution with mean $\lambda$.
The value of $\lambda$ has previously been 6 but Jonty claims that since the Council increased the speed limit, the value of $\lambda$ has increased.

Jonty records the number of accidents in *Daftstown* in the first year after the speed limit was increased. He plans to test, at the 5% significance level, whether or not there is evidence of an increase in the mean number of accidents in *Daftstown* per year.

(a) Stating your hypotheses clearly, calculate the probability of a Type I error for this test.

**(4)**

Given that there were 9 accidents in the first year after the speed limit was increased,

(b) state, giving a reason, whether or not there is evidence to support Jonty's claim.

**(2)**

(c) Given that the value of $\lambda$ has actually increased to 8, calculate the probability of drawing the conclusion, using this test, that the number of accidents per year in *Daftstown* has not increased.

**(2)**

**(Total for question = 8 marks)**

**Q5.**

A manufacturer has a machine that produces lollipop sticks.
The length of a lollipop stick produced by the machine is normally distributed with unknown mean $\mu$ and standard deviation 0.2

Farhan believes that the machine is not working properly and the mean length of the lollipop sticks has decreased.
He takes a random sample of size $n$ to test, at the 1% level of significance, the hypotheses

$$H_0: \mu = 15 \qquad H_1: \mu < 15$$

(a) Write down the size of this test.

**(1)**

Given that the actual value of $\mu$ is 14.9

(b) (i) calculate the minimum value of $n$ such that the probability of a Type II error is less than 0.05
Show your working clearly.

**(6)**

(ii) Farhan uses the same sample size, $n$, but now carries out the test at a 5% level of significance. Without doing any further calculations, state how this would affect the probability of a Type II error.

**(1)**

**(Total for question = 8 marks)**

## Mark Scheme – Quality of Tests

**Q1.**

| Qu | Scheme | Marks | AO |
|---|---|---|---|
| **(a)** | $H_0 : \lambda = 2.5$ (or $\mu = 7.5$) $H_1 : \lambda \neq 2.5$ (or $\mu \neq 7.5$) | B1 | 2.5 |
| | [$X$ = no. of accidents in a 3-month period] $X \sim \text{Po}(7.5)$ | M1 | 3.3 |
| | $P(X \leqslant 2) = 0.0203$ (calc: 0.020256…) { or $P(X \leqslant 3) = 0.0591$} | | |
| | $P(X \leqslant 13) = 0.9784$ so $P(X \geqslant 14) = 0.0216$ (calc: 0.0215646…) {or $P(X \geqslant 15) = 0.0103$} | M1 | 3.4 |
| | Giving Critical region of: $\boldsymbol{X \leqslant 2}$ | A1 | 1.1b |
| | $\boldsymbol{X \geqslant 14}$ | A1 | 1.1b |
| | | (5) | |
| **(b)** | [0.0203 +0.0216] =awrt **<u>0.0419</u>** <u>or</u> (calc: 0.041821366… awrt **<u>0.0418</u>**) | B1ft | 1.2 |
| | | (1) | |
| **(c)** | [Let $M$ = no of 3-month periods with a significant result] | | |
| | $M \sim B(8, \text{“}0.0419\text{”})$ | M1 | 3.3 |
| | $[P(M \geqslant 2)] = 1 - P(M \leqslant 1)$ | M1 | 1.1b |
| | $[= 1 - 0.9584…]$ | | |
| | =0.04153…(calc: 0.041394…) **<u>[0.04139~ 0.04154]</u>** | A1cso | 1.1b |
| | | (3) | |
| **(d)** | $Y \sim \text{Po}(6.3)$ | M1 | 3.3 |
| | P(Type II error) $= P(3 \leqslant Y \leqslant 13)$ or $P(Y \leqslant 13) - P(Y \leqslant 2)$ | M1 | 3.4 |
| | $[= 0.9945147… - 0.049846…]$ | | |
| | $= 0.9446…$ awrt **<u>0.945</u>** | A1 | 1.1b |
| | | (3) | |
| | | **(12 marks)** | |

| | **Notes** |
|---|---|
| **(a)** | B1 for both hypotheses in terms of $\lambda$ or $\mu$ (either way around)<br>1st M1 for selecting the correct Po model. Sight or use of Po(7.5) may be implied by 2nd M1<br>2nd M1 for using the correct model to find one of these probs with correct label (2sf or better)<br>1st A1 for one end correct<br>2nd A1 for a fully correct CR<br>Allow any letter, even CR $\leqslant 2$ or set notation but **not** $P(X \leqslant 2)$<br>Can have $X < 3$ and $X > 13$ etc |
| **(b)** | B1ft for awrt 0.0419 <u>or</u> awrt 0.0418<br><u>or</u> ft addition of their **two** probs provided both are $0 < \text{prob} < 0.025$ (awrt 3sf) |
| **(c)** | 1st M1 for selecting a correct binomial model, ft their answer to part (b)<br>2nd M1 for a correct probability statement of $1 - P(M \leqslant 1)$ **dep on** <u>a</u> binomial selected<br>A1cso for answer in range [0.04139, 0.04154] **dep on** use of B(8, “0.0419”) or better |
| **(d)** | 1st M1 for selecting a Po(6.3) model<br>2nd M1 for a correct probability statement using their Poisson model and their CR in (a) which may have just one tail.<br>A1 for awrt 0.945 |

**Q2.**

| Question | Scheme | Marks | AOs |
|---|---|---|---|
| (a) | $X \sim B(20, 0.2)$ and seek $c$ such that $P(X \leqslant c) < 0.10$ | M1 | 3.3 |
| | $[P(X \leqslant 1) = 0.0692]$ CR is $X \leqslant 1$ | A1 | 1.1b |
| | | (2) | |
| (b) | Size = $\underline{0.0692}$ | B1ft | 1.2 |
| | | (1) | |
| (c) | $Y$ = no. of spins until red obtained so $Y \sim \text{Geo}(0.2)$ | M1 | 3.3 |
| | $\mu = \frac{1}{p}$ so if $p < 0.2$ then mean is <u>larger</u> so seek $d$ so that $P(Y \geqslant d) < 0.10$ | M1 | 2.4 |
| | $P(Y \geqslant d) = (0.8)^{d-1}$ | M1 | 3.4 |
| | $(0.8)^{d-1} < 0.10 \Rightarrow d - 1 > \frac{\log(0.1)}{\log(0.8)}$ | M1 | 1.1b |
| | $d > 11.3..$ | A1 | 1.1b |
| | CR is $Y \geqslant 12$ | A1 | 2.2b |
| | | (6) | |

| Question | Scheme | Marks | AOs |
|---|---|---|---|
| (d) | Size = $[0.8^{11} = 0.085899\ldots] = \underline{0.0859}$ | B1 | 1.1b |
| | | (1) | |
| (e)(i) | Power = P(reject $H_0$ when it is false) = $P(X \leqslant 1 \mid X \sim B(20, p))$ | M1 | 2.1 |
| | $= (1-p)^{20} + 20(1-p)^{19} p$ | M1 | 1.1b |
| | $= (1-p)^{19}(1+19p)$ * | A1*cso | 1.1b |
| (ii) | Power = $(1-p)^{11}$ | B1 | 1.1b |
| | | (4) | |
| (f) | Sam's test has smaller P(Type I error) (or size) so is better | B1 | 2.2a |
| | Power of Sam's test = 0.1755… | B1 | 1.1b |
| | Power of Tessa's test = $0.85^{11} = 0.1673\ldots$ | B1 | 1.1b |
| | So for $p = 0.15$ **Sam's test** is recommended | B1 | 2.2b |
| | | (4) | |
| | | | **(18 marks)** |
| | **Notes** | | |
| (a) | **M1:** Realising the need to use the model Using $B(20,0.2)$ with method for finding the CR or implied by a correct CR<br>**A1**: $X \leqslant 1$ or $X < 2$ | | |
| (b) | **B1:** awrt 0.0692 | | |

| Notes (continued) | |
|---|---|
| **(c)** | **M1:** Realising that the model Geo(0.2)is needed. This may be written or used<br>**M1:** Realising the key step that they need to find $P(Y \geqslant d) < 0.10$<br>**M1:** Using the model $(0.8)^{d-1}$<br>**M1:** Using the model $(0.8)^{d-1} < 0.10$ and finding a method to solve leading to a value/range of values for $d$<br>**A1:** For $d > 11.3..$<br>**A1:** For $Y \geqslant 12$ or $Y > 11$ (a correct inference) |
| **(d)** | **B1ft:** awrt 0.0692. ft their answer to part (c) |
| **(e)(i)** | **M1:** Using B(20, $p$) and realizing they need to find $P(X \leqslant 1)$ o.e. This may be used or written<br>**M1:** Using $P(X = 0) + P(X = 1)$<br>**A1*cso:** Fully correct proof ( no errors) |
| **(ii)** | **B1:** For $(1-p)^{11}$ |
| **(f)** | **B1:** Making a deduction about the tests using the answers to part(b) and (d)<br>**B1:** awrt 0.0176<br>**B1:** awrt 0.167<br>**B1:** A correct inference about which test is recommended |

**Q3.**

| Question | Scheme | Marks | AOs |
|---|---|---|---|
| **(a)(i)** | $X \sim \text{Po}(24)$ | B1 | 3.4 |
| | $P(X = 26) = 0.071912...$ awrt <u>0.0719</u> | B1 | 1.1b |
| | | **(2)** | |
| **(ii)** | $P(X \geq 21) = 1 - P(X \leq 20)$ $[= 1 - 0.24263...]$ | M1 | 3.4 |
| | $= 0.75736...$ awrt <u>0.757</u> | A1 | 1.1b |
| | | **(2)** | |
| **(b)** | $H_0: \lambda = 2 \quad [\mu = 16]$<br>$H_1: \lambda < 2 \quad [\mu < 16]$ | B1 | 2.5 |
| | $P(Y \leq 10 \mid Y \sim \text{Po}(16)) = 0.077396...$ awrt <u>0.0774</u> | B1 | 1.1b |
| | Not significant / Do not reject $H_0$ / 10 is not in the CR | M1 | 1.1b |
| | There is <u>not</u> sufficient evidence to suggest a decrease/change in the rate of <u>customers</u> entering Jeff's supermarket. | A1 | 2.2b |
| | | **(4)** | |
| **(c)** | Use of Po(8) to attempt critical region | M1 | 2.1 |
| | Critical region is $Y \leq 3$/ $H_0$ is not rejected when $Y \geq 4$ | A1 | 1.1b |
| | True distribution is $W \sim \text{Po}(4)$ | B1 | 2.1 |
| | $P(W \geq 4 \mid W \sim \text{Po}(4)) = 1 - P(W \leq 3)$ $[= 1 - 0.43347...]$ | M1 | 1.1b |
| | $=0.56652...$ awrt <u>0.567</u> | A1 | 1.1b |
| | | **(5)** | |
| | | | **(13 marks)** |

| Notes | |
|---|---|
| **(a)(i)** **(ii)** | **B1**: For realising the distribution is Po(24) (May be seen or implied in part (ii))<br>**B1**: awrt 0.0719<br>**M1**: Writing or using $1 - P(X \leq 20)$<br>**A1**: awrt 0.757 |
| **(b)** | **B1**: Both hypotheses correct (must use $\mu$ or $\lambda$ )<br>**B1**: awrt 0.0774 Allow awrt 0.08 from a correct probability statement.<br>allow CR: $X \leq 9$<br>**M1**: Correct non-contextual conclusion (may be implied by correct contextual conclusion). Allow a f.t. comparison of 'their $p$' with 0.05<br>(Ignore any contradictory contextual comments for this mark)<br>**A1**: A fully correct solution drawing a correct inference in context **with all previous marks in (b) scored.** |
| **(c)** | **M1**: Use of Po(8) to attempt critical region [$P(Y \leq 3)$=0.0423.. $P(Y \leq 4)$=0.0996..]<br>**A1**: Finding critical region for the test $Y \leq 3$ which must come from Po(8).<br>**B1**: Identifying the need to use Po(4) as the true distribution.<br>Allow Po(4) seen or used for this mark.<br>**M1**: Writing or using $P(W \geq$ '4') or $1 - P(W \leq$ '3') from Po(4). Allow f.t. on their **identified** CR but must be using Po(4)<br>**A1**: awrt 0.567 |

**Q4.**

| Question Number | Scheme | | Marks |
|---|---|---|---|
| (a) | $H_0 : \lambda = 6,\ H_1 : \lambda > 6$ | both | B1 |
| | $P(X \geq 10) = 0.0839$ | | M1 |
| | $P(X \geq 11) = 0.0426$ | | |
| | CR $X \geq 11$ | | A1 |
| | P (Type I Error) = 0.0426 | | A1 |
| | | | **(4)** |
| (b) | 9 is not in the critical region therefore<br>there is no evidence of an increase in the number of accidents per year<br>**or** there is no evidence to support Jonty's claim | | M1<br>A1ft |
| | | | **(2)** |
| (c) | $\lambda$=8 | | |
| | $P(X \leq 10 \mid \lambda = 8) = 0.8159$ | | M1A1 |
| | | | **(2)** |

| | Notes | Total 8 |
|---|---|---|
| (a) | B1 both hypotheses, allow use of $\mu$<br><br>M1 for seeing $[P(X \geq 10) =]0.0839$ or $[P(X \geq 11) =]0.0426$ or $[P(X \leq 9) =]0.9161$ or $[P(X \leq 10) =]0.9574$ oe allow a sideways slip of 1. ie 6.5/5.5<br>A1 for seeing $P(X \leq 10) = 0.9574$ or $P(X \geq 11) = 0.0426$ or CR $X \geq 11$<br><br>A1 0.0426<br><br>**NB** An answer of 0.0426 implies will get M1A1A1 | |
| (b) | M1 must have 9/ value oe is not in CR allow $0.153 > 0.05$<br>A1ft correct statement in context – need accidents or Jonty | |
| (c) | M1 $P(X \leq c-1 \mid \lambda = 8)$ with $c-1$ being correct or using their $c$. Allow if a CR is stated in the form $X \leq c$ for $1 - P(X \leq c \mid \lambda = 8)$<br>A1 awrt 0.816 | |

**Q5.**

| Question | Scheme | Marks | AOs |
|---|---|---|---|
| (a) | Size of the test = 0.01 | B1 | 1.2 |
| | | (1) | |
| (b)(i) | Let CR be $\bar{L} < k$ | | |
| | $\dfrac{k-15}{\frac{0.2}{\sqrt{n}}} = -2.3263$ | M1 | 3.4 |
| | $k = 15 - \dfrac{0.46526}{\sqrt{n}}$ | A1 | 1.1b |
| | $\dfrac{"15 - \frac{0.46526}{\sqrt{n}}" - 14.9}{\frac{0.2}{\sqrt{n}}} > 1.6449$ | M1d<br>A1ft | 3.4<br>1.1b |
| | $\dfrac{0.79424}{\sqrt{n}} < 0.1 \quad \sqrt{n} > 7.9424$ oe | M1d | 1.1b |
| | $n = 64$ | A1cso | 2.1 |
| | | (6) | |
| (ii) | The probability of a Type II error would decrease. | B1 | 2.2a |
| | | (1) | |
| | | | **(8 marks)** |

| Notes | | |
|---|---|---|
| (a) | B1: | 0.01 |
| (b)(i) | M1: | Finding the CR using the Normal distribution must have $1.5<\lvert z\rvert<3.5$ |
| | A1: | A correct equation in the form $k=\ldots$ and for use of awrt 2.326 (implied by awrt 0.46526 or awrt 0.46527) |
| | M1d: | Dependent on previous M being awarded. Standardising using their $k$ and equating to a $z$ value $1.5<\lvert z\rvert<3$ to form an equation to able $n$ to be found. May use = rather than > |
| | A1ft: | Ft their $k$ for a correct equation with awrt 1.645 |
| | M1d: | Dependent on previous M being awarded. Isolating $\sqrt{n}$ or squaring both sides leading to a value for $n$. Condone $n = 7.9424$ |
| | A1cso: | 64 with correct working |
| (ii) | B1: | Suitable comment |
| | | |

| | | | |
|---|---|---|---|
| ALT (b)(i) | $\dfrac{k-14.9}{\frac{0.2}{\sqrt{n}}}=1.6449$ | M1 | 3.4 |
| | $k=14.9+\dfrac{0.32898}{\sqrt{n}}$ | A1 | 1.1b |
| | $\dfrac{\text{"}14.9+\frac{0.32898}{\sqrt{n}}\text{"}-15}{\frac{0.2}{\sqrt{n}}}>-2.3263$ | M1d<br>A1ft | 3.4<br>1.1b |
| | $\dfrac{0.79424}{\sqrt{n}}<0.1 \quad \sqrt{n}>7.9424$ oe | M1d | 1.1b |
| | $n=64$ | A1cso | 2.1 |
| | | (6) | |