Question 1

(a)

A1A1

[2 marks]

(b) (i) $P(Y) = 0.8 \times 0.1 + 0.2 \times 0.3$ *M1*

$= 0.14$ *A1*

(ii) $P(\text{Star} \mid Y) = \dfrac{0.8 \times 0.1}{0.14}$ *M1*

$= 0.571 \left(\dfrac{4}{7}, \ 0.571428\ldots \right)$ *A1*

[4 marks]

(c) the colours of the sweets are distributed according to
manufacturer specifications   *A1*

[1 mark]

(d)

Colour	Brown	Red	Green	Orange	Yellow	Purple
Expected Frequency	12	20	16	16	8	8

A2

(e)   5   *A1*

[1 mark]

(f)   0.469 (0.4688117…)   *A2*

[2 marks]

(g)   since $0.469 > 0.05$ *R1*

fail to reject the null hypothesis. There is insufficient evidence to
reject the manufacturer's specifications   *A1*

[2 marks]

Total [14 marks]

Question 2

(a)   (i)   evidence of correct probability   *(M1)*
e.g sketch **OR** correct probability statement, $P(X < 6.5)$

0.0151   *A1*

(ii)   0.0228   *A1*

Note: Answers should be given to 4 decimal place.

[3 marks]

(b)   (i)   multiplying **their** probability by 1000   *(M1)*
451.7   *A1*

(ii)   510.5   *A1*

[3 marks]

Note: Answers should be given to 4 sf.

(c) H_0 : stopping distances can be modelled by $N(6.76, 0.12^2)$
H_1 : stopping distances cannot be modelled by $N(6.76, 0.12^2)$ *A1A1*

[2 marks]

(d)   15.1 or 22.8 seen   *(M1)*

0.0727 (0.0726542…, 7.27%)   *A2*

[3 marks]

(e) $0.05 < 0.0727$ *R1*
there is insufficient evidence to reject H_0 (or "accept H_0")   *A1*

Note: Do not award **R0A1**.

[2 marks]
Total [13 marks]

Question 3

(a) Quota sampling   *A1*
                                                                                                          [1 mark]

(b) 10 (hours)   *A1*
                                                                                                          [1 mark]

(c) $15 - 7$ *(M1)*

> **Note:** Award *M1* for 15 and 7 seen.

 8   *A1*
                                                                                                          [2 marks]

(d) indication of a valid attempt to find the upper fence   *(M1)*
 $15 + 1.5 \times 8$
 27   *A1*

 $25 < 27$ (accept equivalent answer in words)   *R1*
 Jason is correct   *A1*

> **Note:** Do not award *R0A1*. Follow through **within** this part from *their* 27,
> but only if their value is supported by a valid attempt **or** clearly
> and correctly explains what their value represents.

                                                                                                          [4 marks]

(e) "negative" seen   *A1*

> **Note:** Strength cannot be inferred visually; ignore "strong" or "weak".

                                                                                                          [1 mark]

(f) correct substitution   *(M1)*
 $y = -1.54 \times 1.5 + 98.8$

 96.5 (%) (96.49)   *A1*
                                                                                                          [2 marks]

(g) not reliable   *A1*
 extrapolation **OR** outside the given range of the data   *R1*

> **Note:** Do not award *A1R0*. Only accept reasoning that includes reference to the range
> of the data. Do not accept a contextual reason such as 1.5 hours is too short to
> read the book.

                                                                                                          [2 marks]

(h)

	Book							
	A	B	C	D	E	F	G	H
Rank – Number of pages	1	3	5	2	6	8	4	7
Rank – Top 50 Rating	1	2	3	4	5	6	7	8

A1A1

> **Note:** Award *A1* for correct ranks for 'number of pages'. Award *A1* for correct ranks for 'top 50 rating'.

[2 marks]

(i)   (i)   0.714 (0.714285...)   *A2*

> **Note:** *FT* from their table.

(ii)   **EITHER**
there is a (strong/moderate) positive association between the number of pages and the top 50 rating.   *A1*

OR
there is a (strong/moderate) agreement between the rank order of number of pages and the rank order top 50 rating.   *A1*

OR
there is a (strong/moderate) positive (linear) correlation between the rank order of number of pages and the rank order top 50 rating.   *A1*
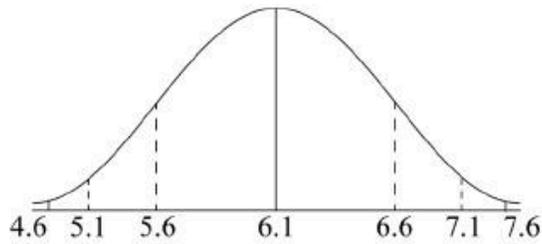
> **Note:** Follow through from their value of r_s .

[3 marks]

Total [18 marks]

Question 4

(a)

A1A1

Note: Award **A1** for a normal curve with mean labelled 6.1 or μ, **A1** for indication of SD (0.5): marks on horizontal axis at 5.6 and/or 6.6 **OR** $\mu - 0.5$ and/or $\mu + 0.5$ on the correct side and approximately correct position.

[2 marks]

(b) $X \sim N(6.1, 0.5^2)$

$P(5.5 < X < 6.5)$ **OR** labelled sketch of region (M1)

$= 0.673 \ (0.673074...)$ A1

[2 marks]

(c) $(P(X < 5.3) =) \ 0.0547992...$ (A1)

$0.0547992... \times 80$ (M1)

$= 4.38 \ (4.38393...)$ A1

[3 marks]

(d) 0.15 **OR** 0.85 (A1)

$P(X > x) = 0.15$ **OR** $P(X < x) = 0.85$ **OR** labelled sketch of region (M1)

6.62 (6.61821...) A1

[3 marks]

(e) $(P(X > 6.25) =) \ 0.382088...$ (A1)

recognition of binomial (M1)

e.g. B(10, 0.382088...)

0.0502 (0.0501768...) A2
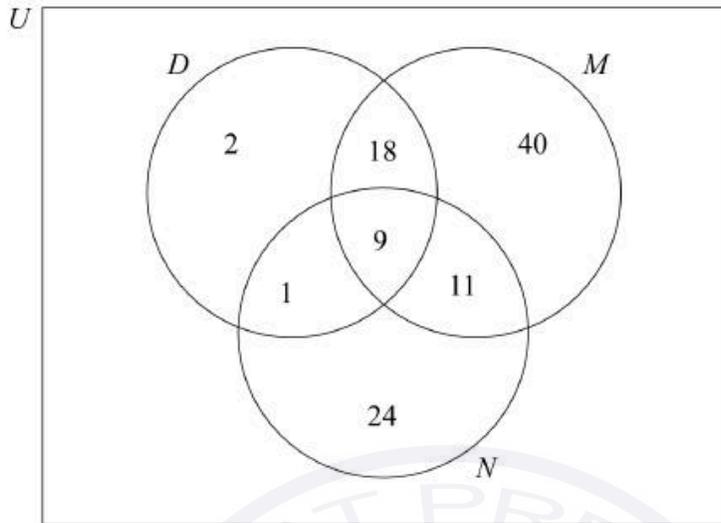
[4 marks]

Total [14 marks]

Question 5

(a) convenience sampling   **A1**

[1 mark]

(b) (i) 95%   **A1**

    (ii) 1%   **A1**

    (iii) 2%   **A1**

    (iv) 98%   **A1**

[4 marks]

(c) (i) 0.95×0.02 **(M1)**
       0.019   **A1**

    (ii) $0.05 \times 0.01 + 0.95 \times 0.98$ **(M1)(M1)**

> **Note:** Award **M1** for summing two products and **M1** for correct products seen.

       0.932 (0.9315)   **A1**

    (iii) recognition of conditional probability   **(M1)**

$$\frac{0.05 \times 0.01}{0.05 \times 0.01 + 0.95 \times 0.98}$$ **A1**

       0.000537 (0.000536768…)   **A1**

> **Note:** Accept 0.000536 if 0.932 used.

[8 marks]

(d) **EITHER**
sample may not be representative of population   **A1**
OR
sample is not randomly selected   **A1**
OR
unrealistic to think expected and observed values will be exactly equal   **A1**

[1 mark]

(e)

                                                                                    A1A1A1

| **Note:** | Award **A1** for rectangle and 3 labelled circles and 9 in centre region; **A1** for 2, 40, 24; **A1** for 18, 1, and 11. |

                                                                                    [3 marks]

(f) $18+9+1+11+2+40+24$ **(M1)**
     105 **A1**

| **Note:** | Follow through from the entries on their Venn diagram in part (e). Working required for **FT**. |

                                                                                    [2 marks]

                                                                            Total [19 marks]

Question 6

(a) (i) Let X be the random variable "distance from O".
$X \sim N(10, 3^2)$
$P(X < 13) = 0.841$ $(0.841344...)$ **(M1)A1**

(ii) $(P(X > 15) =)$ 0.0478 (0.0477903) **A1**

[3 marks]

(b) $P(X > 15) \times P(X > 15)$ **(M1)**
$= 0.00228$ $(0.00228391...)$ **A1**

[2 marks]

(c) $1 - (0.8143)^3$ **(M1)**
0.460 $(0.460050...)$ **A1**

[2 marks]

(d) (i) **METHOD 1**
let Y be the random variable "number of points scored"
evidence of use of binomial distribution   **(M1)**
$Y \sim B(10, 0.539949...)$ **(A1)**
$(P(Y \geq 5) =)$ 0.717 $(0.716650...)$.   **A1**

METHOD 2
let Q be the random variable "number of times a point is not scored"
evidence of use of binomial distribution   **(M1)**
$Q \sim B(10, 0.460050...)$ **(A1)**
$(P(Q \leq 5) =)$ 0.717 $(0.716650...)$ **A1**

(ii) $P(5 \leq Y < 8)$ **(M1)**
0.628 $(0.627788...)$ **A1**

Note: Award **M1** for a correct probability statement or indication of correct lower and upper bounds, 5 and 7.

(iii) $\dfrac{P(5 \leq Y < 8)}{P(Y \geq 5)}$ $\left(= \dfrac{0.627788...}{0.716650...} \right)$ **(M1)**
0.876 $(0.876003...)$ **A1**

[7 marks]
Total: [14 marks]

Question 7

(a) (i) $\dfrac{560}{1280} \left(\dfrac{7}{16}, \; 0.4375 \right)$ **A1A1**

> **Note:** Award **A1** for correct numerator, **A1** for correct denominator.

(ii) $\dfrac{72}{1280} \left(\dfrac{9}{160}, \; 0.05625 \right)$ **A1A1**

> **Note:** Award **A1** for correct numerator, **A1** for correct denominator.

(iii) $\dfrac{153}{348} \left(\dfrac{51}{116}, \; 0.439655... \right)$ **A1A1**

> **Note:** Award **A1** for correct numerator, **A1** for correct denominator.

(iv) $160+224+128+205+131$ **OR** $560+512-224$ **(M1)**

$\dfrac{848}{1280} \left(\dfrac{53}{80}, \; 0.6625 \right)$ **A1A1**

> **Note:** Award **A1** for correct denominator (1280) seen, **(M1)** for correct calculation of the numerator, **A1** for the correct answer.

[9 marks]

(b) H_0 : the variables are independent
   H_1 : the variables are dependent   **A1**

> **Note:** Award **A1** for both hypotheses correct. Do not accept "not correlated" or "not related" in place of "independent".

[1 mark]

(c) 4   **A1**

[1 mark]

(d) (i) $(\chi^2 =)$ 23.3 $(23.3258...)$ **A2**

(ii) 0.000109 $(0.000108991...)$ **OR** 1.09×10^{-4} **A1**

(iii) **EITHER**
$23.3 > 13.277$ **R1**
OR
$0.000109 < 0.01$ **R1**

THEN
(there is sufficient evidence to accept H_1 that) preferred device and age group are not independent   **A1**

> **Note:** For the final **A1** the answer must be in context. Do not award **A1R0**.

[5 marks]
Total: [16 marks]

Question 8

(a) (let T be the number of passengers who arrive)

$$\left(P(T > 72) =\right)\ P(T \geq 73) \quad \textbf{OR} \quad 1 - P(T \leq 72) \qquad \textbf{(A1)}$$

$$T \sim B(74,\ 0.9) \quad \textbf{OR} \quad n = 74 \qquad \textbf{(M1)}$$

$$= 0.00379 \ \ (0.00379124...) \qquad \textbf{A1}$$

Note: Using the distribution $B(74,\ 0.1)$, to work with the 10% that do not arrive for the flight, here and throughout this question, is a valid approach.

[3 marks]

(b) (i) 72×0.9 **(M1)**

            64.8   **A1**

     (ii) $n \times 0.9 = 72$ **(M1)**

            80   **A1**

[4 marks]

(c) **METHOD 1**

EITHER

when selling 74 tickets

	$T \leq 72$	$T = 73$	$T = 74$
Income minus compensation (I)	11100	10800	10500
Probability	0.9962…	0.003380…	0.0004110…

     top row   **A1A1**

     bottom row   **A1A1**

Note: Award **A1A1** for each row correct. Award **A1** for one correct entry and **A1** for the remaining entries correct.

$$E(I) = 11100 \times 0.9962... + 10800 \times 0.00338... + 10500 \times 0.000411 \approx 11099 \qquad \textbf{(M1)A1}$$

OR

income is $74 \times 150 = 11100$ **(A1)**

expected compensation is

$0.003380... \times 300 + 0.0004110... \times 600 \ \ (= 1.26070...)$ **(M1)A1A1**

expected income when selling 74 tickets is $11100 - 1.26070...$ **(M1)**

$= 11098.73.. \ \ (= \$11099)$ **A1**

THEN

income for 72 tickets $= 72 \times 150 = 10800$ **(A1)**

so expected gain $\approx 11099 - 10800 = \$299$ **A1**

Question 9

(a)   (i) $\dfrac{370+472}{2}$ *(M1)*

> **Note:** This *(M1)* can also be awarded for either a correct Q_3 or a correct Q_1 in part (a)(ii).

$Q_3 = 421$ *A1*

(ii)   their part (a)(i) – their Q_1 (clearly stated)   *(M1)*
$IQR = (421-318=)\ 103$ *A1*

[4 marks]

(b) $(Q_3 + 1.5(IQR) =)\ 421+(1.5\times103)$ *(M1)*
$= 575.5$
since $498 < 575.5$ *R1*
Netherlands is not an outlier   *A1*

> **Note:** The *R1* is dependent on the *(M1)*. Do not award *R0A1*.

[3 marks]

(c)   not appropriate ("no" is sufficient)   *A1*
as r is too close to zero / too weak a correlation   *R1*

[2 marks]

(d)   (i)   6   *A1*

(ii)   4.5   *A1*

(iii)   4.5   *A1*

[3 marks]

(e)   (i) $r_s = 0.683\ (0.682646...)$ *A2*

(ii)   **EITHER**
there is a (positive) association between the population size and
the score   *A1*

OR
there is a (positive) linear correlation between the ranks of the population size
and the ranks of the scores (when compared with the PMCC of 0.249).   *A1*

[3 marks]

(f)   lowering the top score by 20 does not change its rank so r_s is unchanged   *R1*

> **Note:** Accept "this would not alter the rank" or "Netherlands still top rank" or similar. Condone any statement that clearly implies the ranks have not changed, for example: "The Netherlands still has the highest score."

[1 mark]
[Total 16 marks]

Question 10

(a) (i)   0.58 (s)   *A1*

   (ii) $0.7 - 0.42$ *(A1)(M1)*

> **Note:** Award *A1* for correct quartiles seen, *M1* for subtraction of their quartiles.

      0.28 s   *A1*
                                                                                              [4 marks]

(b)   9 (people have reaction time ≤ 0.4)   *(A1)*
      31 (people have reaction time > 0.4)   *A1*
                                                                                              [2 marks]

(c) $(90\% \times 40 =)$ 36 **OR** 4   *(A1)*
      0.8 s   *A1*
                                                                                              [2 marks]

(d) (i) $(a =)$ 6   *A1*

   (ii) $(b =)$ 4   *A1*
                                                                                              [2 marks]

(e) $0.6 < t \leq 0.8$ *A1*
                                                                                               [1 mark]

(f)   0.55 s   *A2*
                                                                                              [2 marks]

(g)   the mean will increase   *A1*
      because the incorrect reaction times are moving from a lower interval
      to a higher interval which will increase the numerator of the mean calculation   *R1*

      the median will stay the same   *A1*
      because the median or middle of the data is greater than both intervals
      being changed   *R1*

> **Note:** Do not award *A1R0*.

                                                                                              [4 marks]
                                                                                         Total [17 marks]

Question 11

(a) (i) $(m =)$ 54(%)   **A1**

   (ii) $(n =)$ 14(%)   **A1**

   (iii) $(p =)$ 22(%)   **A1**

   (iv) $(q =)$ 10(%)   **A1**

> **Note:** Based on their n, follow through for parts (i) and (iii), but only if it does not contradict the given information. Follow through for part (iv) but only if the total is 100%.

[4 marks]

(b) 90 (%)   **A1**

> **Note:** Award **A0** for a decimal answer.

[1 mark]

(c) (i) $0.54 \left(\dfrac{54}{100}, \dfrac{27}{50}, 54\% \right)$ **A1**

   (ii) $\dfrac{54}{64} \left(0.844, \dfrac{27}{32}, 84.4\%, 0.84375 \right)$ **A1A1**

> **Note:** Award **A1** for a correct denominator (0.64 or 64 seen), **A1** for the correct final answer.

[3 marks]

(d) (i) recognizing Binomial distribution with correct parameters   **(M1)**
$X \sim \mathrm{B}(10, 0.68)$
$(\mathrm{P}(X = 5) =)$ 0.123 (0.122940..., 12.3%)   **A1**

   (ii) $1 - \mathrm{P}(X \leq 3)$ **OR** $\mathrm{P}(X \geq 4)$ **OR** $\mathrm{P}(4 \leq X \leq 10)$ **(M1)**
0.984 (0.984497..., 98.4%)   **A1**

   (iii) $(0.68)^9 \times 0.32$ **(M1)**
recognition of two possible cases   **(M1)**
$2 \times \left((0.68)^9 \times 0.32 \right)$
0.0199 (0.0198957..., 1.99%)   **A1**

[7 marks]

(e) **EITHER**
the probability is not constant   **A1**
OR
the events are not independent   **A1**
OR
the events should be modelled by the hypergeometric distribution instead   **A1**

[1 mark]
Total [16 marks]

Question 12

(a) continuous $\qquad$ **A1**

[1 mark]

(b) $160 - 50 - 62 - 14 - 8$ $\qquad$ **(M1)**

$(k =)\ 26$ $\qquad$ **A1**

[2 marks]

(c) (i) $20 \leq T < 40$ $\qquad$ **A1**

(ii) 30 $\qquad$ **A1**

[2 marks]

(d) 33.5 minutes $\qquad$ **A2**

Note: **FT** from their value of k and their mid-interval value. Follow through from part (c)(ii) but only if mid-interval value lies in their interval.

[2 marks]

(e) 112 $\qquad$ **A1**

[1 mark]

(f) $\dfrac{22}{160}\ \left[0.138,\ 0.1375,\ 13.75\%,\ \dfrac{11}{80}\right]$ $\qquad$ **A1A1**

Note: Award **A1** for correct numerator, **A1** for correct denominator.

[2 marks]

(g) 26 minutes $\qquad$ **A1**

[1 mark]

(h) $50 - 16$ $\qquad$ **(M1)**

Note: Award **M1** for both correct quartiles seen.

34 minutes $\qquad$ **A1**

[2 marks]

(i) correct substitution into outlier formula   **(M1)**
     $50 + 1.5 \times 34$
     $= 101$ **A1**
     $92 < 101$ **OR** highest value on diagram < 101 **R1**
     not an outlier   **AG**

> **Note:** Award **R1** for their correct comparison. Follow through from their part (h).
> Award **R0** if their conclusion is "it is an outlier", this contradicts Elsie's belief.

                                                                                         [3 marks]

(j) **EITHER**
     the diagram is not symmetric or equivalent
         e.g the median is not in the center of the box or
         the lengths of the whiskers are (very) different or (positive or right) skew

     OR
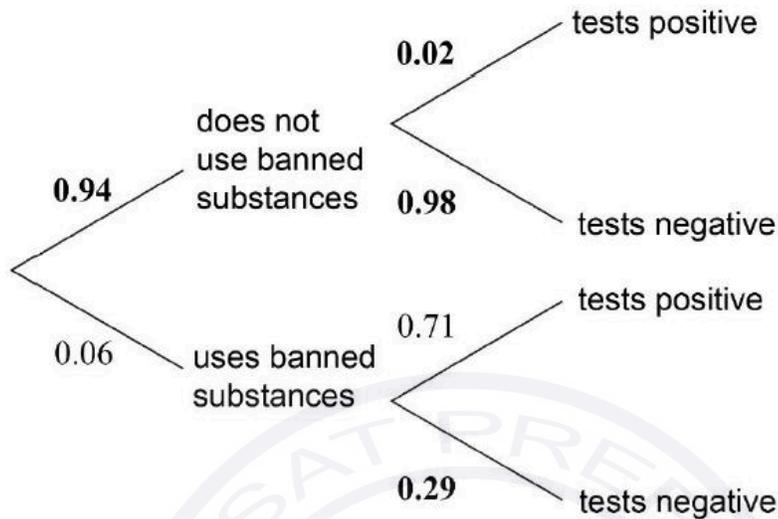   the mean and median are (very) different;   **A1**

                                                                                         [1 mark]
                                                                                 Total [17 marks]

Question 13

(a)

A1A1

> **Note:** Award **A1** for any one value correct, **A1** for other three values correct. Accept percentage responses as equivalent forms on **all** branches.

[2 marks]

(b) (i) multiplication of two probabilities along the tree diagram   **(M1)**

0.94×0.98

$= 0.921 \; (0.9212, \; 92.1\%, \; 92.12\%)$ **A1**

> **Note:** Do not accept the 2sf value for the final **A1**.

(ii) $(0.9212)^2$ **(A1)**

$= 0.849 \; (0.848609\ldots, \; 84.9\%, \; 84.8609\ldots\%)$ **A1**

> **Note:** Accept an answer of 0.848 (0.848241) from use of 3 sf answer from part (b)(i).

[4 marks]

(c) (i) $0.94 \times 0.02 + 0.06 \times 0.29$ *(A1)(M1)*

　　　　　　0.0362 (3.62%)   **A1**

　　(ii)   multiplying their part(c)(i) by 1300

　　　　　　0.0362×1300 *(M1)*

　　　　　　47.1 (47.06)   **A1**

[5 marks]

(d) $p = 0.02$ **OR** $p = 0.98$ *(A1)*

　　　recognition of binomial probability with $n = 20$ *(M1)*

　　　$P(X = 0)$ **OR** $P(X = 20)$ *(M1)*

　　　0.668 (0.667607…)   **A1**

[4 marks]

(e) $P(X \geq 3)$ **OR** $P(X \leq 17)$ *(M1)*

　　　0.00707 (0.00706869…)   **A1**

[2 marks]

[Total: 17 marks]

Question 14

(a) $\dfrac{9.45-8.73}{1958-1708}$ *(M1)*

$=0.00288 \left(\dfrac{9}{3125}\right)$ *A1*

[2 marks]

(b) (i) the (mean) yearly change in (mean annual) temperature *A1*

Note: Accept equivalent statements, e.g. "rate of change of temperature".

(ii) $°C$ / year **OR** degrees C per year *A1*

Note: Do not follow through from part (b)(i) into (b)(ii).

[2 marks]

(c) attempt to substitute point and gradient into appropriate formula *(M1)*

$8.73=0.00288\times1708+c \Rightarrow c=3.81096...$

or

$9.45=0.00288\times1958+c \Rightarrow c=3.81096 .$

equation is $y=0.00288x+3.81$ *A1*

[2 marks]

(d) attempt to substitute 2000 into their part (c) *(M1)*

$0.00288\times2000+3.81096...$

$=9.57\ (°C)\ (9.57096...)$ *A1*

[2 marks]

(e) (i) $y = 0.00256x + 4.46$ $(0.00255714...x + 4.46454...)$ **(M1)A1**

> **Note:** Award **(M1)A0** for answers that show the correct method, but are presented incorrectly (e.g. no "$y =$" or truncated values etc.). Accept 4.465 as the correct answer to 4 sf.

(ii) 0.861 $(0.861333...)$ **A1**

[3 marks]

(f) attempt to substitute 2000 into their part (e)(i) **(M1)**

$0.00255714... \times 2000 + 4.46454...$

$= 9.58(°C)$ $(9.57882...(°C))$ **A1**

> **Note:** Award **A1** for 9.57 from $0.00255714 \times 2000 + 4.46$.

[2 marks]

(g) cannot (always reliably) make a prediction of x from a value of y, when using a y on x line / regression line is not x on y **A1**

extrapolation **A1**

[2 marks]

[Total: 15 marks]

Question 15

(a) recognition of binomial distribution **(M1)**
e.g. $X \sim \mathrm{B}(115,\ 0.82)$ **OR** binompdf$(115,\ 0.82,\ 90)$ etc.

$\big(\big(\mathrm{P}(X=90)=\big)\ 0.0535\ (0.0535325...)$ **A2**

> **Note:** Award **(M1)A1A0** for an answer of 0.054 with or without working shown.

[3 marks]

(b) selecting correct region of distribution **(M1)**
e.g. $\mathrm{P}(X\geq 95)$ **OR** $1-\mathrm{P}(X\leq 94)$ **OR** $1-$binomcdf$(115,\ 0.82,\ 94)$

$0.491\ (0.491036...)$ **A1**

[2 marks]

(c) substitution in the variance formula for binomial distribution **(M1)**
$115\times 0.82\times 0.18$

$17.0\ \ (16.974)$ **A1**

> **Note:** Allow 17 for the final answer.

[2 marks]

(d) **METHOD 1**
attempt to write an expression containing n inside the brackets of $\mathrm{P}()$ **AND**
including 0.3 or 0.7 **(M1)**

$\mathrm{P}(X\geq n)<0.3$ **OR** $\mathrm{P}(X\leq n-1)>0.7$ **(A1)**
$n=98$ **A1**

METHOD 2
using binomcdf in GDC for at least two different values of n greater than 90 **(M1)**

EITHER
$\big(\mathrm{P}(X<97)=\big)\ 0.696683...$ **AND** $\big(\mathrm{P}(X<98)=\big)\ 0.778249...$ (seen) **(A1)**

OR
$\big(\mathrm{P}(X>97)=\big)0.303316...$ **AND** $\big(\mathrm{P}(X>98)=\big)0.221750...$ (seen) **(A1)**

THEN
$n=98$ **A1**

[3 marks]

(e)   (μ_1 : population mean recovery time for new remedy)

     (μ_2 : population mean recovery time for old remedy)

     $H_0 : \mu_1 = \mu_2 \quad (H_0 : \mu_1 - \mu_2 = 0)$ **A1**

     $H_1 : \mu_1 < \mu_2 \quad (H_1 : \mu_1 - \mu_2 < 0)$ **A1**

> **Note:** Accept an equivalent statement in words, must include mean and reference to "**population** mean", e.g. "mean for **all** patients on old remedy", for the first **A1** to be awarded.
>
> Do not accept an imprecise *"the means are equal"*.
>
> Award **A0A1** for reversed hypotheses ($H_0 : \mu_1 < \mu_2$, $H_1 : \mu_1 = \mu_2$).

[2 marks]

(f) $0.0620 \ (0.0620061\ldots)$ **A2**

> **Note:** Allow 0.062 as final answer. Award **A1** for an answer of 0.06 . Award **A1** for an answer of $0.0527756\ldots$ from use of unpooled setting.
>
> Follow through from an incorrect alternative hypothesis as long as their p-value matches their alternative hypothesis.

[2 marks]

(g) $0.0620 < 0.1$ **R1**

     (sufficient evidence to) reject H_0 **A1**

> **Note:** Do not award **R0A1**. Accept "p-value is less than 0.1" provided an answer was seen in part (f).

[2 marks]

(h)   the probability of obtaining results (at least as extreme) as those observed given that the null hypothesis is true   **A1**

[1 mark]

Total [17 marks]

Question 16

(a) (i) 75 (minutes)   *A1*

     (ii) attempt to substitute values in the mean formula with at least one
           mid-interval value multiplied by a corresponding frequency   *(M1)*

           (mean =) 88.2 (88.15) (minutes)   *A1*

                                                                                                   [3 marks]

(b) 91.5 **OR** 84 seen   *(A1)*

Note: These values may be seen in the working for part (c).

     (IQR = 91.5 − 84 =) 7.5 (minutes)   *A1*

                                                                                                   [2 marks]

(c) (upper bound =) $91.5 + 1.5 \times 7.5$ **OR** 102.75 seen   *A1*
     $102.75 > 100$ **OR** $100 - 91.5 < 11.25$ **OR** $100 - 11.25 < 91.5$)   *R1*
     Star Feud is not an outlier   *A1*

Note: Do not award *R0A1*.

                                                                                                   [3 marks]

(d) H_0 : The running times of the movies can be modelled by $N(88, 6.75^2)$

     H_1 : The running times of the movies cannot be modelled by $N(88, 6.75^2)$ *A1A1*

Note: Award *A1* for each correct hypothesis that includes a reference to normal distribution with a mean of 88 and a standard deviation of 6.75 (or variance of 6.75^2). "Correlation", "independence", "association", and "relationship" are incorrect.
Award at most *A0A1* for correctly worded hypotheses that include a reference to a normal distribution but omit the distribution's parameters in one or both hypotheses. Award *A0A1* for correct hypotheses that are reversed.

                                                                                                   [2 marks]

(e) (i) $T \sim N(88, 6.75^2)$

attempt to find normal probability in either correct range   (M1)
$P(85 \leq T < 90)$ **OR** $P(T \geq 95)$

recognition of multiplying either of their probabilities by 200   (M1)
$0.288137... \times 200$ **OR** $0.149859... \times 200$

$a = 57.6 \ (57.6274...)$, $b = 30.0 \ (29.9718...)$ **A1A1**

(ii) df $= 4$ **(A1)**
$(p =) \ 0.0166 \ (= 0.0166282...)$ **A1**

comparing their p-value to 0.05 **R1**
$0.0166 < 0.05$

> **Note:** Accept p value of $0.0165 \ (= 0.0164693...)$ from using a and b to 3 sf.

(Reject H_0, There is sufficient evidence to say that) the data has not

been drawn from the ($N(88, 6.75^2)$) distribution.   **A1**

> **Note:** Do not award **R0A1**.
>
> The conclusion to part (e)(ii) **MUST** follow through from their hypotheses seen
> in part (d); if hypotheses are incorrect/reversed etc., the answer to part (e)(ii)
> must reflect this in order for the **A1** to be credited.

[8 marks]
[Total 18 marks]

Question 17

(a) (i) 150 (cm)   *A1*

     (ii) attempt to substitute values in the mean formula with at least one
           mid-interval value multiplied by a corresponding frequency   *(M1)*

           (mean =) 176 (176.3) (cm)   *A1*

                                                                                       [3 marks]

(b) 183 **OR** 168 seen   *(A1)*

> **Note:** These values may be seen in the working for part (c).

     (IQR = 183 − 168 =) 15 (cm)   *A1*

                                                                                       [2 marks]

(c) (upper bound =) $183 + 1.5 \times 15$ **OR** 205.5 seen   *A1*
     $205.5 > 204$ **OR** $204 - 183 < 22.5$ **OR** $204 - 22.5 < 183$ *R1*
     Laszlo's height is not an outlier   *A1*

> **Note:** Do not award **R0A1**.

                                                                                       [3 marks]

(d) H_0: The heights of the students can be modelled by $N(176, 13.5^2)$

     H_1: The heights of the students cannot be modelled by $N(176, 13.5^2)$ *A1A1*

> **Note:** Award **A1** for each correct hypothesis that includes a reference to normal
>   distribution with a mean of 176 and a standard deviation of 13.5 (or variance of
> 13.5^2). "Correlation", "independence", "association", and "relationship" are
>   incorrect.
>
>   Award at most **A0A1** for correctly worded hypotheses that include a reference to a
>   normal distribution but omit the distribution's parameters in one or both hypotheses.
>   Award **A0A1** for correct hypotheses that are reversed.

                                                                                       [2 marks]

(e) (i) $h \sim \mathrm{N}(176, 13.5^2)$

attempt to find normal probability in either correct range **(M1)**
$\mathrm{P}(170 \leq h < 180)$ **OR** $\mathrm{P}(h \geq 190)$

recognition of multiplying either of their probabilities by 200 **(M1)**
$0.288137... \times 200$ **OR** $0.149859... \times 200$

$a = 57.6 \ (57.6274...)$, $b = 30.0 \ (29.9718...)$ **A1A1**

(ii) df $= 4$ **(A1)**
$(p =) \ 0.0166 \ (= 0.0166282...)$ **A1**

comparing their p-value to 0.05 **R1**
$0.0166 < 0.05$

> **Note:** Accept p value of $0.0165 \ (= 0.0164693...)$ from using a and b to 3 sf.

(Reject H_0, There is sufficient evidence to say that) the data has not

been drawn from the ($\mathrm{N}(176, 13.5^2)$) distribution. **A1**

> **Note:** Do not award **R0A1**.
>
> The conclusion to part (e)(ii) **MUST** follow through from their hypotheses seen in part (d); if hypotheses are incorrect/reversed etc., the answer to part (e)(ii) must reflect this in order for the **A1** to be credited.

[8 marks]
[Total 18 marks]

Question 18

(a) (i) $P(X < 160)$ **OR** labelled sketch of region **OR** calc syntax with correct bounds

(M1)

> **Note:** Accept either zero or a large negative value as the lower bound.

$= 0.0447$ $(0.0447149..., 4.47\%)$ **A1**

(ii) $P(160 < X < 170)$ **OR** labelled sketch of region **OR** calc syntax with correct bounds

(M1)

$= 0.411$ $(0.411049..., 41.1\%)$ **A1**

> **Note:** Award **A0A2** for answers of 0.045 and 0.41 both given to 2 sf.

[4 marks]

(b) $P(X > h) = 0.27$ **OR** labelled sketch of region **OR** calc syntax with correct bounds

(M1)

$= 1.75$ (m) $(1.74560...)$ **A1**

> **Note:** Accept 175 (cm).

[2 marks]

(c) (i) 82.21 **A1**

(ii) 94.86 **A1**

> **Note:** Follow through from an incorrect part (c)(i) if fourth value is found by subtracting first three values from 200. Award at most **A0A1** if both answers are not given to four significant figures.
> Award **A0A1** for an answer of $a = 82.2$ and $b = 94.8$.

[2 marks]

(d) 3 **A1**

[1 mark]

(e) p-value $= 0.616$ $(0.615583....)$ **OR** $\chi^2 = 1.80$ $(1.79702....)$ **A2**

> **Note:** Award **A1A0** if the p-value or χ^2-value is given correct to 2 dp.

$0.615583... > 0.05$ **OR** $1.79702.... < 7.815$ **R1**
EITHER
fail to reject the null hypothesis **A1**
OR
the heights are normally distributed with mean 170.7 cm and standard deviation 6.3 cm **A1**

> **Note:** Do not award **R0A1**. Condone "accept" in place of "fail to reject".
>
> The **R1A1** can be awarded as follow through within part (e) from their (explicitly labelled) p-value or χ^2-value. Accept comparison in words.

[4 marks]

(f)   **EITHER**

$H_0 : \mu_L = \mu_D$ *A1*

$H_1 : \mu_L > \mu_D$ *A1*

OR

H_0 : The (population) mean height of Latvian women is equal to the (population) mean height
of Dutch women   *A1*

H_1 : The (population) mean height of Latvian women is greater than the (population) mean
height of Dutch women   *A1*

Note:	Award at most *A0A1* if the hypotheses explicitly refer to the "sample" and not the population. For $H_0 : m_1 = m_2$ and $H_1 : m_1 > m_2$ award *A0A1*.

[2 marks]

(g) p-value $= 0.673$ $(0.673205\ldots)$ *A2*

Note:	In this question the p-value is the same 3 sf value for unpooled GDC settings so will be awarded *A2*.
	If using a two-tailed test, the answer is p-value $=0.654$ $(0.653589\ldots)$; award *A1* if alternative hypothesis was correct or *A2* if it follows through correctly from their alternative hypothesis (i.e. two-tailed test was penalized in part (f)).

$0.673205 > 0.05$ *R1*

fail to reject the null hypothesis (Gundega is not correct)   *A1*

Note:	Do not award *R0A1*. Condone "accept" in place of "fail to reject".
	The *R1A1* can be awarded as follow through within part (g) from their (explicitly labelled) p-value. Accept comparison in words.

[4 marks]

[Total: 19 marks]

Question 19

1.   (a)   (i)   0.9   (ii)   0.3   (iii)   0.7   **A2**

> **Note:** Award **A1A0** if one of the values is incorrect, **A0A0** otherwise.

[2 marks]

       (b) $(0.1 \times 0.3 =)$ 0.03   **A1**

[1 mark]

       (c) $P(\text{no fail}) = 0.63$ **A1**

                multiplying by 200   **M1**

                $= 126$ **AG**

> **Note:** Award **A0M0** for a flawed approach to find $P(\text{no fail}) = 0.63$, e.g. $\dfrac{126}{200} = 0.63$, which is reverse engineering.

[2 marks]

       (d)   **EITHER**

                attempt to find probability one switch failing   **(M1)**

                $P(\text{one failing}) = 0.34$ **(A1)**

                OR

                expected value for two switches failing $= 6$ **(A1)**

                expected value for one switch failing $= 200 - 126 - 6$ **(M1)**

                THEN

No switch fails	One switch fails	Two switches fail
126	68	6

                                                                      (A1)

                degrees of freedom $= 2$ **(A1)**

> **Note:** Award **A1** for df = 2 seen anywhere and may be awarded independent of the **M1** mark. The df cannot be implied from chi sq statistic = 3.40989

                p-value 0.182 $(0.181781\ldots)$ **A1**

                $0.182 > 0.05$ **R1**

                hence insufficient evidence to reject H_0 (that the manufacturers claims are correct) **A1**

[7 marks]

[Total 12 marks]

Question 20

(a) 0.5   **A1**
                                                                        [1 mark]

(b) 0.452 (0.452209…)   **A2**
                                                                        [2 marks]
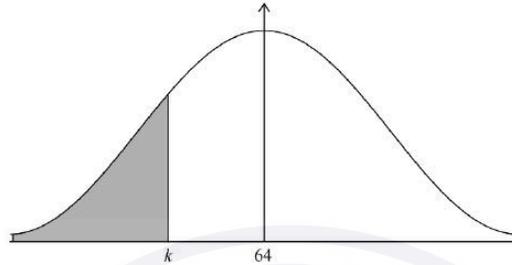
(c) (i)

                                                                        A1A1

ote: Award **A1** for a normal curve (with symmetry and some evidence of change of curvature towards the extreme values).
Award **A1** for a shaded region $x < k$, where $k<$mean.

(ii) $P(T < k) = 0.3$
solving a cumulative distribution function **OR**
        use of inverse function on GDC   **(M1)**
        $k = 57.7$ (57.7071…)   **A1**
                                                                        [4 marks]

(d) recognizing binomial distribution   **(M1)**
     $B(5, 0.3)$ $(P(X = 2))$ **(A1)**

     0.309 (0.3087)   **A1**
                                                                        [3 marks]

(e) $2(x-1)+4.5$ **OR** $2x+2.5$ **A1A1**

---

te: Award **A1** for a linear expression with a gradient of 2,
A1 for a completely correct expression in x.

                                                                        [2 marks]

(f) ($)13.10 (accept 13.1)   **A1**
                                                                        [1 mark]

(g)   attempt to solve $2(x-1)+4.5 = 7.2$ **OR** $2x+2.5 = 7.2$ **(M1)**
      2.35 (kg)   **A1**

---

te: Award **M1A1FT** for an answer of 1.35 (kg) from $2x+4.5$ seen in (e).

                                                                        [2 marks]
                                                                  [Total 15 marks]

Question 21

.   (a) $r_s = -1$ **A1**

   [1 mark]

   (b)   (i) $r = -0.979$ $(-0.979191\ldots)$ **A2**

         (ii)   strong **AND** negative   **A1A1**

Note: Award at most **A1A0** if additional answers are seen.
   Due to the demand of the question, do not accept "negative (from the graph)" if their r value is
   positive.

   [4 marks]

   (c)   (i) $a = -0.0992$ $(a = -0.0992075\ldots)$ **A1**

         (ii) $b = 3.19$ $(b = 3.19150\ldots)$ **A1**

         (iii) b represents the (typical) price of an apartment in the centre (of the city)   **A1**

Note: To award the **A1**, some reference to "centre" or "zero distance from the city" needs to be seen.

   [3 marks]

   (d)   (i)   attempt to substitute 19.6 for x **(M1)**
         $y = -0.0992075\ldots \times 19.6 + 3.19150\ldots$
         $= 1.25$ $(1.24704\ldots)$ **A1**
         price $= 1.25$ million (AUD) $(1.24704\ldots$ million)   **A1**

         (ii)   interpolation   **R1**
               strong correlation.   **R1**
   [5 marks]

   (e) $\mu_A \neq \mu_B$ **A1**

   [1 mark]

   (f) $p = 0.0224$ $(0.0223977\ldots)$ **A2**

Note: Award **A1** for 0.022 (2sf)
   Award **A1** for an answer of $p = 0.0265$ $(0.0265017\ldots)$, from use of unpooled GDC settings.

   [2 marks]

   (g) $0.0223977\ldots < 0.05$ **R1**
         (there is sufficient evidence to) reject the null hypothesis   **A1**

lote: Do not award **R0A1**.

   [2 marks]

   (h)   (the two populations are) normally distributed   **A1**

lote: Do not accept "independent" as that applies to the samples, not the populations.

   [1 mark]

   [Total 19 marks]

Question 22

(a) (i) $(Q_3 =)\ 7.2$ **A2**

> **Note:** Award **A1A0** for a correct ordered list of happiness scores, when the correct Q_3 is not seen.

(ii) $Q_1 = 6.5$ **(A1)**
$IQR = 7.2 - 6.5$
$= 0.7$ **A1**
[4 marks]

(b) $Q_3 + 1.5 \times IQR$ **(A1)**
$(7.2 + 1.5 \times 0.7 =)\ 8.25$ **A1**
since $7.5 < 8.25$ **R1**
Switzerland is not an outlier **AG**

> **Note:** Do not award **A0A0R1**.

[3 marks]

(c) (i) $a = 3.5$ **A1**

(ii) $b = 8$ **A1**

(iii) $c = 3.5$ **A1**
[3 marks]

(d) (i) $(r_s =)\ 0.164\ (0.164134\ldots)$ **A2**

(ii) France rank (of sixth) is unchanged (so the r_s is unchanged) **R1**
[3 marks]

(e) Because r_s is too close to zero and hence Jose's conclusion is not appropriate **R1**

> **Note:** Award **R1** for a comment on the value r_s and "not appropriate" as a conclusion.
> Accept "r_s indicates a weak correlation of the ranks".
> Do not accept "r_s indicates a weak correlation of happiness score and country rank".

[1 mark]
[Total 14 marks]

Question 23

(a) (2, 2) **A1**

> **Note:** Award **A0** if paretheses are omitted.

[1 mark]

(b) (i) attempt to use midpoint formula (at least one correct) **(M1)**

$$\left(\frac{9+10}{2}, \frac{0+6}{2}\right)$$

(9.5, 3) **A1**

(ii) $\left(m = \dfrac{6-0}{10-9} = \right)6$ **(A1)**

finding negative reciprocal of their gradient **(M1)**

$$m_\perp = -\frac{1}{6}$$

attempt to substitute their midpoint and their gradient into equation of straight line **(M1)**

eg. $y - 3 = -\dfrac{1}{6}\left(x - \dfrac{19}{2}\right)$ **OR** $3 = -\dfrac{1}{6}(9.5) + c$

$y = -\dfrac{1}{6}x + \dfrac{55}{12}$ **OR** $y = -0.167x + 4.58$ $(y = -0.166666...x + 4.58333...)$ **A1**

> **Note:** Substituting $m = 6$ (no negative reciprocal) and their point into the equation of straight line would receive at most **A1M0M1A0**.

[6 marks]

(c) (closest to B so) 5 **A1**
[1 mark]

(d) **EITHER**
$(\mathrm{H}_1 :)\ \mu_1 < \mu_2$ **A1**

OR
mean amount of money spent by all customers on weekdays, is less than the **mean** amount of money spent by all customers on weekends **A1**

> **Note:** Award **A0** for "sample mean" or if 'mean' is omitted. Eg "money spent on weekdays, is less than the money spent on weekends".

[1 mark]

(e)   (i)   convenience   *A1*

       (ii)   not representative of the whole population   *A1*

              Do not accept the following reasons
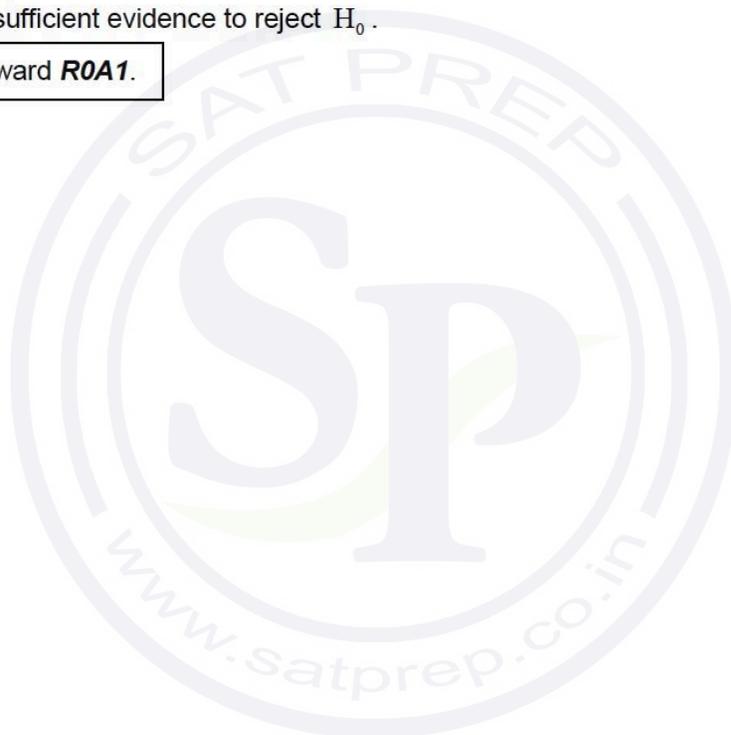   - disadvantages of other sampling methods
              - there are five days during the week, only two on a weekend
              - anything based on speculation/information not presented in the question
              - people spend different amounts of money
              - this is a sample (doesn't collect every data point) more data should have been collected

                                                                                                               [2 marks]

(f) $p > 0.05 \ (\alpha)$ **OR** $0.0897\ldots > 0.05$ *R1*
       There is insufficient evidence to reject H_0.   *A1*

       Note: Do not award *R0A1*.

                                                                                                               [2 marks]
                                                                                                       [Total 13 marks]

Question 24

(a) (i) $(Q_3 =)\ 7.2$ **A2**

> **Note:** Award **A1A0** for a correct ordered list of happiness scores, when the correct Q_3 is not seen.

    (ii) $Q_1 = 6.5$ **(A1)**

        $IQR = 7.2 - 6.5$

        $= 0.7$ **A1**

        [4 marks]

(b) $Q_3 + 1.5 \times IQR$ **(A1)**

    $(7.2 + 1.5 \times 0.7 =)\ 8.25$ **A1**

    since $7.5 < 8.25$ **R1**

    Switzerland is not an outlier **AG**

> **Note:** Do not award **A0A0R1**.

    [3 marks]

(c) (i) $a = 3.5$ **A1**

    (ii) $b = 8$ **A1**

    (iii) $c = 3.5$ **A1**

        [3 marks]

(d) (i) $(r_s =)\ 0.164\ (0.164134\ldots)$ **A2**

    (ii) France rank (of sixth) is unchanged (so the r_s is unchanged) **R1**

        [3 marks]

(e) Because r_s is too close to zero and hence Jose's conclusion is not appropriate **R1**

> **Note:** Award **R1** for a comment on the value r_s and "not appropriate" as a conclusion.
> Accept "r_s indicates a weak correlation of the ranks".
> Do not accept "r_s indicates a weak correlation of happiness score and country rank".

[1 mark]
[Total 14 marks]

Question 25

(a) $(2, 6)$ **A1**

Note: Award **A0** if parentheses are omitted.

[1 mark]

(b)   (i)   attempt to use midpoint formula (at least one correct)   **(M1)**

$$\left(\frac{11+12}{2}, \frac{0+6}{2}\right)$$

$(11.5, 3)$ **A1**

(ii) $\left(m = \dfrac{6-0}{12-11} = \right) 6$ **(A1)**

finding negative reciprocal of their gradient   **(M1)**

$$m_{\perp} = -\frac{1}{6}$$

attempt to substitute their midpoint and their gradient into equation of straight line   **(M1)**

eg. $y - 3 = -\dfrac{1}{6}\left(x - \dfrac{23}{2}\right)$ **OR** $3 = -\dfrac{1}{6}(11.5) + c$

$y = -0.167x + 4.92\left(y = -0.166666...x + 4.91666..., y = -\dfrac{1}{6}x + \dfrac{59}{12}\right)$ **A1**

Note: Substituting $m = 6$ (no negative reciprocal) and their point into the equation of straight line would receive at most **A1M0M1A0**.

[6 marks]

(c)   (closest to Q so) 4   **A1**
                                            [1 mark]

(d)   **EITHER**

$.(H_1 :) \; \mu_1 < \mu_2 .$ **A1**

OR
mean amount of money spent by all customers on weekdays is less than the
mean amount of money spent by all customers on weekends   **A1**

Note: Award **A0** for 'sample mean' or if 'mean" is omitted. For example, 'money spent on weekdays is less than the money spent on weekends'.

[1 mark]

(e) (i)   convenience   *A1*

     (ii) not representative of the whole population   *A1*

           Do not accept the following reasons
           - disadvantages of other sampling methods
           - there are five days during the week, only two on a weekend
           - anything based on speculation/information not presented in the question
           - people spend different amounts of money
           - this is a sample (doesn't collect every data point) more data should have
             been collected

                                                                               [2 marks]

(f) $p > 0.05 \ (\alpha)$ **OR** $0.0897.... > 0.05$ *R1*

       There is insufficient evidence to reject H_0 *A1*

       Note: Do not award *R0A1*.

                                                                               [2 marks]
                                                                          [Total 13 marks]