MA4413 2015 Solutions.

Question 1

(a)(i) $\boxed{1}$

$$Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$$
$$= 0.7 + 0.6 - 0.5$$
$$= 0.8 \xrightarrow{\text{correct}} 1$$

(ii) $\boxed{2}$

$$Pr(A^c \cup B^c) = 1 - Pr((A^c \cup B^c)^c)$$
$$= 1 - Pr(A \cap B) \xrightarrow{\text{Complement rule}} 1$$
$$= 1 - 0.5$$
$$= 0.5 \xrightarrow{\text{correct}} 1$$

(iii) $\boxed{2}$

$$Pr(A|B) = \underbrace{\frac{Pr(A \cap B)}{Pr(B)}}_{\text{formula } 1} = \frac{0.5}{0.6}$$
$$= 0.833 \xrightarrow{\text{correct}} 1$$

(b)(i) $\boxed{3}$

Position	1	2	3	4	5	6	7	8	9	10	11	12
Value	17	19	24	24	24	26	29	32	32	33	34	34

Position of $Q1 = \frac{n+1}{4} = \frac{12+1}{4} = \frac{13}{4} = 3.25 \Rightarrow$ between 3 & 4

$Q2 = 2\left(\frac{n+1}{4}\right) = 2(3.25) = 6.5 \Rightarrow$ between 6 & 7

$Q3 = 3\left(\frac{n+1}{4}\right) = 3(3.25) = 9.75 \Rightarrow$ between 9 & 10.

$$\left.\begin{aligned} \Rightarrow Q1 &= \frac{24+24}{2} = 24 \\ Q2 &= \frac{26+29}{2} = 27.5 \\ Q3 &= \frac{32+33}{2} = 32.5 \end{aligned}\right\} 3$$

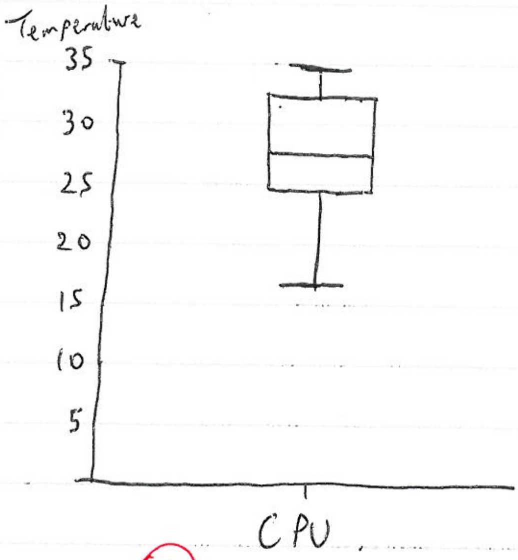
(ii) $\boxed{2}$

$IQR = Q3 - Q1 = 32.5 - 24 = 8.5$

$$\left.\begin{aligned} UF &= Q3 + 1.5\,IQR = 32.5 + 1.5(8.5) = 45.25 \\ LF &= Q1 - 1.5\,IQR = 24 - 1.5(8.5) = 11.25. \end{aligned}\right] \text{correct } 1$$

No values above UF or below LF $\Rightarrow$ no outliers $\rightarrow$ 1

(ii)
$\boxed{3}$



min/max correct : 1
Box correct (i.e., Q1, Q2, Q3) : 1
Overall layout (neatness) : 1

$\boxed{2}$ (iv) Ho - most values are above 22°. 1 → 1

justification

(c) (i) $\boxed{2}$ μ = unknown → 1
→ 1

(ii) $\boxed{2}$

$$\left.\begin{array}{l} H_0: \quad \mu = 2 \\ H_a: \quad \mu \neq 2 \end{array}\right\} 2$$

(iii) $\boxed{3}$ P-value = 0.001195 is very small (less than 0.01) → 1

⇒ there is strong evidence to reject Ho. → 1

i.e., it appears that $\mu \neq 2$. → 1

(iv) $\boxed{3}$ We are 95% confident that μ is in the range [2.36, 3.11]. 1

interpret C.I.

This does not support the possibility that $\mu = 2$ since $\mu = 2$ is not in the C.I., i.e., we reject Ho. 1

Reject Ho

1

since $\mu = 2$ not in C.I.

Question 2

(a) (i) Class width $= \dfrac{\text{range}}{\#\text{ of classes}} = \dfrac{16}{7} = 2.28 \Rightarrow \underline{3} \rightarrow 1$

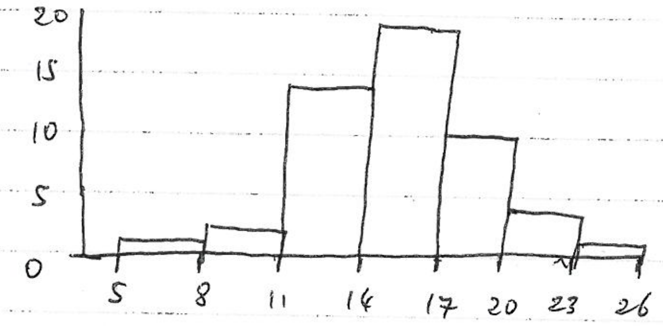
$\boxed{4}$

Class	Frequency
5 – 7.9	1
8 – 10.9	2
11 – 13.9	14
14 – 16.9	18
17 – 19.9	10
20 – 22.9	4
23 – 25.9	1
	50

Classes correct 2

frequencies correct 1

(ii) $\boxed{3}$



Layout : 1

Agrees with table in part (i) : 1

Correct : 1

(iii) $\boxed{2}$ <u>Symmetric</u> → 1 $\Rightarrow$ the <u>mean</u> → 1 is an appropriate measure of centrality.

(b) (i) $\boxed{2}$ Standard deviation : a measure of spread around the mean → 1

IQR : the range of the middle 50% of data. → 1

(ii) $\boxed{2}$ When the data is <u>skewed</u>. → 2

(iii) $\boxed{2}$ It is usually <u>not feasible (or possible) to access the whole population</u>. → 1
Thus, we collect a sample of data and calculate a statistic (e.g. $\bar{x}$ or $\hat{p}$) to <u>estimate</u> a <u>parameter</u> (μ or p). → 1

(c) (i) $\boxed{4}$ $P_r(R_1) = 0.3$ $\quad P_r(L \mid R_1) = 0.15$

$P_r(R_2) = 0.7$ $\quad P_r(L \mid R_2) = 0.04$

$\Rightarrow P_r(L \cap R_1) = P_r(R_1) P_r(L \mid R_1) \xrightarrow{\text{formula}} 1$

$= 0.3(0.15) = 0.045 \rightarrow 1$ answer

$P_r(L \cap R_2) = P_r(R_2) P_r(L \mid R_2) \xrightarrow{\text{formula}} 1$

$= 0.7(0.04) = 0.028 \rightarrow 1$ answer

(ii) $\boxed{2}$ $P_r(L) = P_r(L \cap R_1) + P_r(L \cap R_2) \longrightarrow$ formula 1

$= 0.045 + 0.028$

$= 0.073 \xrightarrow{\text{answer}} 1$

(iii) $\boxed{4}$ $P_r(R_1 \mid L^c) = \dfrac{P_r(R_1 \cap L^c)}{P_r(L^c)} \xrightarrow{\text{formula}} 1$

$= \dfrac{P_r(R_1) P_r(L^c \mid R_1)}{P_r(L^c)}$

$= \dfrac{P_r(R_1)[1 - P_r(L \mid R_1)]}{1 - P_r(L)}$

$\rightarrow$ using complement rule 1

$= \dfrac{0.3(1 - 0.15)}{1 - 0.073} \rightarrow$ complement rule 1

$= 0.2751 \rightarrow 1$ correct

Question 3

(a) (i) $\boxed{1}$ $\bar{x} = \frac{\sum x}{n} = \frac{5+2+2+3+1+3}{6} = \frac{16}{6} = 2.66667$. Correct 1

(ii) $\boxed{3}$ $\sum x^2 = 5^2 + 2^2 + 2^2 + 3^2 + 1^2 + 3^2$
$= 25 + 4 + 4 + 9 + 1 + 9 = 52$ 1

$$\Rightarrow s^2 = \frac{\sum x^2 - n\bar{x}^2}{n-1} = \frac{52 - 6(2.66667^2)}{5}$$

$$= \frac{9.333}{5} = 1.86667$$ correct 1

$$\Rightarrow s = \sqrt{s^2} = \sqrt{1.86667} = 1.36626$$ 1

Square root of previous answer.

(b) (i) $\boxed{1}$ Categorical ("Yes" / "No") 1

(ii) $\boxed{2}$ p 1 = unknown 1

(iii) $\boxed{2}$ $\hat{p}$ 1 $= \frac{50}{168} = 0.2976$ 1

(iv) $\boxed{4}$ $\hat{p} \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ where $n = 168$

formula 1

$\alpha = 0.01 \Rightarrow \alpha/2 = 0.005$

$\Rightarrow Z_{0.005} = 2.58$ 1

$$0.2976 \pm 2.58\sqrt{\frac{0.2976(0.7024)}{168}}$$

$$0.2976 \pm 0.091$$

$\underbrace{[0.2066, 0.3886]}_{\text{Correct. } 1}$

This does not support the researcher's belief that $p = 0.4$ since $p = 0.4$ is not contained in the interval. 1 Conclusion.

(v) $\boxed{3}$ We require $\hat{p} \pm 0.03$

i.e., $2.58\sqrt{\dfrac{p(1-p)}{n}} = 0.03$ → setting equal to 0.03 (1)

using $\hat{p}$ in place of p we have

$$2.58\sqrt{\frac{0.2976(0.7024)}{n}} = 0.03$$

$$\sqrt{\frac{0.2976(0.7024)}{n}} = \frac{0.03}{2.58}$$ workings → (1)

$$\frac{1}{n} = \left(\frac{0.03}{2.58}\right)^2 \frac{1}{0.2976(0.7024)}$$

$$\Rightarrow n = \left(\frac{2.58}{0.03}\right)^2 0.2976(0.7024)$$

$$\simeq 1546.$$ → (1)

Correct, i.e., ≈ 1500.

(c)(i) $\boxed{2}$

$$\left.\begin{aligned} H_0 &: \mu_1 - \mu_2 = 0 \\ H_a &: \mu_1 - \mu_2 \neq 0. \end{aligned}\right\} (2)$$

(ii) $\boxed{4}$

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{83.1 - 80.1}{\sqrt{\frac{30.6}{40} + \frac{18.5}{40}}}$$

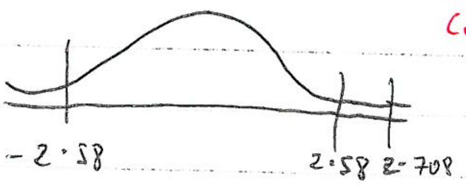
formula (1)

$$= \frac{3}{1.108} = 2.708$$ → correct (1)

Two-tailed test $\Rightarrow$ critical values are $\pm Z_{\alpha/2} = \pm Z_{0.005} = \pm 2.58$ (since $\alpha = 0.01$, $\alpha/2 = 0.005$).

$0.01 \div 2$ (1) Correct (1)

(iii) $\boxed{3}$



Compare to critical value → (1)

- 2.708 is outside of ± 2.58 $\Rightarrow$ we reject H_0, i.e., it appears $\mu_1 \neq \mu_2$.

Conclusion (1)

- The hours of gameplay are not equal, "Game 1" has more gameplay hours.

plain English → (1)

Question 4

(a) (i) $k = 1-(0.1+0.4+0.2) = 1-0.7 = 0.3$ 1

$\boxed{1}$

(ii) $EX = 0(0.1) + 3(0.4) + 6(0.3) + 9(0.2)$ $\xrightarrow{\text{multiplying } x \cdot p(x)}$ 1

$\boxed{2}$

$= 0 + 1.2 + 1.8 + 1.8$

$= 4.8$ $\xrightarrow{\text{correct}}$ 1

(iii) $EX^2 = 0^2(0.1) + 3^2(0.4) + 6^2(0.3) + 9^2(0.2)$

$\boxed{2}$

$= 30.6$

$\text{Var}X = EX^2 - (EX)^2$ $\xrightarrow{\text{formula}}$ 1

$= 30.6 - (4.8)^2 = 7.56$

$\text{Sd}X = \sqrt{\text{Var}X} = \sqrt{7.56} = 2.749$ $\xrightarrow{\text{correct}}$ 1

(b) (i) $EX = np$ $\xrightarrow{\text{formula}}$ 1 $= 80(0.04) = 3.2$ $\xrightarrow{\text{correct}}$ 1

$\boxed{2}$

(ii) $n = 15 \Rightarrow p(x) = \binom{n}{x} p^x (1-p)^{n-x} = \binom{15}{x} 0.04^x \cdot 0.96^{15-x}$

$\boxed{3}$

$P_r(2 \leq X \leq 5) = p(2) + p(3) + p(4) + p(5)$ $\xrightarrow{\text{summing probabilities}}$ 1

$= 0.0988 + 0.0178 + 0.0022 + 0.0002$ $\longrightarrow$ 2

$= 0.119$

(iii) $P_r(X > 8) = P_r(X \geq 9)$ $\longrightarrow$ 1 $= 0.0190$ $\longrightarrow$ 2 (from tables)

$\boxed{3}$

$(n=100, p=0.04, r=9)$.

(iv) The binomial distribution arises as a sequence of independent $\longrightarrow$ 2

$\boxed{3}$

Bernoulli trials. (If the disease is contagious then the occurrence of the disease is not independent as one person can pass it on to another. Similarly, individuals in the same family will be more alike and, hence, not independent.)

$\longrightarrow$ 1 something along these lines.

(c) (i) $\lambda = 7/hr$.

$\boxed{3}$ $\Rightarrow \lambda = 7(\frac{1}{2}) = 3.5 \,/\, \frac{1}{2}hr.$ $\Rightarrow p(x) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{3.5^x e^{-\lambda}}{x!}$

(1)

$$\begin{aligned} Pr(X \geq 3) &= 1 - Pr(X < 3) \xrightarrow{\text{Complement}} (1) \\ &= 1 - [p(0) + p(1) + p(2)] \\ &= 1 - [0.0302 + 0.1057 + 0.1850] \\ &= 1 - 0.3209 = 0.6791 \xrightarrow{\text{correct}} (1) \end{aligned}$$

This replaces the above workings if using tables, i.e., 2 marks.

[Or, using, tables $Pr(X \geq 3) = 0.6792$ (m = 3.5, r = 3)]

(ii) $\lambda = 7(3) = 21 \,/\, 3\text{-hrs}.$ (1)

$\boxed{3}$

$$\begin{aligned} Pr(15 \leq X \leq 25) &= Pr(X \geq 15) - Pr(X \geq 26) \rightarrow (1) \\ &= \underset{(m=21,\ r=15)}{0.9284} - \underset{(m=21,\ r=26)}{0.1623} \\ &= 0.7661 \xrightarrow{\text{Correct}} (1) \end{aligned}$$

(iii) $T \sim \text{Exponential}(\lambda = 7)$

$\boxed{3}$

$$\begin{aligned} Pr(T < \tfrac{5}{60}) &= 1 - Pr(T > \tfrac{5}{60}) = 1 - e^{-7(\frac{5}{60})} \\ &= 1 - 0.558 \\ &= 0.442 \end{aligned}$$

Since we are working in hours

$\div 60$ (1)

Complement (1)

Correct (1)

Question 5

(a) (i) $\boxed{3}$

x	a	b	c	d
$p(x)$	0.4	0.1	0.35	0.15
$h(x)$	1.323	3.323	1.515	2.737

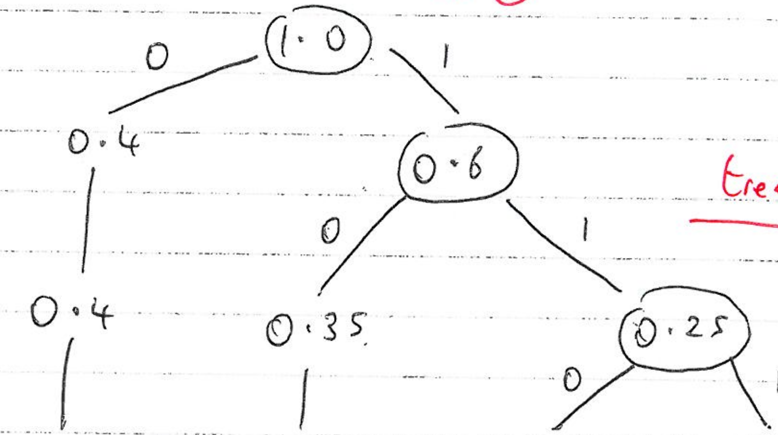
correct information content values → 1

$$H(X) = E(h(X))$$
$$= 1.323(0.4) + 3.323(0.1) + 1.515(0.35) + 2.737(0.15)$$
$$= 1.8018 \text{ bits}$$

correct → 1

(multiplying $h \times p$ and summing) → 1

(ii) $\boxed{3}$



tree → 1

$P(x)$	0.4	0.35	0.15	0.1
x	a	c	d	b
$C(x)$	0	10	110	111
$\ell(x)$	1	2	3	3

sorting → 1

correct codes → 1

(iii) $\boxed{4}$

$$E(L) = 1(0.4) + 2(0.35) + 3(0.15) + 3(0.1)$$
$$= 0.4 + 0.7 + 0.45 + 0.3$$
$$= 1.85$$

correct → 1

multiplying $\ell \times p$ and summing → 1

$$\Rightarrow \text{efficiency} = \frac{H(X)}{E(L)} = \frac{1.8018}{1.85} = 0.97$$

correct. → 1

formula → 1

i.e., 97%.

(b) (i) $\boxed{3}$

$$Pr(X < 25) = Pr\left(Z < \frac{25-20}{3}\right)$$
$$= Pr(Z < 1.67) \xrightarrow{\text{z-score}} 1$$
$$= 1 - Pr(Z > 1.67) \xrightarrow{\text{complement}} 1$$
$$= 1 - 0.0475$$
$$= 0.9525. \xrightarrow{\text{Correct}} 1$$

(ii) $\boxed{3}$

$$Pr(23.5 < X < 28.4) = Pr(X > 23.5) - Pr(X > 28.4) \longrightarrow 1$$
$$= Pr\left(Z > \frac{23.5-20}{3}\right) - Pr\left(Z > \frac{28.4-20}{3}\right)$$
$$= Pr(Z > 1.17) - Pr(Z > 2.8) \longrightarrow 1$$
$$= 0.1210 - 0.00256$$
$$= 0.11844 \xrightarrow{\text{Correct}} 1$$

(iii) $\boxed{3}$

$$Pr(X > x) = 0.35$$
$$\Rightarrow Pr\left(Z > \frac{x-20}{3}\right) = 0.35$$

But $Pr(Z > 0.38) = 0.3483 \simeq 0.35$

(0.38, 0.385 or 0.39) $\longrightarrow$ 1

$$\Rightarrow \frac{x-20}{3} = 0.39 \quad 1 \quad \Rightarrow x = 20 + 0.39(3)$$
$$= 21.17 \xrightarrow{\text{Correct}} 1$$

(iv) $\boxed{3}$

$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) = N\left(20, \frac{3}{\sqrt{45}} = 0.4472\right) \longrightarrow 1$$

$$\Rightarrow Pr(\bar{X} > 20.8) = Pr\left(Z > \frac{20.8-20}{0.4472}\right)$$
$$= Pr(Z > 1.79) = 0.0367 \xrightarrow{\text{correct}} 1$$

Correct $\rightarrow$ 1

(v) $\boxed{3}$

$$X_1 + X_2 \sim N\left(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2}\right) = N(40, \sqrt{18} = 4.2426).$$

1 1

$$Pr(X_1 + X_2 > 45.7) = Pr\left(Z > \frac{45.7-40}{4.2426}\right)$$
$$= Pr(Z > 1.34) = 0.0901$$

Correct $\rightarrow$ 1

Question 6

(a) (i) $H_0: \mu_1 - \mu_2 = 0$
$\boxed{2}$ $H_a: \mu_1 - \mu_2 \neq 0$ } 2

(ii) $n_1 = 8, \quad \bar{x}_1 = 7.96, \quad s_1 = 0.73$
$\boxed{5}$ $n_2 = 7, \quad \bar{x}_2 = 6.83, \quad s_2 = 2.36$

95% C.I. $\Rightarrow \alpha = 0.05 \Rightarrow \alpha/2 = 0.025$

Two small samples, unequal variance $\Rightarrow t_{v, \alpha/2}$

where $v = \dfrac{(a+b)^2}{\frac{a^2}{n_1-1} + \frac{b^2}{n_2-1}}$ formula → 1 $\left[\begin{array}{l} a = \frac{s_1^2}{n_1} = \frac{0.73^2}{8} = 0.0666 \\ b = \frac{s_2^2}{n_2} = \frac{2.36^2}{7} = 0.7957 \end{array} \right]$

$= \dfrac{(0.0666 + 0.7957)^2}{\frac{0.0666^2}{7} + \frac{0.7957^2}{6}} = 7.004 \simeq 7.$ → 1 Correct degrees of freedom

t-tables only have whole numbers.

$(\bar{x}_1 - \bar{x}_2) \pm t_{7, 0.025} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

$(7.96 + 6.83) \pm 2.365 \sqrt{\frac{0.73^2}{8} + \frac{2.36^2}{7}}$ → 1

$1.13 \pm 2.365(0.4286)$ workings → 1

1.13 ± 2.196

$[-1.066, 3.326]$ Correct → 1

(iii) The interval includes $\mu_1 - \mu_2 = 0$ which supports H_0, i.e., no difference between means.
$\boxed{3}$ It appears that customers spend equal amounts using both website designs.

reason → 1 conclusion → 1 Plain English → 1 (b)(i) $\boxed{2}$ $\lambda_s = \frac{1}{E(T_s)} = \frac{1}{0.05} = 20$ cust/hr

reciprocal (1)

Correct (1)

(ii) $\boxed{2}$ $\rho = \frac{\lambda_a}{\lambda_s} = \frac{15}{20} = 0.75$ correct (1)

$\Rightarrow$ The service node is in use 75% of the time and idle 25% of the time. interpretation (1)

(iii) $\boxed{2}$ $T \sim Exp(\lambda)$ where $\lambda = \lambda_s - \lambda_a = 20 - 15 = 5$.

$\Rightarrow E(T) = \frac{1}{\lambda} = \frac{1}{5}$ hours $= \frac{1}{5}(60) = 12$ minutes. hours or minutes (1)

$E(T_q) = E(T) - E(T_s) = \frac{1}{5} - 0.05$

$= 0.2 - 0.05$

$= 0.15$ hours hours or minutes (1)

$= 0.15(60) = 9$ minutes

(iv) $\boxed{3}$ $E(N) = \lambda_a E(T) = 15(0.2) = 3$ customers correct (1)

use of formula (1)

$E(N_q) = \lambda_a E(T_q) = 15(0.15) = 2.25$ customers correct (1)

(v) $\boxed{3}$ $P_r(T > \frac{45}{60}) = P_r(T > 0.75) = e^{-5(0.75)}$ use of formula (1)

convert to hours (1)

$= 0.0235$ correct (1)

(vi) $\boxed{3}$ $ET = \frac{1}{\lambda} = \frac{1}{\lambda_s - \lambda_a} = \frac{1}{\lambda_s - 15} = \frac{5}{60}$ (5 mins in hours).

setting $\frac{1}{\lambda_s - 15}$ = to anything! (1)

$\Rightarrow \lambda_s - 15 = \frac{60}{5} = 12$

$\lambda_s = 15 + 12 = 27$ cust/hr

workings (1)

$EN = \lambda_a ET = 15(\frac{5}{60}) = 1.25$ customers.

correct (1)