Teacher Support Materials

Maths GCE

Paper Reference SS03

The Assessment and Qualifications Alliance (AQA) is a company limited by guarantee registered in England and Wales (company number 3644723) and a registered charity (registered charity number 1073334). Registered address: AQA, Devas Street, Manchester M15 6EX.
Dr Michael Cresswell, Director General.

Question 1

1 A manufacturer of digital radios seeks the opinions of customers about the performance of its radios before and after introducing a new component.

The manufacturer selects, at random, 10 customers. Each customer is given a radio without the new component and a radio with the new component. Each customer then rates the performance of each radio on a scale from 1 to 20.

The results of the survey are shown in the table.

Customer	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
Without new component	16	18	14	18	15	13	16	8	8	13
With new component	12	16	19	18	11	19	19	15	17	12

Carry out a Wilcoxon signed-rank test, at the 5% significance level, to investigate whether customers rated the radio differently after the new component was introduced.

Interpret your conclusion in context. *(9 marks)*

Student Response

Question number

Leave blank

1\.

H_0: population mean difference $= 0$

H_1: population mean diffence $\neq 0$, 5% 2 tail ✓

	Diffence (w/o-w)	Sign +	−
A	4	4.5	
B	2	2	
C	−5		6
D	0	discarded	
E	4	4.5	
F	−6		7
G	−3	✓	3
H	−7		8
I	−9		9 ✓
J	1	1	

$T^+ = 12$ $T^- = 33$ ✓ ✓

ts = 12

n = 9, cv = 6 ✓

ts > cv

accept H_0 ✓

no significant evidence supperting diffence in population mean

No evidence suggests that customers rated the radio differently after the new component was introduced ✓ bod

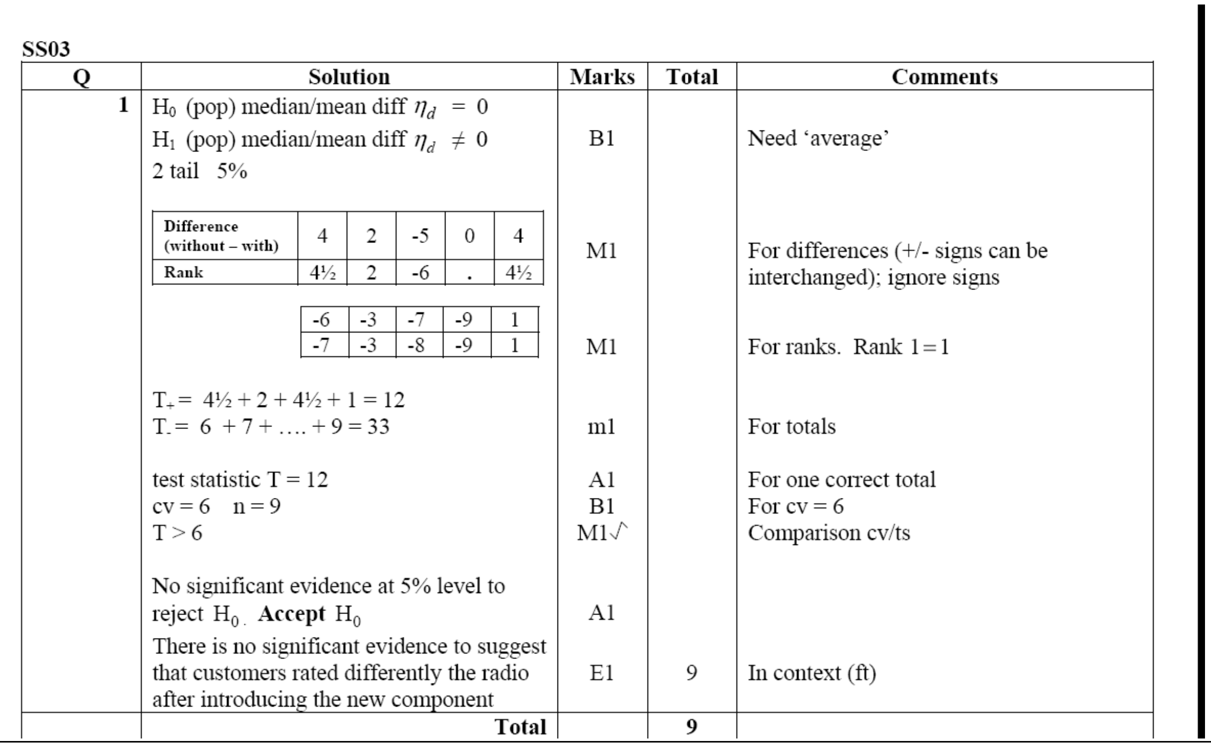
9

9

Commentary

Some candidates seemed to expect a sign test for question 1 and so carried out an incorrect procedure. Many candidates made a very good effort at this question and the majority showed the differences between each value and 6 as well as the rank values. Several ranked 0 with rank 1 or ranked -9 with rank 1. A difference of 0 should be excluded and the difference with the smallest absolute value is assigned rank 1.
A substantial number of candidates carried out a 1 tail test in error.
Conclusions were generally fairly well done and in context.

Mark Scheme



SS03

<table>
<tr><th>Q</th><th>Solution</th><th>Marks</th><th>Total</th><th>Comments</th></tr>
<tr><td>1</td><td>H_0 (pop) median/mean diff $\eta_d = 0$
H_1 (pop) median/mean diff $\eta_d \neq 0$
2 tail 5%</td><td>B1</td><td></td><td>Need ‘average’</td></tr>
<tr><td></td><td><table><tr><td>Difference (without – with)</td><td>4</td><td>2</td><td>-5</td><td>0</td><td>4</td></tr><tr><td>Rank</td><td>4½</td><td>2</td><td>-6</td><td>.</td><td>4½</td></tr></table></td><td>M1</td><td></td><td>For differences (+/- signs can be interchanged); ignore signs</td></tr>
<tr><td></td><td><table><tr><td>-6</td><td>-3</td><td>-7</td><td>-9</td><td>1</td></tr><tr><td>-7</td><td>-3</td><td>-8</td><td>-9</td><td>1</td></tr></table></td><td>M1</td><td></td><td>For ranks. Rank 1 = 1</td></tr>
<tr><td></td><td>$T_+ = 4½ + 2 + 4½ + 1 = 12$
$T_- = 6 + 7 + …. + 9 = 33$</td><td>m1</td><td></td><td>For totals</td></tr>
<tr><td></td><td>test statistic T = 12
cv = 6 n = 9
T > 6</td><td>A1
B1
M1√</td><td></td><td>For one correct total
For cv = 6
Comparison cv/ts</td></tr>
<tr><td></td><td>No significant evidence at 5% level to reject H_0. Accept H_0</td><td>A1</td><td></td><td></td></tr>
<tr><td></td><td>There is no significant evidence to suggest that customers rated differently the radio after introducing the new component</td><td>E1</td><td>9</td><td>In context (ft)</td></tr>
<tr><td></td><td>Total</td><td></td><td>9</td><td></td></tr>
</table>

Question 2

2 An American study investigated the weight gains, x kg, of mothers during pregnancy and the weights, y kg, of their children at 3 years of age.

The table gives the results for a random sample of 10 mothers and their children.

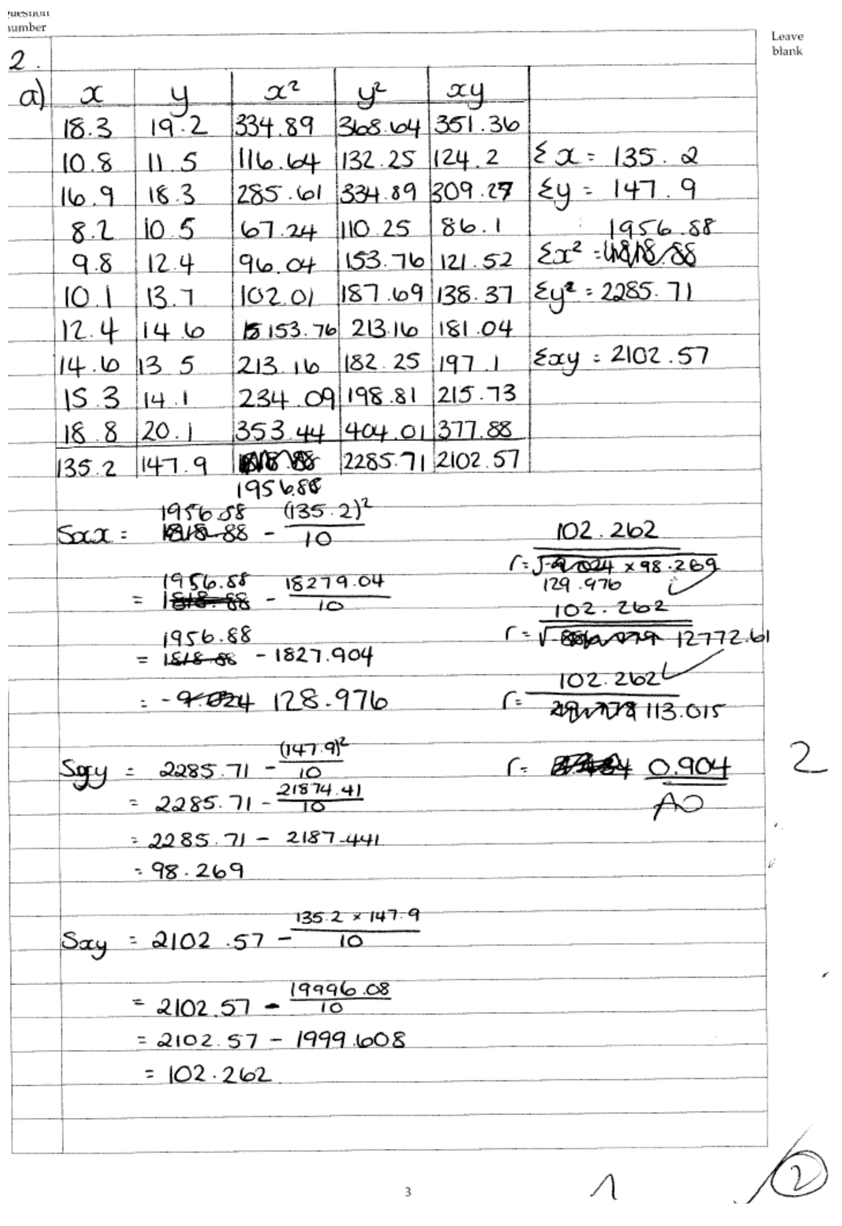
Mother	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
x	18.3	10.8	16.9	8.2	9.8	10.1	12.4	14.6	15.3	18.8
y	19.2	11.5	18.3	10.5	12.4	13.7	14.6	13.5	14.1	20.1

(a) Calculate the value of the product moment correlation coefficient between the weight gains of mothers during pregnancy and the weights of their children at 3 years of age.
(3 marks)

(b) Carry out a hypothesis test, at the 1% level of significance, to determine whether the value that you calculated in part (a) indicates a positive association between the weight gains of mothers during pregnancy and the weights of their children at 3 years of age.

Interpret your conclusion in context. *(5 marks)*

Student response



Question number

Leave blank

2.

a)

x	y	x^2	y^2	xy
18.3	19.2	334.89	368.64	351.36
10.8	11.5	116.64	132.25	124.2
16.9	18.3	285.61	334.89	309.27
8.2	10.5	67.24	110.25	86.1
9.8	12.4	96.04	153.76	121.52
10.1	13.7	102.01	187.69	138.37
12.4	14.6	153.76	213.16	181.04
14.6	13.5	213.16	182.25	197.1
15.3	14.1	234.09	198.81	215.73
18.8	20.1	353.44	404.01	377.88
135.2	147.9	~~1818.88~~ 1956.88	2285.71	2102.57

$\Sigma x = 135.2$

$\Sigma y = 147.9$

$\Sigma x^2 = $ ~~1818.88~~ 1956.88

$\Sigma y^2 = 2285.71$

$\Sigma xy = 2102.57$

$S_{xx} = $ ~~1818.88~~ $1956.88 - \frac{(135.2)^2}{10}$

$= $ ~~1818.88~~ $1956.88 - \frac{18279.04}{10}$

$= $ ~~1818.88~~ $1956.88 - 1827.904$

$= $ ~~−9.024~~ 128.976

$S_{yy} = 2285.71 - \frac{(147.9)^2}{10}$

$= 2285.71 - \frac{21874.41}{10}$

$= 2285.71 - 2187.441$

$= 98.269$

$S_{xy} = 2102.57 - \frac{135.2 \times 147.9}{10}$

$= 2102.57 - \frac{19996.08}{10}$

$= 2102.57 - 1999.608$

$= 102.262$

$r = \frac{102.262}{\sqrt{\cancel{-9.024} \times 98.269}}$

129.976 ✓

$r = \frac{102.262}{\sqrt{12772.61}}$ ✓

$r = \frac{102.262}{113.015}$

$r = \underline{0.904}$

A0

2

3

1

②

Commentary

Part (a) was answered correctly by most candidates but a number calculated Spearman's rank correlation coefficient in error. Most candidates successfully obtained the coefficient from a calculator but many rounded incorrectly or only quoted the answer to 2 sig figs. In part (b), incorrect critical values; either two tail, from *n* = 20 or from the Spearman's rank table, were common.

Mark Scheme

	Total		**9**	
2(a)	From calculator $r = 0.915$ (0.91456)	B3	3	AWRT B2 for 0.914 or 0.91 – 0.92 B1 for 0.9
	or $r = \dfrac{2102.57 - \left(\dfrac{135.2 \times 147.9}{10}\right)}{\sqrt{128.976} \times \sqrt{98.269}}$ $= \dfrac{102.962}{11.35 \times 9.913}$ $= 0.915$			**Alternative:** $n = 10$ $\sum x = 135.2$ $\sum y = 147.9$ $\sum x^2 = 1956.88$ $\sum y^2 = 2285.71$ $\sum xy = 2102.57$ (M1) sub in formula (m1) (A1)
(b)	H_0 $\rho = 0$ H_1 $\rho > 0$ 1 tail 1% sig level	B1		Or words
	test statistic $r = 0.915$			
	cv = 0.7155 n = 10	B1		For cv
	since ts > 0.7155	M1		For comparison ts/cv
	Reject H$_0$	A1		
	Significant evidence at 1% level to suggest a positive linear association between the weight gain of mothers during pregnancy and the weight of their children at 3 years of age	E1	5	In context (ft)
	Total		**8**	

Question 3

3 A long-term investigation was carried out into disease in childhood.

One part of this investigation considered the height of a child at age one year and the income they achieved at age 50 years. The results are summarised in **Table 1**.

Table 1

Income (£) / **Height (cm)**	**Under 20 500**	**20 500 and over**
Under 75	14	6
75 to under 80	12	18
80 and over	8	22

(a) (i) Use a χ^2 distribution and the 5% level of significance to investigate whether income at age 50 years is associated with height at age one year. *(8 marks)*

(ii) Interpret your result in part (a)(i) in the context of the question. *(2 marks)*

(b) Another part of the investigation considered the effect of a parent travelling with their seriously ill child during the child's transfer by a medical team to a paediatric intensive care unit.

For each of 147 transfers, a record was kept of whether or not the child required emergency medical treatment. The results are summarised in **Table 2**.

Table 2

Parent / **Emergency medical treatment**	**Travelled with child**	**Did not travel with child**
Required	10	8
Not required	92	37

Using a 5% significance level, examine whether the presence of a parent travelling with a child is associated with the occurrence of emergency medical treatment.

Interpret your conclusion in context. *(8 marks)*

Student Response

Question number

Leave blank

3a(i) H_0 income aged 50 is independant of height aged 1
H_1 income aged 50 not independant of height aged 1. ✓

height \ income	under 20,500	20500 and over	
under 75	14 8.5	6 11.5	20
75 to under 80	12 12.75	18 17.25	30
80 and above	8 12.75	22 17.25	30
	34	46	✓ 80

$\frac{(O-E)^2}{E}$

$$X^2 = \frac{(14-8.5)^2}{8.5} + \frac{(6-11.5)^2}{11.5} + \frac{(12-12.75)^2}{12.75} + \frac{(18-17.25)^2}{17.25}$$

$$+ \frac{(8-12.75)^2}{12.75} + \frac{(22-17.25)^2}{17.25} =$$

$= 3.5588 + 2.6304 + 0.0441 + 0.0326 + 1.7696$
$+ 1.308 \quad = 9.34$ ✓

$v = (2-1)(3-1) = 2 \quad 5\% \quad CV = 5.991$ ✓

8

Reject H_0 ✓

3a(ii) income aged 50 is associated with height aged 1. It appears the taller the person is aged 1, the more income they get aged 50

1

b) χ^2 H_0 parent travelling with child independant of occurance of emergency treatment
H_1 parent travelling with child is not independant of occurance of emergency treatment.

$5\% \quad v = 1 \quad CV = 3.841$ ✓

Question number

	travelled with child	did not travel with child	
required	10 12.49	8 5.51	18
not required	92 89.51	37 39.49	129
	102	45	147

O	E	O − E	|O − E| − 0.5	$\frac{(\|O-E\|-0.5)^2}{E}$
10	12.49	−2.49	1.99	0.31706
8	5.51	2.49	1.99	0.71871
92	89.51	2.49	1.99	0.04424
37	39.49	−2.49	1.99	0.10028
				1.18029

3.9601

Accept Ho ✓

whether travelling with the child or not the parent didnt influence whether emergency medical treatment was required.

8

⑰

4a) Ho population median = 14

H₁ population median > 14 ✓

10 % 1 tailed n = 9

✓✓

− + + + + + + − +

0, 1, 2, 3, 4, 5, 6, [7, 8, 9

*

B(9, 0.5)

P(0 'more than 14') = 0.0020

P(1 or fewer 'more than 14') = 0.0195 ✓ M1

p(2 or fewer 'more than 14) = 0.0898 ← = P(7 or less) bod

p(3 or fewer 'more than 14) = 0.2539

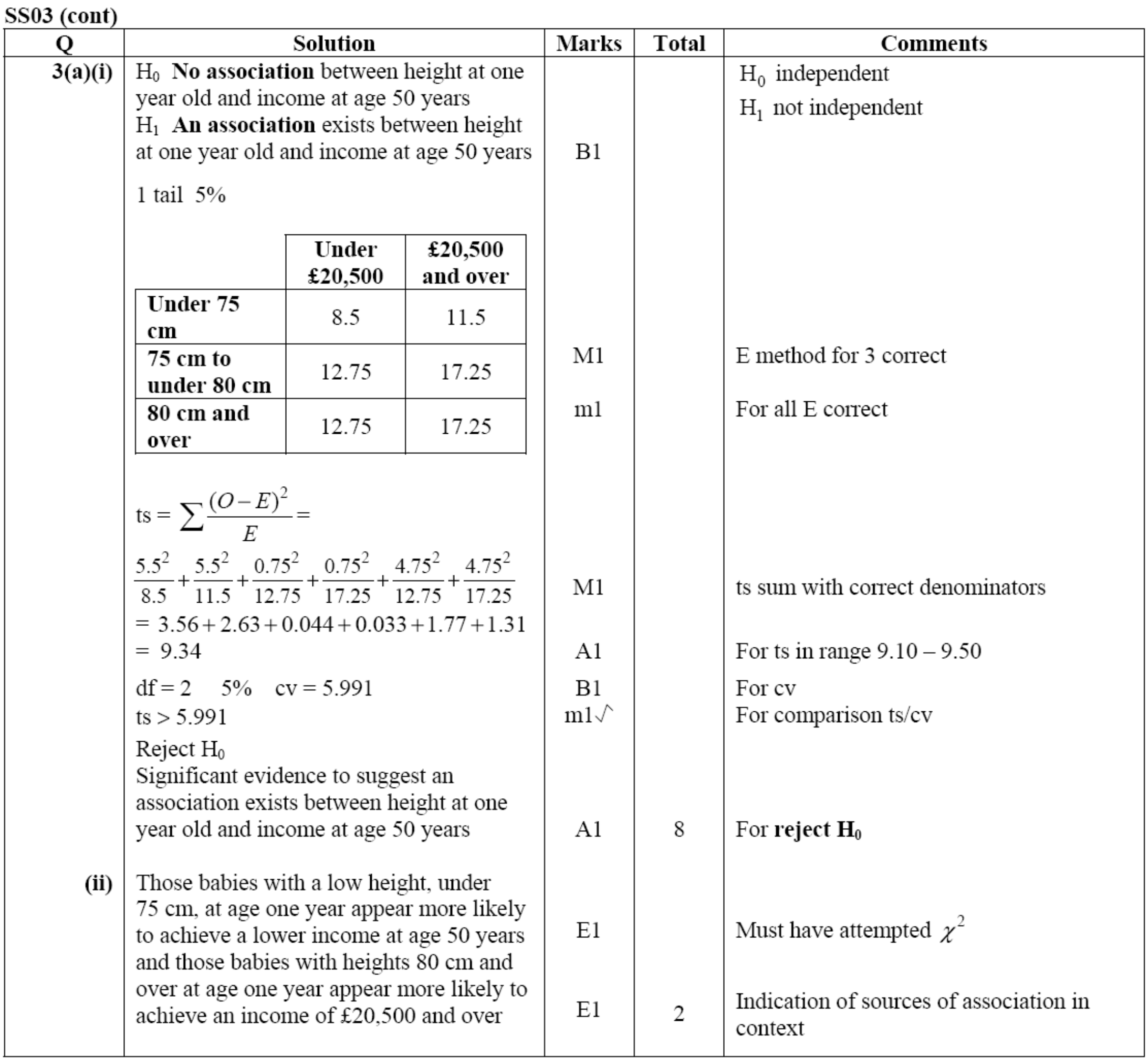
critical region = 7, 8, 9

Leave blank

Commentary

Hypotheses were generally stated correctly.
In part (a)(i), several candidates pooled the rows ‘75 to under 80’ and ‘80 and over’ although expected values were clearly larger than 5. Also, some candidates applied Yates’ correction to the 3x2 table.
In part (a) (iii) candidates often simply repeated the conclusion made in part (a) (i) but did not identify the sources of association by referring to the observed and expected frequencies for height and income.
In part (b), most candidates applied Yates correction but very few applied it correctly. A common error was to use (O – E – 0.5)² .

Mark Scheme



SS03 (cont)

<table>
<tr><th>Q</th><th>Solution</th><th>Marks</th><th>Total</th><th>Comments</th></tr>
<tr><td>3(a)(i)</td><td>H_0 No association between height at one year old and income at age 50 years
H_1 An association exists between height at one year old and income at age 50 years</td><td>B1</td><td></td><td>H_0 independent
H_1 not independent</td></tr>
<tr><td></td><td>1 tail 5%</td><td></td><td></td><td></td></tr>
<tr><td></td><td><table><tr><th></th><th>Under £20,500</th><th>£20,500 and over</th></tr><tr><td>Under 75 cm</td><td>8.5</td><td>11.5</td></tr><tr><td>75 cm to under 80 cm</td><td>12.75</td><td>17.25</td></tr><tr><td>80 cm and over</td><td>12.75</td><td>17.25</td></tr></table></td><td>M1
m1</td><td></td><td>E method for 3 correct
For all E correct</td></tr>
<tr><td></td><td>$\text{ts} = \sum \frac{(O-E)^2}{E} =$
$\frac{5.5^2}{8.5}+\frac{5.5^2}{11.5}+\frac{0.75^2}{12.75}+\frac{0.75^2}{17.25}+\frac{4.75^2}{12.75}+\frac{4.75^2}{17.25}$</td><td>M1</td><td></td><td>ts sum with correct denominators</td></tr>
<tr><td></td><td>$= 3.56+2.63+0.044+0.033+1.77+1.31$
$= 9.34$</td><td>A1</td><td></td><td>For ts in range 9.10 – 9.50</td></tr>
<tr><td></td><td>df = 2   5%   cv = 5.991</td><td>B1</td><td></td><td>For cv</td></tr>
<tr><td></td><td>ts > 5.991</td><td>m1√</td><td></td><td>For comparison ts/cv</td></tr>
<tr><td></td><td>Reject H_0
Significant evidence to suggest an association exists between height at one year old and income at age 50 years</td><td>A1</td><td>8</td><td>For reject H_0</td></tr>
<tr><td>(ii)</td><td>Those babies with a low height, under 75 cm, at age one year appear more likely to achieve a lower income at age 50 years</td><td>E1</td><td></td><td>Must have attempted χ^2</td></tr>
<tr><td></td><td>and those babies with heights 80 cm and over at age one year appear more likely to achieve an income of £20,500 and over</td><td>E1</td><td>2</td><td>Indication of sources of association in context</td></tr>
</table>

SS03 (cont)

<table>
<tr><th>Q</th><th>Solution</th><th>Marks</th><th>Total</th><th>Comments</th></tr>
<tr><td>3(b)</td><td>H_0 Presence of parent is independent of occurrence of emergency medical treatment
H_1 Presence of parent is not independent of occurrence of emergency medical treatment
1 tail 5%</td><td>B1</td><td></td><td>Or as in (a)</td></tr>
<tr><td></td><td><table><tr><th></th><th>Travelled</th><th>Did not travel</th></tr><tr><td>Required</td><td>12.490</td><td>5.510</td></tr><tr><td>Not required</td><td>89.510</td><td>39.490</td></tr></table></td><td>M1</td><td></td><td>For E values method</td></tr>
<tr><td></td><td>$\text{ts} = \sum \frac{(|O-E|-0.5)^2}{E}$</td><td>M1</td><td></td><td>For ts</td></tr>
<tr><td></td><td>$= \frac{1.990^2}{12.490} + \frac{1.990^2}{5.510} + \frac{1.990^2}{89.510} + \frac{1.990^2}{39.490}$</td><td>m1</td><td></td><td>For Yates' corr</td></tr>
<tr><td></td><td>$= 0.317 + 0.719 + 0.044 + 0.100$
$= 1.18$</td><td>A1</td><td></td><td>For ts 1.00 – 1.30</td></tr>
<tr><td></td><td>df = 1 5% cv =3.84</td><td>B1</td><td></td><td>For cv</td></tr>
<tr><td></td><td>ts < 3.84</td><td>M1√</td><td></td><td>For comparison ts/cv</td></tr>
<tr><td></td><td>Accept H_0
No significant evidence to doubt that presence of parent is independent of occurrence of emergency medical treatment</td><td>A1</td><td>8</td><td>In context</td></tr>
<tr><td></td><td>Total</td><td></td><td>18</td><td></td></tr>
</table>

SSO3

Question 4

4 Scientists carried out research in 2005 to investigate the extent of drug abuse in large Italian towns. Waste water from each of a random sample of nine large Italian towns was analysed. For each town, the estimated cocaine use, measured as the number of 100 g daily doses per 1000 young adults, was calculated.

Results from the research gave the following estimates for cocaine use:

9 26 17 18 21 16 19 13 15

The average estimated cocaine use in large Italian towns during the year 2000 was 14 daily doses per 1000 young adults.

(a) Carry out a sign test, at the 10% level of significance, to investigate the claim that the median cocaine use in large Italian towns has increased since the year 2000.

Interpret your conclusion in context. *(7 marks)*

(b) A random sample of young adults from each of the same nine towns in Italy was asked how difficult they thought it was to buy cocaine in their town. They were asked to respond on a scale of 0 to 10, where 0 represents 'not difficult at all' and 10 represents 'extremely difficult'.

The estimated cocaine use, x, together with the average response, y, of young adults (excluding non-responses) from each town are given in the following table.

Town	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
x	9	26	17	18	21	16	19	13	15
y	8.2	3.1	5.1	4.5	2.6	6.1	4.3	7.3	5.8

(i) Calculate the value of Spearman's rank correlation coefficient between x and y. *(6 marks)*

(ii) Interpret your answer to part (b)(i) in the context of the question. *(1 mark)*

Student Response

4a. H_0 population median = 14 ✓ one-tailed test, 10% sig. level.
H_1 population median > 14 ✓ $n = 9$. (0.10)

9 26 17 18 21 16 19 13 15
− + + + + + + − + ✓ ✓

ts = 2− (or 7+)

$p = P(x \leq 2) = P(x \geq 7) = 0.0898$ ✓ < 0.10 ✓ so ~~we as~~ the test statistic must lie in the critical region and so we reject Ho. There is significant evidence to suggest that the median number of daily cocaine doses has increased. ✓

7

8

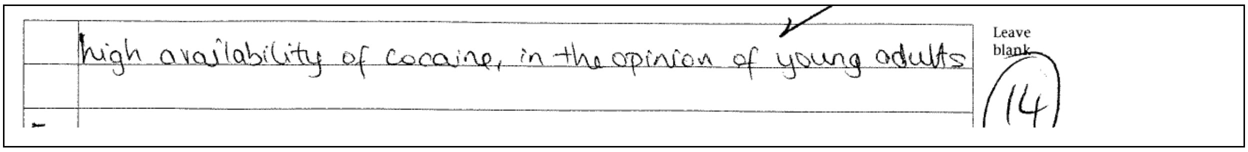
b. i)

Town	A	B	C	D	E	F	G	H	I
x	1	9	5	6	8	4	7	2	3
y	9	2	5	4	1	7	3	8	6

$r_s \overset{-}{=} 0.967$ ✓

6

ii) This value is a ~~positive~~ negative value and is close to −1, showing a very strong ~~positive~~ negative relationship between estimated cocaine use and average response of young adults. It would appear therefore that with high cocaine use there will also be



high availability of cocaine, in the opinion of young adults

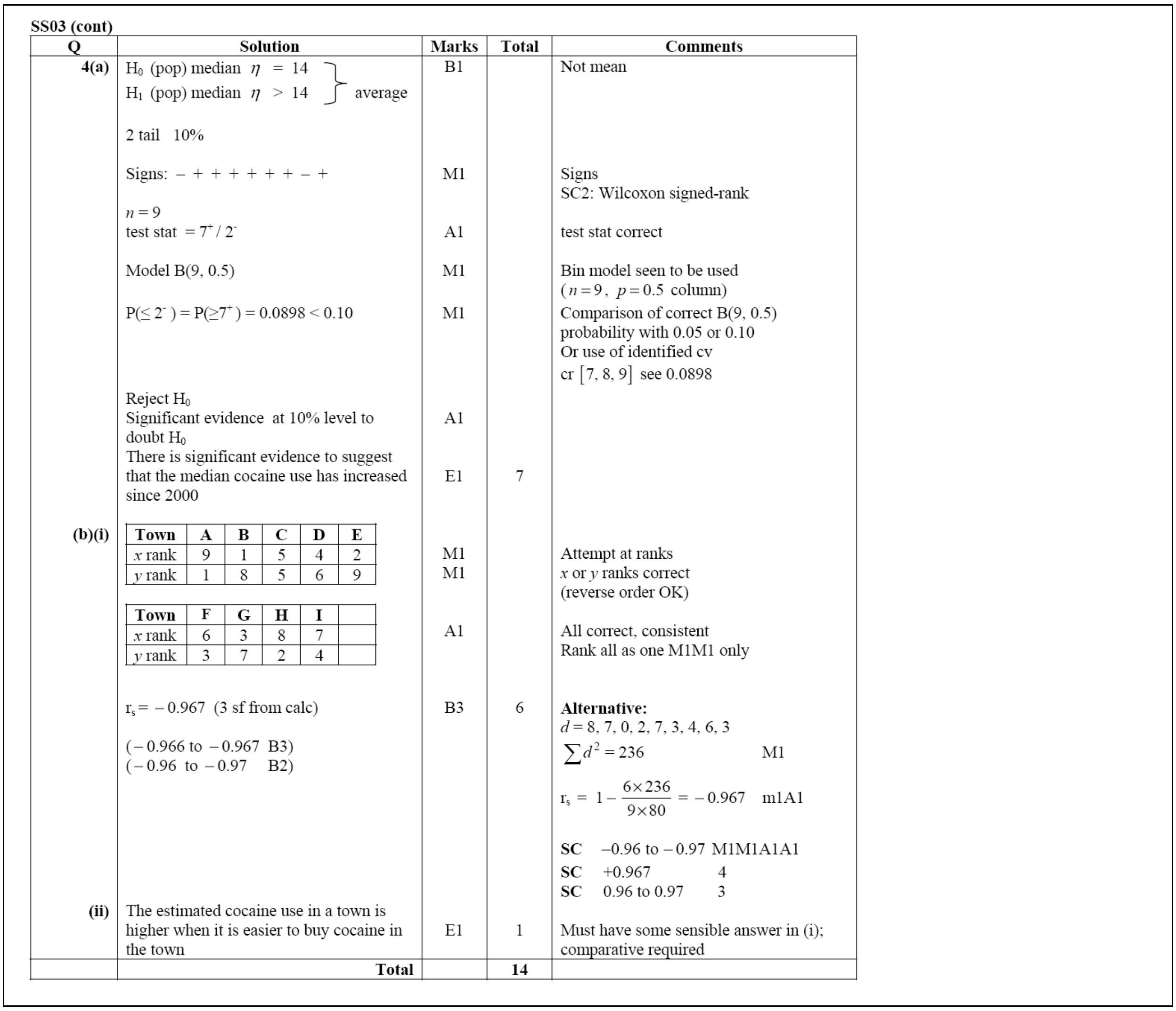
Leave blank

(14)

Commentary

Some excellent solutions to part (a). The majority of candidates quoted the binomial probability of 0.0898 and showed a comparison with 0.10.
Candidates lost marks if probabilities from the binomial tables were not stated or a critical region was identified without the relevant probability being quoted.
Some candidates carried out a Wilcoxon signed-rank test that gained some marks but was not the test specified. Some found differences between the given values and the median/mean of all the given data. This did not gain any marks.
The hypotheses were often poorly stated.
In part (b) (i) most candidates knew how to evaluate a Spearman's rank correlation coefficient and showed the relevant ranks. In part (b) (ii) the interpretation of the coefficient – 0.967 was often poor. Usually no reference was made to the fact that towns that had a higher usage of cocaine were those where it was easier to buy the drug.

Mark Scheme



SS03 (cont)

<table>
<tr><th>Q</th><th>Solution</th><th>Marks</th><th>Total</th><th>Comments</th></tr>
<tr><td>4(a)</td><td>H_0 (pop) median $\eta = 14$
H_1 (pop) median $\eta > 14$ } average</td><td>B1</td><td></td><td>Not mean</td></tr>
<tr><td></td><td>2 tail 10%</td><td></td><td></td><td></td></tr>
<tr><td></td><td>Signs: – + + + + + + – +</td><td>M1</td><td></td><td>Signs
SC2: Wilcoxon signed-rank</td></tr>
<tr><td></td><td>$n = 9$
test stat $= 7^+ / 2^-$</td><td>A1</td><td></td><td>test stat correct</td></tr>
<tr><td></td><td>Model B(9, 0.5)</td><td>M1</td><td></td><td>Bin model seen to be used
($n = 9$, $p = 0.5$ column)</td></tr>
<tr><td></td><td>$P(\leq 2^-) = P(\geq 7^+) = 0.0898 < 0.10$</td><td>M1</td><td></td><td>Comparison of correct B(9, 0.5) probability with 0.05 or 0.10
Or use of identified cv
cr $[7, 8, 9]$ see 0.0898</td></tr>
<tr><td></td><td>Reject H_0
Significant evidence at 10% level to doubt H_0</td><td>A1</td><td></td><td></td></tr>
<tr><td></td><td>There is significant evidence to suggest that the median cocaine use has increased since 2000</td><td>E1</td><td>7</td><td></td></tr>
<tr><td>(b)(i)</td><td><table><tr><th>Town</th><th>A</th><th>B</th><th>C</th><th>D</th><th>E</th></tr><tr><td>x rank</td><td>9</td><td>1</td><td>5</td><td>4</td><td>2</td></tr><tr><td>y rank</td><td>1</td><td>8</td><td>5</td><td>6</td><td>9</td></tr></table></td><td>M1
M1</td><td></td><td>Attempt at ranks
x or y ranks correct
(reverse order OK)</td></tr>
<tr><td></td><td><table><tr><th>Town</th><th>F</th><th>G</th><th>H</th><th>I</th><th></th></tr><tr><td>x rank</td><td>6</td><td>3</td><td>8</td><td>7</td><td></td></tr><tr><td>y rank</td><td>3</td><td>7</td><td>2</td><td>4</td><td></td></tr></table></td><td>A1</td><td></td><td>All correct, consistent
Rank all as one M1M1 only</td></tr>
<tr><td></td><td>$r_s = -0.967$ (3 sf from calc)

(-0.966 to -0.967 B3)
(-0.96 to -0.97 B2)</td><td>B3</td><td>6</td><td>Alternative:
$d = 8, 7, 0, 2, 7, 3, 4, 6, 3$
$\sum d^2 = 236$ M1
$r_s = 1 - \frac{6 \times 236}{9 \times 80} = -0.967$ m1A1

SC -0.96 to -0.97 M1M1A1A1
SC $+0.967$ 4
SC 0.96 to 0.97 3</td></tr>
<tr><td>(ii)</td><td>The estimated cocaine use in a town is higher when it is easier to buy cocaine in the town</td><td>E1</td><td>1</td><td>Must have some sensible answer in (i); comparative required</td></tr>
<tr><td></td><td>Total</td><td></td><td>14</td><td></td></tr>
</table>

Question 5

5 Ms Testum wishes to investigate whether students will score differently in a test depending on whether the test is taken in a morning session or an afternoon session.

She selects a group of 19 students of similar ability and randomly assigns some of them to take the test in the morning and the remainder to take the same test in the afternoon.

The students taking the test in the morning are kept apart from the students taking the test in the afternoon until all the students have taken the test.

The ordered scores are given in the table.

Session	Score
morning	44
afternoon	46
afternoon	47
afternoon	51
morning	53
morning	54
morning	56
afternoon	58
afternoon	59
afternoon	61
afternoon	62
morning	63
morning	63
morning	65
afternoon	67
afternoon	68
morning	72
morning	74
morning	81

(a) Carry out a Mann-Whitney U test, at the 5% level of significance, to investigate whether there is any difference in the average test score between mornings and afternoons.

Interpret your conclusion in context. *(10 marks)*

(b) A matched-pairs design was suggested for this investigation.

(i) Explain why a matched-pairs design might be preferred when comparing two groups. *(2 marks)*

(ii) Explain how Ms Testum tried to ensure that her test was not biased. *(2 marks)*

Student Response

town.

5 (a) H_0: samples are from identical population

H_1: samples are not from identical population (average test score do differ in ~~[illegible]~~ samples), 5% sig level 2-tail 14

Ranks

Morning	1	5	6	7	12.5	12.5	14	17	18	19
Afternoon	2	3	4	8	9	10	11	15	16	

$T_M = 112$ $T_A = 78$

$U_M = T_M - \frac{m(m+1)}{2}$ $U_A = T_A - \frac{n(n+1)}{2}$

$= 112 - \frac{10(11)}{2}$ $= 78 - \frac{9(10)}{2}$

$= \underline{57}$ $= \underline{33}$

$ts = \underline{33}$

$m=10, n=9$, $cv = \underline{21}$

$ts > cv$

accept H_0

no evidence suggesting saple are not from identical population.

~~the average~~ no significant evidence suggesting average test score in the morning differs from that in the afternoon. 10

(ie. no evidence suggesting they score differently in the morning than in the afternoon sessions)

(b), i) Could eliminate the experimental errors associated to the individual variation of students (eg intelligence), and ensure that any difference detected, if one exists, is due to the difference in time of taking the test (ie morning/afternoon) 2

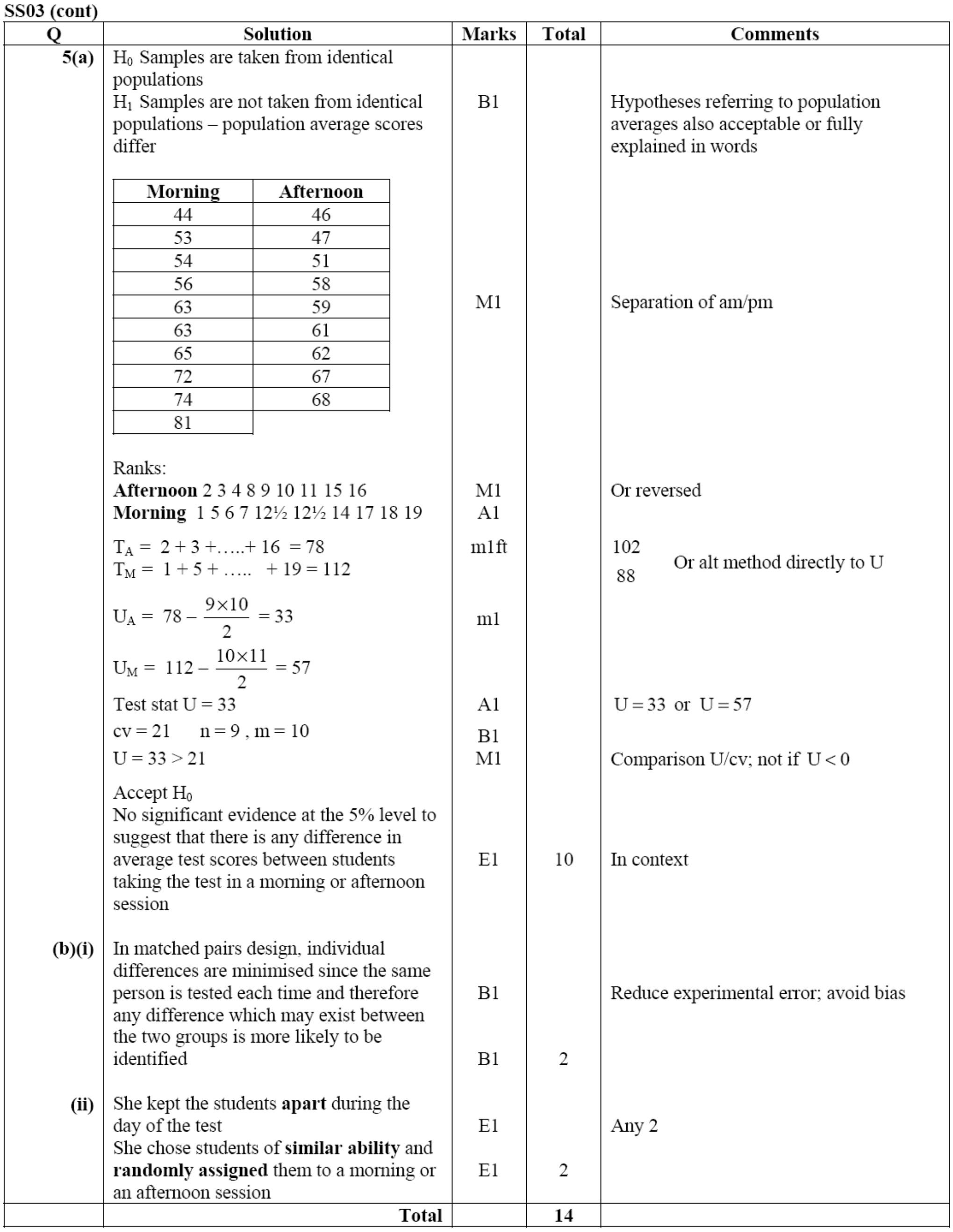
(ii) SHe tried to choose students of similar ability so that the students won't score very different because of individual difference in ability. Also, she rand[illegible]ssigned the students to the two sessions, which might eliminate the effects of preferences of individual students ~~[illegible]~~ if they were to choose by themselves which session to sit. 2

14

Commentary

The Mann-Whitney test was generally well done in part (a) and candidates sorted and ranked the data efficiently but not always as one group.
Hypotheses were well worded in most cases with ‘population’ mentioned or an explanation about average scores for morning and afternoon.
Conclusions were usually in context and correct.
Candidates usually only gained 1 mark for part(b)(i).Avoiding bias was usually mentioned but not the fact that any difference between morning and afternoon scores is more likely to be detected if a paired test is carried out.
Part (b) (ii) was very well answered.

Mark Scheme



SS03 (cont)

Q	Solution	Marks	Total	Comments
5(a)	H_0 Samples are taken from identical populations H_1 Samples are not taken from identical populations – population average scores differ	B1		Hypotheses referring to population averages also acceptable or fully explained in words
	Morning **Afternoon** 44 46 53 47 54 51 56 58 63 59 63 61 65 62 72 67 74 68 81	M1		Separation of am/pm
	Ranks: **Afternoon** 2 3 4 8 9 10 11 15 16	M1		Or reversed
	Morning 1 5 6 7 12½ 12½ 14 17 18 19	A1		
	$T_A = 2 + 3 + \ldots + 16 = 78$ $T_M = 1 + 5 + \ldots + 19 = 112$	m1ft		102 88 Or alt method directly to U
	$U_A = 78 - \frac{9\times10}{2} = 33$ $U_M = 112 - \frac{10\times11}{2} = 57$	m1		
	Test stat U = 33	A1		U = 33 or U = 57
	cv = 21 n = 9 , m = 10	B1		
	U = 33 > 21	M1		Comparison U/cv; not if $U < 0$
	Accept H_0 No significant evidence at the 5% level to suggest that there is any difference in average test scores between students taking the test in a morning or afternoon session	E1	10	In context
(b)(i)	In matched pairs design, individual differences are minimised since the same person is tested each time and therefore any difference which may exist between the two groups is more likely to be	B1		Reduce experimental error; avoid bias
	identified	B1	2	
(ii)	She kept the students **apart** during the day of the test	E1		Any 2
	She chose students of **similar ability** and **randomly assigned** them to a morning or an afternoon session	E1	2	
	Total		**14**	

Question 6

6 It is believed that a happy marriage can offer increased immunity to infections in those aged over 65.

A sample of 15 males, aged over 65, had their levels of protective antibodies measured one month after their flu jabs. Each male selected his marital category as either 'Happily Married', 'Unhappily Married' or 'Unmarried'.

The results are given in the table.

Happily Married	Unhappily Married	Unmarried
194	185	146
215	192	150
242	210	155
285	236	168
291		195
292		

Carry out a distribution-free test, using the 5% significance level, to investigate whether there is any difference between the average level of protective antibodies found in males aged over 65 for the three marital categories.

Interpret your conclusion in context. *(12 marks)*

Student Response

If $H > \chi^2 \rightarrow$ reject Ho

Test Stat

no

$H = \frac{12}{N(N+1)} \sum \frac{T_i^2}{n_i} - 3(n+1)$ ranks

$\frac{12}{15(15+1)} = 0.05$

$3(n+1) = 48$

?

~~$\frac{T_i^2}{N_i}$~~ $\frac{T_i^2}{N_i} = \frac{1096}{15} = 73.0666667$

≈ 73.1 ∴ X

~~Happily married Row~~

$H = 0.05 \sum 73.1 - 48$

$= 0.05 \sum 125.1$

$= 6.255$

≈ 6.26 X

Comparison + conclusion

H crit region

$6.26 > 0.068359$ ∴ reject Ho

X

We conclude there is significant evidence at the 5% significance level, to suggest that at least two of the pop. median ~~differences~~ are different.

no valid method

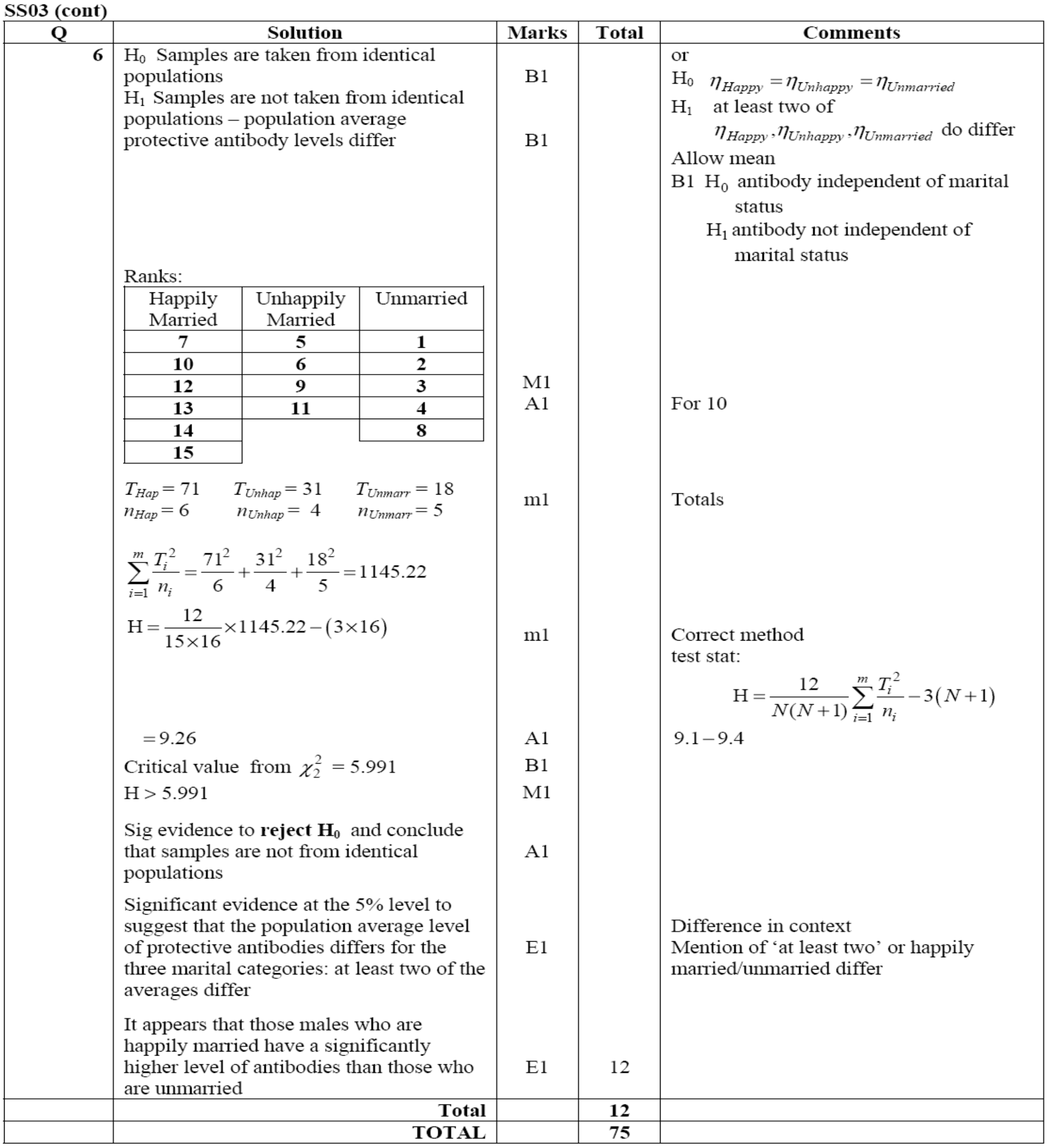
0

$\textcircled{0}$

Commentary

Candidates frequently incorrectly stated the hypotheses and, if referring to population medians, failed to mention that the alternative hypothesis should be that **at least two** of the average antibody levels from the three marital categories differ.
The Kruskal Wallis test was carried out accurately by most candidates but some did not show their rank values or made errors in ranking. A substantial number were unable to substitute the relevant values into the formula.
Critical values were frequently obtained from n =15 rather than n = 3.
Most candidates identified ‘Happily married’ men as having the highest average level of antibodies but few realised that the significant difference found was between the average levels for ‘Happily married’ and ‘Unmarried’ men.

Mark Scheme



SS03 (cont)

<table>
<tr><th>Q</th><th>Solution</th><th>Marks</th><th>Total</th><th>Comments</th></tr>
<tr><td>6</td><td>H_0 Samples are taken from identical populations</td><td>B1</td><td></td><td>or
H_0 $\eta_{Happy} = \eta_{Unhappy} = \eta_{Unmarried}$</td></tr>
<tr><td></td><td>H_1 Samples are not taken from identical populations – population average protective antibody levels differ</td><td>B1</td><td></td><td>H_1 at least two of $\eta_{Happy}, \eta_{Unhappy}, \eta_{Unmarried}$ do differ
Allow mean
B1 H_0 antibody independent of marital status
H_1 antibody not independent of marital status</td></tr>
<tr><td></td><td>Ranks:<table>
<tr><th>Happily Married</th><th>Unhappily Married</th><th>Unmarried</th></tr>
<tr><td>**7**</td><td>**5**</td><td>**1**</td></tr>
<tr><td>**10**</td><td>**6**</td><td>**2**</td></tr>
<tr><td>**12**</td><td>**9**</td><td>**3**</td></tr>
<tr><td>**13**</td><td>**11**</td><td>**4**</td></tr>
<tr><td>**14**</td><td></td><td>**8**</td></tr>
<tr><td>**15**</td><td></td><td></td></tr>
</table></td><td>M1
A1</td><td></td><td>
For 10</td></tr>
<tr><td></td><td>$T_{Hap} = 71$ $T_{Unhap} = 31$ $T_{Unmarr} = 18$
$n_{Hap} = 6$ $n_{Unhap} = 4$ $n_{Unmarr} = 5$</td><td>m1</td><td></td><td>Totals</td></tr>
<tr><td></td><td>$\sum_{i=1}^{m} \frac{T_i^2}{n_i} = \frac{71^2}{6} + \frac{31^2}{4} + \frac{18^2}{5} = 1145.22$
$H = \frac{12}{15 \times 16} \times 1145.22 - (3 \times 16)$</td><td>m1</td><td></td><td>Correct method
test stat:
$H = \frac{12}{N(N+1)} \sum_{i=1}^{m} \frac{T_i^2}{n_i} - 3(N+1)$</td></tr>
<tr><td></td><td>$= 9.26$</td><td>A1</td><td></td><td>9.1 – 9.4</td></tr>
<tr><td></td><td>Critical value from $\chi_2^2 = 5.991$</td><td>B1</td><td></td><td></td></tr>
<tr><td></td><td>$H > 5.991$</td><td>M1</td><td></td><td></td></tr>
<tr><td></td><td>Sig evidence to **reject H_0** and conclude that samples are not from identical populations</td><td>A1</td><td></td><td></td></tr>
<tr><td></td><td>Significant evidence at the 5% level to suggest that the population average level of protective antibodies differs for the three marital categories: at least two of the averages differ</td><td>E1</td><td></td><td>Difference in context
Mention of ‘at least two’ or happily married/unmarried differ</td></tr>
<tr><td></td><td>It appears that those males who are happily married have a significantly higher level of antibodies than those who are unmarried</td><td>E1</td><td>12</td><td></td></tr>
<tr><td></td><td>**Total**</td><td></td><td>**12**</td><td></td></tr>
<tr><td></td><td>**TOTAL**</td><td></td><td>**75**</td><td></td></tr>
</table>