OCR
RECOGNISING ACHIEVEMENT

2641/01 Statistics 1

June 2004

Mark Scheme

<table>
<tr>
<td>1 (i)</td>
<td>
<table>
<tr><td>Aspect</td><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td></tr>
<tr><td>Person 1</td><td>71</td><td>63</td><td>84</td><td>57</td><td>64</td></tr>
<tr><td>Person 2</td><td>12</td><td>62</td><td>20</td><td>85</td><td>31</td></tr>
<tr><td>Rank 1</td><td>2</td><td>4</td><td>1</td><td>5</td><td>3</td></tr>
<tr><td>Rank 2</td><td>5</td><td>2</td><td>4</td><td>1</td><td>3</td></tr>
<tr><td>d</td><td>-3</td><td>2</td><td>-3</td><td>4</td><td>0</td></tr>
</table>
$\Sigma d^2 = 9 + 4 + 9 + 16 + 0 = 38$

Spearman's rank correlation

Coefficient = $1 - \frac{6 \times 38}{5 \times 24} = \frac{-9}{10} = -0.9$
</td>
<td>B1

M1

M1

A1

4</td>
<td>Correct ranks (or reverse)

Attempt to find d (or d^2) from ranked or ordered data

Correct formula for Spearman used and $|r| < 1$
Correct answer -0.9 or $\frac{-9k}{10k}$
cao</td>
</tr>
<tr>
<td>(ii)</td>
<td>Spearman's rank correlation coefficient shows that the two people have different, opposite views, or no or little agreement when considering aspects of their job</td>
<td>B1

1</td>
<td>Comment in context, consistent with r_s value
$|r| < 1$</td>
</tr>
<tr>
<td>2 (i)

(ii)</td>
<td>Number of possible arrangements = $\frac{5!}{2}$ = 60

Number of arrangements in which the white bricks are at each end = 3!
or
Number of arrangements in which both bricks are at either end = $3! \times 2!$

Therefore P(white bricks are at each end)
$= \frac{3!}{60} = \frac{6}{60}$ = 0.1

or P(both white bricks at either end)
$= \frac{3! \times 2!}{60}$ = 0.2

or P(white at each end or both at either end) = 0.1 + 0.2 = 0.3</td>
<td>M1
A1
2

M1

M1

A1
3</td>
<td>5! or 120 seen (not in $^5C_3$)
60, cao

3! Seen for either case

their 3! Divided by their (i)

0.1 or 0.2 or 0.3 $\frac{k}{10k}$ or $\frac{k}{5k}$ or $\frac{3k}{10k}$</td>
</tr>
</table>

<table>
<tr><td>3 (i)</td><td>Let X = number of heads in 4 randomly chosen spins of the coin.
$P(W = -5) = P(X = 0 \text{ or } 1) \quad = \frac{1}{16} + \frac{4}{16}$
$= \frac{5}{16}$ **AG**</td><td>M1

A1

2</td><td>Attempt to find P(X = 0) or P(X = 1)

Wholly correct and convincing attempt (allow decimals)</td></tr>
<tr><td>(ii)</td><td>Distribution table for X

x	0	1	2	3	4
$P(X = x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

Distribution tale for W

w	-5	5	10
$P(W = w)$	$\frac{5}{16}$	$\frac{6}{16}$	$\frac{5}{16}$

</td><td>M1

A1

2</td><td>A clear attempt to derive $P(W = 5)$ or $P(W = 10)$

Wholly correct table</td></tr>
<tr><td>(iii)</td><td>$E(W) = (-5) \times \frac{5}{16} + 5 \times \frac{6}{16} + 10 \times \frac{5}{16}$
$= \frac{-25+30+50}{16} = \frac{55}{16} (= 3.4375)$ **AG**</td><td>M1

A1
2</td><td>Use of Σwp for *their* distribution table, at least 2 wp terms added
Wholly correct method</td></tr>
<tr><td>(iv)</td><td>$E(W^2) = (-5)^2 \times \frac{5}{16} + 5^2 \times \frac{6}{16} + 10^2 \times \frac{5}{16}$
$= \frac{775}{16}$
So $\text{Var}(W) = E(W^2) - [E(W)]^2$
$= \left(\frac{775}{16}\right) - \left(\frac{55}{16}\right)^2$
$= \frac{9375}{256} = 36.62.. = 36.6$ (3 sf)</td><td>M1

M1

A1
3</td><td>Use of Σw^2p for *their* distribution table, at least 2 w^2p terms added
Subtracting (*their* mean)^2

$\frac{9375k}{256k}$ or a.r.t. 36.6</td></tr>
</table>

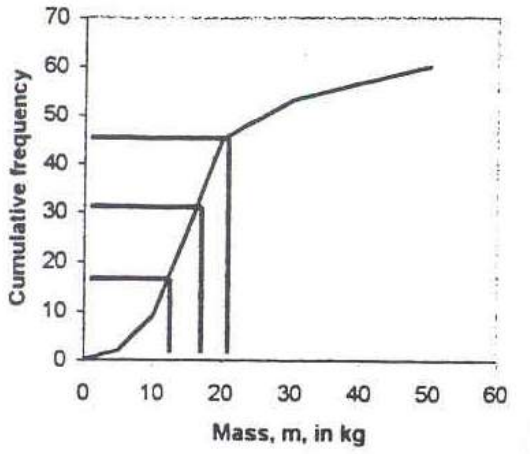
4 (i)	Since x is a controlled variable, only the y on x line is appropriate	M1		Use of y on x line
	$S_{xy} = \Sigma xy - \frac{(\Sigma x)(\Sigma y)}{n} =$ $= 244260 - \frac{550\times3717}{10} = 39825$ $S_{xx} = \Sigma x^2 - \frac{(\Sigma x)^2}{n}$ $= 38500 - \frac{550^2}{10} = 8250$ $\bar{x} = 55,\ \bar{y} = 371.7$			
	$b = \frac{S_{xy}}{S_{xx}} = \frac{39825}{8250} = 4.82727...$	M1		$\frac{S_{xy}}{S_{xx}}$ used
	$a = \bar{y} - b\bar{x} = 371.7 - 4.82727\ldots\ 55$ $= 106.2$			May be implied if calculator routine is used
	Equation of line is	M1		Using $= \bar{y} - b\bar{x}$ with *their* b
	$y = 4.82727\ldots x + 106.2$ $y = 4.83x + 106$ (3 sf)	A1		$y = 4.83x + 106.2$, or correct equivalent (does **not** need to be in the form $y = a + bx$)
	Estimated value of x $= \frac{(220 - 106.2)}{4.82727...}$ $= 23.57438$ **= 23.6 (3 sf)**	M1		Substitute $y = 220$ into *their* equation
		A1	6	a.r.t. 23.6
				Use of the x on y line
		M1		For $\frac{S_{xy}}{S_{xx}}$ used
		M1		Using $\bar{x} - b'\bar{y}$ with *their b′*
		A1		$x = -21.1 + 0.205y$
		M1		Substitute $y = 220$ into *their* equation
(ii)	$S_{yy} = \Sigma y^2 - \frac{(\Sigma y)^2}{n}$ $= 1576075 - \frac{3717^2}{10}$ $= 194466.1$			
	$r = \frac{S_{xy}}{\sqrt{S_{xx} \times S_{yy}}}$	M1		Calculator of formula correctly used or equivalent (may be implied)
	$= \frac{39825}{\sqrt{8250\times194466.1}}$ $= 0.99427\ldots =$ **0.994 (3 sf)**	A1		Correct answer, a.r.t. 0.994
	This is a very high positive correlation so the estimate is likely to be reliable	B1	3	Comment consistent with *their r* value, provided $\|r\| < 1$

5 (i)	$X \sim B(10, \frac{1}{57})$ Independence: whether Andy wins a particular lottery game is independent of whether he has won any other game. Two possible outcomes: for each game Andy either wins or loses.	B1) B1) B1 3	Binomial stated $n = 10$ and $p = \frac{1}{57}$ stated clearly One valid comment in context
(ii)(a)	$P(X = 2) = {}^{10}C_2 \times \left(\frac{1}{57}\right)^2 \times \left(\frac{56}{57}\right)^8$ $= 0.0120217633$ **= 0.012**	M1 M1 A1 3	Their ${}^{n}C_2 \times p^2 \times (1 - p)^{n-2}$ used Wholly correct method a.r.t. 0.012
(b)	$P(X > 2)$ $= 1 - P(X = 0) - P(X = 1) - P(X = 2)$ $= 1 - [0.83778\ldots + 0.14960\ldots + 0.01202\ldots]$ $= 0.00059074\ldots$ **= 0.000591 (3 sf)**	M1 M1 A1 3	1 – [P(X=0)+P(X=1)+P(X=2)] with at least 2 probs attempted Wholly correct method a.r.t. 0.0006
(iii)	$E(X) = np = 10 \times \frac{1}{57} = \frac{10}{57}$ = 0.175438… = **0.175 (3 sf)**	B1 1	$= \frac{10k}{57k}$ or a.r.t. 0.175

6 (i)

Mass, m in kg	Cumulative frequency
$m < 5$	2
$m < 10$	9
$m < 15$	26
$m < 20$	45
$m < 30$	53
$m < 50$	60

M1 — At least one correct cumulative frequency seen, other than 2



M1 — At least 4 points correct with the correct (u.c.b., cum freq)

A1 — Wholly correct diagram

3

6 (ii)

From the graph

M1 — Correct method for *either* the median or for *a* quartile

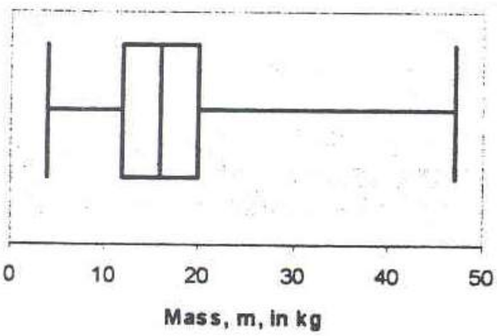
Reading from the CF axis at 15 (or 15.25) for Q_1, at 30 (or 30.5) for Q_2, at 45 (or 45.75) for Q_3

A1 — *Their* Q_2 from *their* CF curve, provided u.c.b's used

A1 — *Their IQR* from *their* CF curve

3

6 (iii)



M1 — A recognisable attempt at a boxplot

M1 — At least 4 from 4: *their* Q_1: *their* Q_2 *their* Q_3: 47 correctly plotted

A1 — Wholly correct diagram

3

6 (iv)

Comment on skewness, range, IQR, 5 summary numbers, max and/or min values, symmetry

B1 — One valid feature of data which can be deduced more easily from a boxplot, but do not allow median and/or quartiles.

1

7 (i)	Possible routes: ABA → prob $= \frac{2}{3}\times\frac{3}{4}$ ACA → prob $= \frac{1}{3}\times\frac{4}{5}$ P(back at A) $= \frac{1}{2}+\frac{4}{15}=\frac{15}{30}+\frac{8}{30}$ $= \frac{23}{30}$ **AG**	M1 M1 A1 3	One correct product seen Both correct routes identified (letters, probs, tree diagram) **and** one correct product. No other routes allowed. Wholly convincing and correct
7 (ii)	Possible routes= ABCA or ACBA So prob $= \frac{2}{3}\times\frac{1}{4}\times\frac{4}{5}+\frac{1}{3}\times\frac{1}{5}\times\frac{3}{4}$ $= \frac{2}{15}+\frac{1}{20}=\frac{8}{60}+\frac{3}{60}$ $= \frac{11}{60}$ or **0.183.. = 0.183** (3 sf)	M1 M1 M1 A1 4	One correct route identified Both correct routes identified and one correct product Wholly correct method (no other routes) $\frac{11k}{60k}$ or a.r.t. 0.183
7 (iii)	Possible routes ACBCB → $\frac{1}{3}\times\frac{1}{5}\times\frac{1}{4}\times\frac{1}{5}$ ACBAB → $\frac{1}{3}\times\frac{1}{5}\times\frac{3}{4}\times\frac{2}{3}$ ACACB → $\frac{1}{3}\times\frac{4}{5}\times\frac{1}{3}\times\frac{1}{5}$ ABACB → $\frac{2}{3}\times\frac{3}{4}\times\frac{1}{3}\times\frac{1}{5}$ ABCAB → $\frac{2}{3}\times\frac{1}{4}\times\frac{4}{5}\times\frac{2}{3}$ $= \frac{1}{300}+\frac{1}{30}+\frac{4}{225}+\frac{1}{30}+\frac{4}{45}=\frac{53}{300}=$ 0.176666. = **0.177** (3 s.f.)	M1 M1 M1 M1 A1 5	At least 4 correct routes chosen 2 correct routes identified and one correct 4-termed product 3 correct products all products correct and added (no other routes) $\frac{53k}{300k}$ or a.r.t. 0.177
7 (iii)	ALITER: (i) $\times\frac{1}{3}\times\frac{1}{5}+$ (ii) $\times\frac{2}{3}+\frac{1}{3}\times\frac{1}{5}\times\frac{1}{4}\times\frac{1}{5}$		