EDEXCEL STATISTICS S3 (6685) – JUNE 2004 **PROVISIONAL MARK SCHEME**

Question Number	Scheme	Marks
1a)	Allocate a number between 1 and N (or equiv) to each pupil.	M1
	Use random number tables, computer or calculator to select 15 different numbers between 1 and 120 (or equiv).	B1
	Pupils corresponding to these numbers become the sample.	B1 (3)
(b)	Allocate numbers 1 – 64 to girls and 1 – 56 to boys. Idea of different sets for boys and girls	M1
	Select $\frac{64}{120}\times 15 = 8$ random numbers between 1 – 64 for girls attempt find no	M1
	Select 7 random numbers between 1 – 56 for boys. Both 7 and 8	A1 (3)
2a)		
	$H_0: \rho = 0$; $H_1: \rho > 0$ both and ρ	B1 B1
	5% CV – PMCC 0.6215	M1
	0.572 < 0.6215 / not in critical region / not significant	A1
	No evidence of positive correlation	B1
	Spearman 0.6429	B1 (6)
(b)	Evidence of positive correlation	B1
	No evidence to suggest that as Statistics marks increased Geography marks increased. Context and not correlation	B1
	Evidence that students ranked highly in Statistics were also ranked highly in Geography ranked	(2)

EDEXCEL STATISTICS S3 (6685) – JUNE 2004 **PROVISIONAL MARK SCHEME**

Question Number	Scheme	Marks
3a)	$H_0: \mu_A = \mu_B$; $H_1: \mu_B > \mu_A$ both and μ	B1
	$z = \pm \dfrac{249 - 251}{\sqrt{\dfrac{2.5^2}{10} + \dfrac{2.3^2}{15}}}$ 249,251 accept; $\sqrt{\frac{2.5}{10}+\frac{2.3}{15}}$ for M	M1 A1
	$= \pm 2.0227\ldots$ awrt ± 2.02	A1
	CV = ± 1.6449 or $P(Z \geq 2.02) = 0.0212 - 0.0217$, or $P(Z \leqslant 2.02) = 0.9788 - 0.9783$	B1
	$-2.0227 < -1.6449$ or $2.0227 > 1.6449$, or $0.0212 - 0.0217 < 0.05$ comparison and consistency needed or $0.9788 - 0.9783 > 0.95$	M1
	There is evidence that the <u>mean amount of coffee</u> dispensed by B <u>is greater</u> than A. context	A1√ (7)
b)	Machine B amounts are normally distributed.	B1 (1)

EDEXCEL STATISTICS S3 (6685) – JUNE 2004 **PROVISIONAL MARK SCHEME**

Question Number	Scheme	Marks
4a)	$\bar{x} = 75.3$	B1
	$s^2 = \frac{1}{9}\left\{57455 - \frac{753^2}{10}\right\}$	M1
	$= 83.7\dot{8}$, $83\frac{71}{90}$, 83.8 awrt 83.8	A1 (3)
b)	1.96	B1
	$74.8 \pm 1.96\sqrt{\frac{84.6}{100}}$ any z value, may use 75.3,83.8 for M	M1 A1√ on z only
	(73.0, 76.6) awrt 73.0,76.6	A1, A1 (5)
c)	Journey times independent	
	Sample large enough to use central limit theorem any 2	B1,B1
	Same distribution / population	(2)

Question Number	Scheme	Marks
5.		M1 convert % to freq
	Never	A1 (26, 91, 30, 132)
	Sometimes	
	Regularly	A1 (143, 78)
	Totals	
	Males	B1
	30	
	132	B1
	78	
	240	
	Females	
	26	
	143	
	91	M1
	260	A1 at least 3sf
	56	
	275	
	169	
	500	B1; B1√
	H_0 : No association (independent) between gender and exercise	
	H1 : association (not independent) between gender and exercise	M1 A1
	Expected Values	A1√
		(12)
	Never	
	Sometimes	
	Regularly	
	Totals	
	Males	
	26.88	
	132	
	81.12	
	240	
	Females	
	29.12	
	143	

	87.88 260 56 275 169 500 $\alpha = 0.05$ $\underline{\nu = 2}$; CV $\chi^2 > \underline{5.991}$ $\Sigma\frac{(O-E)^2}{E}$ *OR* $\Sigma\frac{O^2}{E} - N = 0.9271$ answers in range 0.90 – 0.95 Not in critical region – no evidence of association between gender and exercise	

EDEXCEL STATISTICS S3 (6685) – JUNE 2004 **PROVISIONAL MARK SCHEME**

Question Number	Scheme	Marks
6a)	$X \sim B(3,1/6)$ bino 3, 1/6	B1 B1 (2)
b)	X Prob Expected freq — prob – must show working and use B(3,p) or may be implied by correct answer	M1
	0 $\left(\frac{5}{6}\right)^3$ 144.68 — expected	M1
	1 $3\times\left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right)$ 86.81	
	2 $3\times\left(\frac{5}{6}\right)\left(\frac{1}{6}\right)^2$ 17.36 — awrt 145,86.8,17.4,1.15/1.16	B2 (-1 ee)
	3 $\left(\frac{1}{6}\right)^3$ 1.15 (1.16)	
	H_0 : Binomial model is a good fit H_1 : Binomial model is not a good fit — both, no ditto	B1
	Amalgamate 3 with another group	M1
	$\alpha = 0.01$ $\nu = 2$; CR $\chi^2 > \underline{9.210}$	B1 ; B1√
	$\Sigma\frac{(O-E)^2}{E}$ OR $\Sigma\frac{O^2}{E} - N = 8.6894...$ answers in range 8.67 – 8.70 or	M1 A1
	Evidence that Binomial is a good model.	A1√ (11)

EDEXCEL STATISTICS S3 (6685) – JUNE 2004 **PROVISIONAL MARK SCHEME**

Question Number	Scheme	Marks
6.c)	Estimate p	B1
	Degrees of freedom reduced by 1	B1 (2)
	Special case	
	Use of B(3,0.192) in part (b)	
	Expected frequencies	M1
	131.8785 94.01242 22.339	M1
	1.769	B0
	H_0 : Binomial model is a good fit H_1 : Binomial model is not a good fit	both, no ditto B1
	Amalgamate 3 with another group	M1
	$\alpha = 0.01$ $\nu = 1$; CR $\chi^2 > 6.635$	B1 ; B1√
	$\Sigma\frac{(O-E)^2}{E}$ OR $\Sigma\frac{O^2}{E} - N$ in range 5.45 -5.50	M1 A1
	Evidence that Binomial is a good model.	A1√ (11)

EDEXCEL STATISTICS S3 (6685) – JUNE 2004 **PROVISIONAL MARK SCHEME**

Question Number	Scheme		Marks
7a)	$E(D) = E(A) - 3E(B) + 4E(C)$		M1
	$= 20$		A1
	$Var(D) = Var(A) + 9Var(B) + 16Var(C)$	Use of a^2Var X Adding 3 Var ie 4 +…	M1 M1
	$=341$		A1
	$P(D < 44) = P\left(z < \frac{44-20}{\sqrt{341}}\right)$	standardising their mean and sd	M1, A1√
	$= P(z < 1.30)$	awrt 1.30	A1
b)	$= \underline{0.9032}$		A1 (9)
	$E(X) = 20$		B1
	$Var(X) = Var(A) + 3Var(B) + 16\ Var(C)$	+ and 16 3 Var (B)	M1 M1
	$= 287$		A1
	$P(X > 0) = P\left(z > \frac{-20}{\sqrt{287}}\right)$	standardising their mean and sd	M1
	$= P(z > -1.18)$	awrt -1.18	A1
	$= 0.8810$		A1 (7)