RECOGNISING ACHIEVEMENT

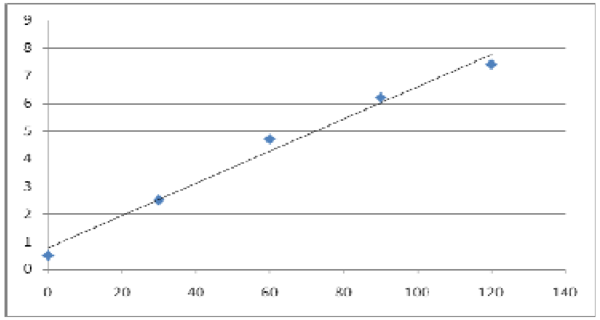
GCE

Mathematics (MEI)

Advanced GCE

Unit **4767:** Statistics 2

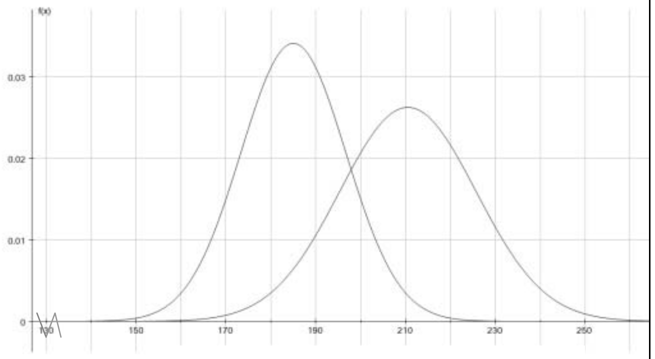
Mark Scheme for June 2011

1 (i)		G1 for axes G1 For values of x G1 for values of y	**3**	Condone axes drawn either way. Axes should show some indication of scale. If not then Max G1 if points 'visibly correct'. If axes are scaled and only one point is incorrectly plotted, allow max G2/3.
1 (ii)	$\bar{x} = 60$, $\bar{y} = 4.26$ $b = \frac{S_{xy}}{S_{xx}} = \frac{1803 - 300 \times 21.3/5}{27000 - 300^2/5} = \frac{525}{9000} = 0.0583$ OR $b = \frac{1803/5 - 60 \times 4.26}{27000/5 - 60^2} = \frac{105}{1800} = 0.0583$ hence least squares regression line is: $y - \bar{y} = b(x - \bar{x})$ $\Rightarrow y - 4.26 = 0.0583\,(x - 60)$ $\Rightarrow y = 0.0583x + 0.76$	B1 for $\bar{x}$ and $\bar{y}$ used appropriately (SOI) M1 for attempt at gradient (b) A1 for 0.0583 **cao** M1 for equation of line A1 FT for complete equation	**5**	B1 for means can be implied by a correct value of b using either method. Allow $\bar{y} = 4.3$ Attempt should be correct – e.g. evidence of either of the two suggested methods should be seen. Allow 0.058 Condone $0.058\dot{3}$ and $\frac{7}{120}$ Dependent on first M1. Values must be substituted to earn M1. Condone use of their b for FT provided $b>0$. Final equation must be simplified. $b = 0.058$ leads to $y = 0.058x + 0.78$
1 (iii)	Regression line plotted on graph The fit is good	G1 G1 E1 for good fit	**3**	Line must pass through their ($\bar{x}$, $\bar{y}$) and y-intercept. E0 for notably inaccurate graphs/lines

1 (iv)	$x = 30 \Rightarrow$ predicted $y = 0.0583\times30 + 0.76 = 2.509$ Residual = 2.5 – 2.509 = –0.009	B1 for prediction M1 for subtraction A1 FT	**3**	Using their **equation** Subtraction can be 'either way' but for the final mark the sign of the residual must be correct. FT sensible equations only – e.g. no FT for $y = 0.071x$ leading to +0.37. [$c = 0.78$ leads to a residual of -0.02]
1 (v)	(*A*) For $x = 45$, $y = 0.0583\times45 + 0.76 = 3.4$ (*B*) For $x = 150$, $y = 0.0583\times150 + 0.76 = 9.5$	M1 for at least one prediction attempted A1 for **both answers** (FT their equation provided their $b>0$)	**2**	Prediction obtained from their equation.
1 (vi)	This is **well** below the predicted value … …suggesting that the model breaks down for larger values of *x*.	E1 for **well** below E1 extrapolation	**2**	Some indication that the value (8.7) is **significantly** below what is expected (9.5) is required for the first E1. Simply pointing out that it is 'below' is not sufficient. The second E1 is available for a suitable comment relating to the model being suitable only for values within the domain of the given points. Allow other sensible comments for either E1. E.g. The data might be better modelled by a curve', 'there may be other factors affecting yield',
			18	

2 (i)	Independently means that the arrival time of each car is unrelated to the arrival time of any other car. Randomly means that the arrival times of cars are not predictable. At a uniform average rate means that the average rate of car arrivals does not vary over time.	E1 E1 E1	**3**	NOTE **Each answer must be 'in context' and 'clear'** Allow sensible alternative wording. SC1 For ALL answers not in context but otherwise correct.
2 (ii)	P(At most 1 car) $= e^{-0.62}\frac{0.62^0}{0!} + e^{-0.62}\frac{0.62^1}{1!}$ $= 0.5379\ldots + 0.3335\ldots = 0.871$	M1 for either M1 for sum of both A1 CAO	**3**	$1.62e^{-0.62}$ Allow 0.8715 not 0.872 or 0.8714 Allow 0.87 without wrong working seen
2 (iii)	New $\lambda = 10 \times 0.62 = 6.2$ P(more than 5 in 10 mins) = 1 – 0.4141 = 0.5859	B1 for mean (SOI) M1 for probability A1 CAO	**3**	Use of $1 - \text{P}(X \le 5)$ with any λ Allow 0.586
2 (iv)	Poisson with mean 37.2	B1 for Poisson B1 for mean 37.2	**2**	Dependent on first B1 Condone P(37.2, 37.2)
2 (v)	Use Normal approx with $\mu = \sigma^2 = \lambda = 37.2$ $\text{P}(X \ge 40) = \text{P}\left(Z > \frac{39.5 - 37.2}{\sqrt{37.2}}\right)$ $= \text{P}(Z > 0.377) = 1 - \Phi(0.377) = 1 - 0.6469$ $= 0.3531$	B1 for Normal (SOI) B1 for parameters B1 for 39.5 M1 for correct use of Normal approximation using correct tail A1 **cao**	**5**	Allow 0.353
			16	

3 (i)	P(Apple weighs at least 220g) $= \text{P}\left(Z > \dfrac{220-210.5}{15.2}\right)$ $= \text{P}(Z > 0.625)$ $= 1 - \Phi(0.625) = 1 - 0.7340$ $= 0.2660$	M1 for standardising M1 for correct structure A1 CAO inc use of diff tables	**3**	Condone numerator reversed but penalise continuity corrections i.e. 1 – Φ(positive z value) Allow 0.266 but not 0.27
3 (ii)	P(All 3 weigh at least 220g) = $0.2660^3 = 0.0188$	M1 A1 FT	**2**	M1 for their answer to part (i) cubed Allow 0.019 and 0.01882
3 (iii)	(*A*) Binomial (100, 0.0188) (*B*) Use a Poisson distribution with $\lambda = 100\times0.0188 = 1.88$ P(At most one) = $e^{-1.88}\dfrac{1.88^0}{0!} + e^{-1.88}\dfrac{1.88^1}{1!}$ $= 0.1525 + 0.2869 = 0.4394$ (*C*) n is large and p is small	B1 for binomial B1 for parameters B1 for Poisson SOI B1 for Poisson mean M1 for either probability M1 for sum of both A1 CAO For 0.44 or better B1	**2** **5** **1**	Second B1 dependent on first B1 FT their answer to part (ii) for second B1 Consistent with $p < 0.1$ from part (iii) (*A*) FT answer to part (ii) with $p < 0.1$ Dependent on both previous B1 marks Allow 0.4395 but not 0.4337 Dependent on use of Poisson in part (iii) *B* Allow n is large and np <10 & n is large and $np \approx npq$
3(iv)(*A*)	$\Phi^{-1}(0.1) = -1.282$ $\dfrac{170-185}{\sigma} = -1.282$ $1.282\,\sigma = 15$ $\sigma = 11.70$	B1 for ±1.282 M1 for correct equation as written o.e. A1 CAO	**3**	Do not allow 1 – 1.282 Allow M1 if different ***z*-value** used Without incorrect working seen. Allow 11.7

3(iv)(*B*)	 Cox's Braeburns	G1 for shape G1 for means, shown explicitly or by scale G1 for lower max height for Braeburns G1 for greater width (variance) for Braeburns	**4**	Ignore labelling of vertical axis. Two intersecting, adjacent Normal curves Means at 185 and 210.5
		TOTAL	**20**	
4(a)(i)	H_0: no association between amount spent and sex H_1: some association between amount spent and sex	B1 for both	**1**	Hypotheses must be the right way round, in context and must not mention 'correlation'.
4(a)(ii)	Expected frequency = 62×102÷200 = 31.62 Contribution = $(34 – 31.62)^2$ / 31.62 = 0.1791	B1 M1 A1 for valid attempt at $(O-E)^2/E$ ***NB Answer given***	**3**	Do not allow 31.6

4(a)(iii)	Refer to X_4^2 Critical value at 5% level = 9.488 3.205 < 9.488 Result is not significant There is insufficient evidence to suggest any association between amount spent and sex. .	B1 for 4 deg of freedom B1 CAO for cv M1 A1 for not significant E1	**5**	Allow p = 0.524 0.524 > 0.05 Conclusion must be stated to earn A1 here. Allow 'do not reject H_0' and condone 'accept H_0' or 'reject H_1'. FT if cv consistent with their d.o.f. Dependent on previous A1 and final comment must be in context and not mention correlation. SC1 for correct final conclusion where previous A1 omitted but M1 awarded.
4 (b)	H_0: $\mu = 400$; H_1: $\mu < 400$ Where μ denotes the population mean (weight of the loaves). $\bar{x} = 396.5$ Test statistic $= \dfrac{396.5 - 400}{5.7/\sqrt{6}} = \dfrac{-3.5}{2.327} = -1.504$ 5% level 1 tailed critical value of z = –1.645 –1.504 > –1.645 so not significant. There is insufficient evidence to reject H_0 There is insufficient evidence to suggest that the true mean weight of the loaves is lower than the minimum specified value of 400 grams.	B1 for H_0 B1 for H_1 B1 for definition of μ B1 for sample mean M1 must include √6 A1FT their sample mean B1 for ±1.645 M1 for sensible comparison leading to a conclusion A1 for conclusion in context	**9**	Hypotheses in words must refer to population mean. Condone numerator reversed for M1 but award A1 only if test statistic of 1.504 is compared with a positive z-value. Dependent on previous M1 FT their sample mean only if hypotheses are correct.
		TOTAL	**18**	

Additional notes re Q1 parts (ii), (iv) and (v)

Part (ii) ‘x on y’ max B1
Part (iv) $x = 16.9y - 12.02$ leads to a prediction of $x = 30.23$ and a residual of -0.23 B1 M1 A1 available.
Part (v) ‘x on y’ not appropriate here so award 0 if ‘x on y’ used.

Additional notes re Q2 parts (i) & (v)
Part (i)
Independent – Allow ‘not linked to’ or ‘no association’ or ‘unrelated’ ‘not affected by’, ‘not connected to’, ‘not influenced by’
DO NOT ACCEPT ‘not together’ or ‘not dependent’
Random – Allow ‘not predictable’ or ‘no pattern’ or ‘could happen at any time’ or ‘not specific time’
Uniform average rate – Allow ‘average (rate) is constant over time’ DO NOT ACCEPT ‘average constant’ or ‘average rate and uniform’ – be generous over defining ‘average’ and/or ‘rate’.
Part (v) If Binomial distribution stated in part (iv), allow B1 B0 B1 M0 A0 max

Additional notes re Q3 part (iii) where $p > 0.1$
(iii) B – as scheme unless a Normal approximation is more suitable ($p > 0.1$). If so, award B1 B1 for Normal and correct parameters. The remaining marks are dependent on both these B1 marks being awarded. M1 for the correct continuity correction ($P(X < 1.5)$) and depM1 for the correct tail but award A0.
(iii) C – ‘n is large and p is not too small’ or ‘$np > 10$’

Additional notes re Q4(b)

σ estimated

sample mean, 7.079… used in place of 5.7, the given value of the population mean, leads to a test statistic of -1.212… This gets M1A0 & the remaining marks are still available.

Critical Value Method

400 – 1.645 × 5.7 ÷ √6… gets M1B1 …= 396.17…gets A1

400 + 1.645 × 5.7 ÷ √6 gets M1B1A0.

396.5 > 396.2 gets M1for sensible comparison (and B1 for 396.5)

A1 still available for correct conclusion in words & context

90% Confidence Interval Method

CI centred on 396.5 (gets B1 for 396.5)

+ or – 1.645 × 5.7 ÷ √6 gets M1 B1

= (392.67, 400.33) A1

contains 400 gets M1

A1 still available for correct conclusion in words & context

Probability Method

Finding P(sample mean < 396.5) = 0.0663 gets M1 A1 (and B1 for 396.5)

0.0663 > 0.05 gets M1 for a sensible comparison if a conclusion is made and also gets the B1 for 0.0663 (to replace the B1 for cv = 1.645).

A1 still available for correct conclusion in words & context.

Condone P(sample mean >396.5) = 0.9337 for M1 and B1 for 0.9337 but only allow A1 if later compared with 0.95 at which point the final M1and A1 are still available

Two-tailed test

Max B1 B0 B1 B1 M1 A1 B1 (for cv = -1.96) M1 A0