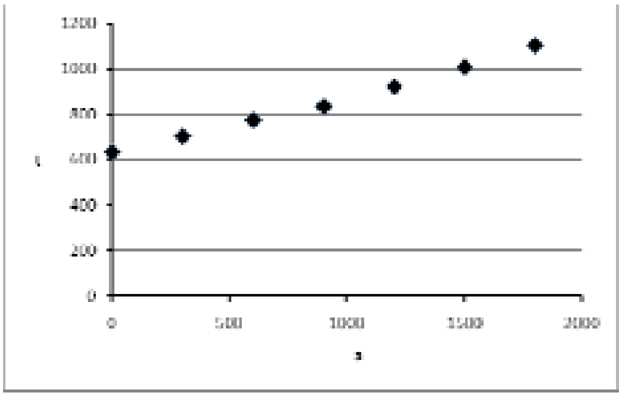
4767 Statistics 2

1	**(i)**		G1 For values of *a* G1 for values of *t* G1 for axes	**[3]**
	(ii)	*a* is independent, *t* is dependent since the values of *a* are not subject to random variation, but are determined by the runways which the pilot chooses, whereas the values of *t* are subject to random variation.	B1 E1dep E1dep	**[3]**
	(iii)	$\bar{a} = 900$, $\bar{t} = 855.2$ $b = \frac{S_{at}}{S_{aa}} = \frac{6037800 - 5987 \times 6300/7}{8190000 - 6300^2/7} = \frac{649500}{2520000} = 0.258$ OR $b = \frac{6037800/7 - 855.29 \times 900}{8190000/7 - 900^2} = \frac{92785}{360000} = 0.258$ hence least squares regression line is: $t - \bar{t} = b(a - \bar{a})$ $\Rightarrow t - 855.29 = 0.258(a - 900)$ $\Rightarrow t = 0.258a + 623$	B1 for $\bar{a}$ and $\bar{t}$ used (SOI) M1 for attempt at gradient (*b*) A1 for 0.258 **cao** M1 for equation of line A1 FT for complete equation	**[5]**
	(iv)	(*A*) For *a* = 800, predicted take–off distance $= 0.258 \times 800 + 623 = 829$ (*B*) For *a* = 2500, predicted take–off distance $= 0.258 \times 2500 + 623 = 1268$ Valid relevant comments relating to the predictions such as: First prediction is interpolation so should be reasonable Second prediction is extrapolation and may not be reliable	M1 for at least one prediction attempted A1 for both answers (FT their equation if $b>0$) E1 (first comment) E1 (second comment)	**[4]**
	(v)	$a = 1200 \Rightarrow$ predicted $t = 0.258 \times 1200 + 623 = 933$ Residual = 923 – 933 = –10 The residual is negative because the observed value is less than the predicted value.	M1 for prediction M1 for subtraction A1 FT E1	**[4]**
			Total	**[19]**

2	**(i)**	P(1 of 10 is faulty) $= \binom{10}{1} \times 0.02^1 \times 0.98^9 = 0.1667$	M1 for coefficient M1 for probabilities A1	**[3]**
	(ii)	n is large and p is small	B1, B1 Allow appropriate numerical ranges	**[2]**
	(iii)	$\lambda = 150 \times 0.02 = 3$ (*A*) $\mathrm{P}(X = 0) = \mathrm{e}^{-3}\frac{3^0}{0!} = 0.0498$ (3 s.f.) or from tables $= 0.0498$	B1 for mean (soi) M1 for calculation or use of tables A1	**[3]**
		(*B*) Expected number = 3 Using tables: $\mathrm{P}(X > 3) = 1 - \mathrm{P}(X \le 3)$ $= 1 - 0.6472 = 0.3528$	B1 expected no = 3 (soi) M1 A1	**[3]**
	(iv)	(*A*) Binomial(2000,0.02)	B1 for binomial B1 for parameters	**[2]**
		(*B*) Use Normal approx with $\mu = np = 2000 \times 0.02 = 40$ $\sigma^2 = npq = 2000 \times 0.02 \times 0.98 = 39.2$ $\mathrm{P}(X \le 50) = \mathrm{P}\left(Z \le \frac{50.5 - 40}{\sqrt{39.2}}\right)$ $= \mathrm{P}(Z \le 1.677) = \Phi(1.677) = 0.9532$ NB Poisson approximation also acceptable for full marks	B1 B1 B1 for continuity corr. M1 for probability using correct tail A1 CAO	**[5]**
			Total	**[18]**

3	**(i)**	(*A*) P(X < 50) $= \mathrm{P}\left(Z < \frac{50-45.3}{11.5}\right)$ = P(Z < 0.4087) = Φ(0.4087) = 0.6585	M1 for standardising M1 for correct structure of probability calc' A1 CAO inc use of diff tables NB When a candidate's answers suggest that (s)he appears to have neglected to use the difference column of the Normal distribution tables penalise the first occurrence only	**[3]**
		(*B*) P(45.3 < X < 50) = 0.6585 – 0.5 =0.1585	M1 A1	**[2]**
	(ii)	From tables Φ^{-1}(0.9) = 1.282 $\frac{k-45.3}{11.5} = 1.282$ k = 45.3 + 1.282 × 11.5 = 60.0	B1 for 1.282 seen M1 for equation in k A1 CAO	**[3]**
	(iii)	P(score = 111) =P(110.5 < Y < 111.5) $= \mathrm{P}\left(\frac{110.5-100}{15} < Z < \frac{111.5-100}{15}\right)$ = P(0.7< Z < 0.7667) = Φ(0.7667) – Φ(0.7) = 0.7784 – 0.7580 = 0.0204	B1 for both continuity corrections M1 for standardising M1 for correct structure of probability calc' A1 CAO	**[4]**
	(iv)	From tables, Φ^{-1}(0.3) = –0.5244, Φ^{-1}(0.8) = 0.8416 22 = μ + 0.8416 σ 15 = μ – 0.5244 σ 7 = 1.3660 σ σ = 5.124, μ = 17.69	B1 for 0.5244 or 0.8416 seen M1 for at least one equation in z, μ & σ A1 for both correct M1 for attempt to solve two appropriate equations A1 CAO for both	**[5]**
			TOTAL	**[17]**

4	**(i)**	H_0: no association between size of business and recycling service used. H_1: some association between size of business and recycling service used.	B1 for both	**[1]**
	(ii)	Expected frequency = 78/285 × 180 = 49.2632 Contribution = $(52 – 49.2632)^2$ / 49.2632 = 0.1520	M1 A1 M1 for valid attempt at $(O\text{-}E)^2/E$ A1 ***NB Answer given*** Allow 0.152	**[4]**
	(iii)	Test statistic $X^2 = 0.6041$ Refer to χ_2^2 Critical value at 5% level = 5.991 Result is not significant There is no evidence to suggest any association between size of business and recycling service used. NB if H_0 H_1 reversed, or 'correlation' mentioned in part (i), do not award B1in part (i) or E1 in part (iii).	B1 B1 for 2 deg of f(seen) B1 CAO for cv B1 for not significant E1	**[5]**
	(iv)	H_0: $\mu = 32.8$; H_1: $\mu < 32.8$ Where μ denotes the population mean weight of rubbish in the bins. Test statistic = $\dfrac{30.9-32.8}{3.4/\sqrt{50}} = -\dfrac{1.9}{0.4808} = -3.951$ 5% level 1 tailed critical value of z = –1.645 –3.951 < –1.645 so significant. There is sufficient evidence to reject H_0 There is evidence to suggest that the weight of rubbish in dustbins has been reduced.	B1 for use of 32.8 B1 for both correct B1 for definition of μ M1 must include √50 A1 B1 for ±1.645 M1 for sensible comparison leading to a conclusion A1 for conclusion in words in context	**[8]**
			TOTAL	**[18]**