Question			Answer	Marks	Guidance	
1	(i)		**EITHER** $S_{xy} = \Sigma xy - \frac{1}{n}\Sigma x \Sigma y = 600.41 - \frac{1}{10} \times 113.69 \times 52.81 = 0.01311$ $S_{xx} = \Sigma x^2 - \frac{1}{n}(\Sigma x)^2 = 1292.56 - \frac{1}{10} \times 113.69^2 = 0.01839$ $S_{yy} = \Sigma y^2 - \frac{1}{n}(\Sigma y)^2 = 278.91 - \frac{1}{10} \times 52.81^2 = 0.02039$ $r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \frac{0.01311}{\sqrt{0.01839 \times 0.02039}} = 0.677$ **OR** $\text{cov}(x,y) = \frac{\sum xy}{n} - \bar{x}\bar{y} = 600.41/10 - 11.369 \times 5.281 = 0.001311$ $\text{rmsd}(x) = \sqrt{\frac{S_{xx}}{n}} = \sqrt{(0.01839/10)} = \sqrt{0.001839} = 0.04288$ $\text{rmsd}(y) = \sqrt{\frac{S_{yy}}{n}} = \sqrt{(0.02039/10)} = \sqrt{0.002039} = 0.04516$ $r = \frac{\text{cov}(x,y)}{rmsd(x)rmsd(y)} = \frac{0.001311}{0.04288 \times 0.04516} = 0.677$	M1 M1 A1 M1 A1 M1 M1 A1 M1 A1 **[5]**	For method for S_{xy} For method for at least one of S_{xx} or S_{yy} For at least one of S_{xy}, S_{xx} or S_{yy} correct For fully correct structure of r For answer rounding to 0.68 For method for cov (x,y) For method for at least one msd or rmsd For at least one of cov (x,y), msd or rmsd correct For fully correct structure of r For answer rounding to 0.68 Methods mixed – max M0M1A1M0A0	
1	(ii)		H_0: $\rho = 0$ H_1: $\rho \neq 0$ (two-tailed test) where ρ is the population correlation coefficient For $n = 10$, 10% critical value = 0.5494	B1 B1 B1	For H_0, H_1 in symbols. Hypotheses in words must refer to population. Do not allow alternative symbols unless clearly defined as the population correlation coefficient. For defining ρ. Condone omission of "population" if correct notation ρ is used, but if ρ is defined as the **sample** correlation coefficient then award **B0**. CAO Note that critical values for a one-tailed test at the 10% level are not available in tables.	

			Since 0.677 > 0.5494 the result is significant. (Thus we have sufficient evidence to) reject H_0 There is sufficient evidence at the 10% level to suggest that there is correlation between times for the first and last sections.	M1 A1* E1dep* **[6]**	For sensible comparison leading to a conclusion provided that $\|r\| < 1$. The comparison can be in the form of a diagram as long as it is clear and unambiguous. Sensible comparison: e.g. 0.677 > 0.5494 is 'sensible' whereas 0.677 > - 0.5494 is 'not sensible'. Reversed inequality sign e.g. 0.677 < 0.5494 etc. gets max M1 A0. For reject H_0 o.e. FT their *r* and critical value from 10% 2-tail column. For correct, non-assertive conclusion in context. Allow '*x* and *y*' for context. E0 if H_0 and H_1 not stated, reversed or mention a value other than zero for ρ in H_0. Do not allow 'positive correlation' or 'association'	
1	**(iii)**		The underlying population must have a bivariate Normal distribution. The points in the scatter diagram should have a roughly elliptical shape.	B1 E1 **[2]**	Condone "bivariate Normal distribution", "underlying bivariate Normal distribution", but **do not allow** "the **data** have a bivariate Normal distribution" Condone 'oval' or suitable diagram	
1	**(iv)**		The hypothesis test has shown that there appears to be correlation. However it could be that there is a third causal factor	E1 E1 **[2]**	For relevant comment relating to the test result or positive value of *r* in supporting (unless FT leads to not supporting) the commentator's suggestion. Or correlation does not imply causation. There may be a third factor. For questioning the use of the word 'must' Allow any two suitable, statistically based comments.	
1	**(v)**	(*A*)	Yes because the critical value at the 1% level is 0.7646 which is larger than the test statistic	B1* E1dep* **[2]**	B1 for 0.7646 seen E1 for comment consistent with their (ii) provided $r\| < 1$	

1	**(v)**	(*B*)	One advantage of a 1% level is that one is less likely to reject the null hypothesis when it is true. One disadvantage of a 1% level is that one is more likely to accept the null hypothesis when it is false.	E1 E1 **[2]**	o.e. Wording must be clear. o.e.	
2	**(i)**		Binomial(1200,1/300)	B1 B1dep **[2]**	For binomial. For parameters Allow B(1200, 1/300) and B(1200, 0.00333)	
2	**(ii)**		Because n is large and p is small	E1, E1 **[2]**	Allow n is large and $np < 10$. Allow "sample is large" for n is large and "mean ≈ variance" for "p is small"	
2	**(iii)**		$\lambda = 1200 \times 1/300 = 4$ (*A*) $P(X = 1) = e^{-4}\frac{4^1}{1!} = 0.0733$ (3 s.f.) or from tables $= 0.0916 - 0.0183 = 0.0733$ (*B*) Using tables: $P(X > 4) = 1 - P(X \leq 4)$ $= 1 - 0.6288 = 0.3712$	B1 M1 A1 M1 A1 **[5]**	For λ FT their p For attempt to find $P(X = 1)$ using Poisson p.d.f. or tables Allow answers which round to 0.073 www. FT their λ (= np). No FT for $\lambda = 1/300$. For finding $1 - P(X \leq 4)$ CAO For answers rounding to 0.371 www	
2	**(iv)**		$\mu = 80$ $\sigma^2 = 80$	B1 B1 **[2]**	If symbols/words used then they must be correct. Allow σ^2 rounding to 79.7 from original binomial. FT their λ (= np)	
2	**(v)**	(*A*)	$P(Y \geq 90) = P\left(Z \geq \frac{89.5 - 80}{\sqrt{80}}\right)$ $= P(Z > 1.062) = 1 - \Phi(1.062)$ $= 1 - 0.8559 = 0.1441$	B1 M1 A1cao **[3]**	For correct continuity correction. For probability using correct tail and structure (condone omission of c.c.) $\sigma^2 = 79.73$ leads to $P(Z > 1.064)$ $\sigma^2 = 79.73$ leads to $1 - 0.8563 = 0.1437$. Allow 0.144 www. NOTE 0.1441 from B(24000, 1/300) gets 0/3	

2	**(v)**	(*B*)	P($Y \leq k$) > 0.05 From tables Φ^{-1}(0.05) = -1.645 $\frac{(k+0.5)-80}{\sqrt{80}}$ = -1.645 k + 0.5 = 80 – (1.645 × √80) = 65.29 k > 64.79 So least value of k = 65	B1 M1 A1 A1 **[4]**	For ±1.645 seen For correct equation in k seen or equivalent – e.g. allow +1.645 used if numerator reversed. FT their μ, σ^2 and z-value. Condone omission of, or incorrect, continuity correction. A1 for 65.29 or 64.79 or 65.79 (σ^2 = 79.73 leads to 65.31 or 64.81 or 65.81) Allow 3s.f. For rounding **64.79 or 64.81 up to give k = 65**. See additional notes for alternative method	
3	**(i)**		P($X \geq 750$) = $P\left(Z \geq \frac{750-751.4}{2.5}\right)$ = P(Z > -0.56) = Φ(0.56) = 0.7123	M1 M1 A1 **[3]**	For standardizing For correct structure (M0 if continuity correction used) CAO Allow 0.712 www	
3	**(ii)**		P(all 6 at least 750ml) = 0.7123^6 = 0.1306	M1 A1 **[2]**	For (their answer to part (i))^6 FT 3s.f.	
3	**(iii)**		P(Y=0) = $\binom{25}{0} \times 0.8694^{25}$ (= 0.0302) P(Y=1) = $\binom{25}{1} \times 0.8694^{24} \times 0.1306$ (= 0.1135) P(Y=0) + P(Y=1) = 0.144 P($Y \geq 2$) = 1 – 0.144 =0.856	M1 M1 M1dep A1 **[4]**	For using Binomial(25, p) with their p from part (ii) For correct structure of either P(Y = 0) or P(Y = 1) with their p from part (ii) M0 if p and q reversed For 1 – sum of both probabilities CAO	

3	**(iv)**		$P(Z < \frac{750-\mu}{2.5}) = 0.02$ $\Phi^{-1}(0.02) = -2.054$ $\frac{750-\mu}{2.5} = -2.054$ $\mu = 750 + 2.054 \times 2.5$ $= 755.1$	 B1 M1 M1 A1 **[4]**	 For ±2.054 seen. Allow ±2.055 For correct equation as seen or equivalent. FT σ = √2.5. M0 if c.c. used. For correctly rearranging their equation (if 750 used in numerator) for μ, FT their *z* cao Condone 755 or 5 s.f. rounding to 755.1 www	
3	**(v)**		$P(Z < \frac{750-751.4}{\sigma}) = 0.02$ $\frac{750-751.4}{\sigma} = -2.054$ $\sigma = \frac{-1.4}{-2.054}$ $= 0.682$	 M1 M1 A1 **[3]**	 For correct equation as seen or equivalent For correctly rearranging their equation (if 750 used in numerator) for σ unless this leads to σ < 0 cao Allow answers rounding to 0.68 www	
3	**(vi)**		Probably easier to change the mean (as reducing the standard. deviation would require a much more accurate filling process). However increasing the mean would result in fewer bottles being filled overall and so less profit for the owners, so **reducing** the **standard deviation** would be **preferable** to the vineyard owners.	E1 E1 **[2]**	 For "preferable to reduce the standard deviation" with valid reason.	
4	**(a)**	**(i)**	Expected frequency = 67/150 × 57 = 25.46 Contribution = $(34 - 25.46)^2 / 25.46$ = 2.8646	B1 M1 A1 **[3]**	For 25.46 For valid attempt at $(O-E)^2/E$ Correct values used to give answer which rounds to 2.8646 ***NB Answer given***	

4	**(a)**	**(ii)**	H_0: no association between type of cake and classification of person. H_1: some association between type of cake and classification of person. Test statistic $X^2 = 12.86$ Refer to X^2_3 Critical value at 1% level = 11.34 Result is significant There is evidence to suggest association between type of cake and classification of person. NB if H_0 H_1 reversed, omitted or 'correlation' mentioned, do not award first B1 or final E1	B1 B1 B1 B1 E1 **[5]**	For both hypotheses in context For 3 degrees of freedom CAO For cv. No FT from here if wrong/omitted For significant For correct, non-assertive conclusion, in context.	
4	**(b)**		$\bar{x} = 4.995$ H_0: $\mu = 5$ H_1 : $\mu < 5$ Where μ denotes the **mean** content of the bags of flour (in the population) Test statistic $= \dfrac{4.995 - 5.0}{0.0072/\sqrt{8}} = \dfrac{-0.005}{0.002546} = -1.964$ Lower 5% level 1 tailed critical value of $z = -1.645$ $-1.964 < -1.645$ so significant. There is sufficient evidence to reject H_0 There is sufficient evidence to suggest that the average contents of bags is less than 5kg.	B1 B1 B1 B1 M1* A1 B1* M1 dep* A1 **[9]**	For 4.995 seen For use of 5 in hypotheses. For both correct. Hypotheses in words must refer to population. Do not allow alternative symbols unless clearly defined as the population mean. For definition of μ in context. Condone omission of "population" if correct notation μ is used, but if μ is defined as the **sample** mean then award **B0**. must include $\sqrt{8}$ FT their $\bar{x}$. Allow +1.964 only if later compared with +1.645 For –1.645 No FT from here if wrong. Must be –1.645 unless it is clear that absolute values are being used. For sensible comparison with correct c.v. leading to a conclusion. For non-assertive conclusion in words and in context. No FT here. See additional notes.	

ADDITIONAL NOTES REGARDING QUESTION 2 (v) *B*

M1 for using a trial and improvement method with N(80,80) or N(80, 79.73) to find P($Y \leq k$) for any k. The distribution being used needs to be made clear.

A1 for P($Y \leq 66$) = 0.0587… (0.0584… from $\sigma^2 = 79.73$) or P($Y \leq 65$) = 0.0467… (0.0464… from $\sigma^2 = 79.73$)
A1 for both
Final A1 not available if 66 and 65 used

Or
A1 for P($Y \leq 65.5$) = 0.0524… (0.0521… from $\sigma^2 = 79.73$) or P($Y \leq 64.5$) = 0.0415… (0.0412… from $\sigma^2 = 79.73$)
A1 for both
A1 for least value of $k = 65$, dependent on previous two A marks earned.

ADDITIONAL NOTES REGARDING QUESTION 4 (b)

Critical Value Method
$5 - 1.645 \times 0.0072 \div \sqrt{8}$ gets M1*B1*
= 4.9958… gets A1
4.995 < 4.99581.. gets M1dep* for sensible comparison
A1 still available for correct conclusion in words & context

"Confidence Interval" Method
$4.995 + 1.645 \times 0.0072 \div \sqrt{8}$ gets M1* B1*
= 4.9991.. gets A1
NOTE that the final M1dep* A1 available only if 1.645 used.
5 > 4.9991… gets M1
A1 still available for correct conclusion in words & context

Probability Method
Finding P(sample mean < 4.995) = 0.0248 gets M1* A1 B1
0.0248 < **0.05*** gets M1dep* for a sensible comparison if a conclusion is made.
A1 available for a correct conclusion in words & context.
Condone P(sample mean > 4.995) = 0.9752 for M1 but only allow A1 B1 if later compared with 0.95, at which point the final M1and A1 are still available

ADDITIONAL NOTE REGARDING OVER-SPECIFICATION OF ANSWERS

Over-specification by providing final answers correct to 5 or more significant figures will be penalised. When this applies, candidates may lose no more than 2 marks per question and no more than 4 marks in total. The only exception to this rule is in Question 3 part (iv) – see guidance note.