Question			Answer	Marks	Guidance
1	(i)		A paired sample is used in this context in order to eliminate any effects due to the surfaces used.	E1 **[1]**	Must refer to (differences between) surfaces.
1	(ii)		A *t* test might be used since … … the sample is small and … the population variance is not known (it must be estimated from the data). Must assume: Normality of population … … of <u>differences</u>.	 E1 E1 B1 B1 **[4]**	 Allow use of "σ", otherwise insist on "population". Allow "underlying" or "distribution" to imply "population".
1	(iii)		H_0: $\mu_D = 0$ H_1: $\mu_D > 0$ Where μ_D is the (population) mean reduction/difference in drying time.	B1 B1	Both. Accept alternatives e.g. $\mu_D < 0$ for H_1, or $\mu_B - \mu_A$ etc provided adequately defined. Hypotheses in words only must include "population". Do NOT allow "$\bar{X} = \ldots$" or similar. unless $\bar{X}$ is clearly and explicitly stated to be a <u>population</u> mean. For adequate verbal definition. Allow absence of "population" if correct notation μ is used.
			<u>MUST</u> be PAIRED COMPARISON *t* test. Differences (reductions) (before – after) are: 0.7 0.7 0.2 –0.3 0.8 –0.1 0.3 –0.1 0.1 0.5		 Allow "after – before" if consistent with alternatives above.
			$\bar{x} = 0.28$ $s_{n-1} = 0.3852(84)$ $(s_{n-1}^2 = 0.1484(44))$	B1	Do not allow $s_n = 0.3655$ $(s_n^2 = 0.1336)$
			Test statistic is $\dfrac{0.28 - 0}{\frac{0.3853}{\sqrt{10}}}$	M1	Allow c's $\bar{x}$ and/or s_{n-1}. Allow alternative: 0 + (c's 1.833) × $\frac{0.3853}{\sqrt{10}}$ (= 0.2233) for subsequent comparison with $\bar{x}$. (Or $\bar{x}$ – (c's 1.833) × $\frac{0.3853}{\sqrt{10}}$ (= 0.0566) for comparison with 0.)
			= 2.298.	A1	c.a.o. but ft from here in any case if wrong. Require 3/4 sf; condone up to 6. Use of $0 - \bar{x}$ scores M1A0, but ft.
			Refer to t_9.	M1	No ft from here if wrong. $P(t > 2.298) = 0.02357$.
			Single-tailed 5% point is 1.833.	A1	No ft from here if wrong.
			Significant.	A1	ft only c's test statistic.
			Seems mean drying time has fallen.	A1 **[9]**	ft only c's test statistic. "Non-assertive" conclusion in context to include "on average" oe.

Question			Answer	Marks	Guidance
1	(iv)		CI is given by $0.28 \pm$ 2.262 $\times \frac{0.3853}{\sqrt{10}}$ $= 0.28 \pm 0.2756 = (0.0044, 0.5556)$	M1 B1 M1 A1 **[4]**	Allow c's $\bar{x}$. Allow c's s_{n-1}. c.a.o. Must be expressed as an interval. Require 3/4 dp; condone 5. If the final answer is centred on a negative sample mean then do not award the final A mark. ZERO/4 if not same distribution as test. Same wrong distribution scores maximum M1 B0 M1 A0. Recovery to t_9 is OK.
2	(a)	(i)	For example, need to take a sample because the population might be too large for it to be sensible to take a complete census. Because the sampling process might be destructive.	E1 E1 **[2]**	Reward 1 mark each for any two distinct, sensible points.
2	(a)	(ii)	For example Sample should be unbiased. Sample should be representative (of the population).	E1 E1 **[2]**	Reward 1 mark each for any two distinct, sensible points that the sample/data should be fit for purpose. Further examples include: data should not be distorted by the act of sampling; data should be relevant.
2	(a)	(iii)	A random sample … enables proper statistical inference to be undertaken …… because we know the probability basis on which it has been selected	E2 **[2]**	Award E2, 1, 0 depending on the quality of response.
2	(b)	(i)	A Wilcoxon signed rank test might be used when nothing is known about the distribution of the background population. Must assume symmetry (about the median).	E1 E1 **[2]**	Do not allow "sample", or "data" unless it clearly refers to the population. Do not allow if "Normality" forms part of the assumption.

Question			Answer	Marks	Guidance
2	(b)	(ii)	H_0: $m = 28.7$ H_1: $m > 28.7$ where m is the population median	B1 B1	Both. Accept hypotheses in words. Adequate definition of m to include "population".
			Speeds / −28.7 / Rank of |diff|: 32.0, 3.3, 8 29.1, 0.4, 3 26.1, −2.6, 6 35.2, 6.5, 12 34.4, 5.7, 11 28.6, −0.1, 1 32.3, 3.6, 9 28.5, −0.2, 2 27.0, −1.7, 5 33.3, 4.6, 10 28.2, −0.5, 4 31.9, 3.2, 7	M1 M1 A1	for subtracting 28.7. for ranks. ft if ranks wrong. If candidate has tied ranks then penalise A0 here but ft from here.
			$W_- = 1 + 2 + 4 + 5 + 6 = 18$ Refer to Wilcoxon single sample tables for $n = 12$. Lower 5% point is 17 (or upper is 61 if 60 used). Result is not significant. No evidence to suggest that the median speed has increased.	B1 M1 A1 A1 A1 **[10]**	($W_+ = 3 + 7 + 8 + 9 + 10 + 11 + 12 = 60$) No ft from here if wrong. ie a 1-tail test. No ft from here if wrong. ft only c's test statistic. ft only c's test statistic. "Non-assertive" conclusion in context to include "on average" oe.
3			$S \sim N(11.07,\ 2.36^2)$ $C \sim N(57.33,\ 8.76^2)$ $R \sim N(24.23,\ 3.75^2)$		When a candidate's answers suggest that (s)he appears to have neglected to use the difference columns of the Normal distribution tables, penalise the first occurrence only.
	(i)		$P(10 < S < 13)$ $= P\left(\frac{10-11.07}{2.36} < Z < \frac{13-11.07}{2.36}\right)$ $= P(-0.4534 < Z < 0.8178)$ $= 0.7931 - (1 - 0.6748)$ $= 0.4679$	M1 A1 A1 **[3]**	For standardising. Award once, here or elsewhere. Cao Accept 0.468(0), 0.4681, 0.4682, but not 0.4683.

Question			Answer	Marks	Guidance
3	(ii)		Want P($R > S + 10$) i.e. P($R - S > 10$) R – S ~ N(24.23 – 11.07 = 13.16, $3.75^2 + 2.36^2 = 19.6321$) P(this > 10) = P($Z > \frac{10 - 13.16}{\sqrt{19.6321}} = -0.7132$) = 0.7621	M1 B1 B1 A1 **[4]**	Allow $S - R$ provided subsequent work is consistent. Mean. Variance. Accept sd = √19.6321 = 4.4308... cao
3	(iii)		Want P($S + R > ⅔C$) i.e. P($S + R - ⅔C > 0$) $S + R - ⅔C$ ~ N(11.07 + 24.23 – ⅔ × 57.33 = –2.92, $2.36^2 + 3.75^2 + (⅔ \times 8.76)^2 = 53.7377$) P(this > 0) = P($Z > \frac{0 - (-2.92)}{\sqrt{53.7377}} = 0.3983$) = 1 – 0.6548 = 0.3452	M1 B1 B1 A1 **[4]**	Allow $⅔L - (S + R)$ provided subsequent work is consistent. Mean Variance. Accept sd = √53.7377 = 7.3306... cao
3	(iv)		$\bar{x} = 98.484$, $s_{n-1} = 10.1594$ CI is given by 98.484 ± 2.201 $\times \frac{10.1594}{\sqrt{12}}$ = 98.484 ± 6.455 = (92.03, 104.94)	B1 M1 B1 M1 A1 **[5]**	Do not allow s_n = 9.7269. ft c's $\bar{x}$ ±. From t_{11}. ft c's s_{n-1}. cao Must be expressed as an interval. Require 1 or 2 dp; condone 3dp.
3	(v)		Normality is unlikely to be reasonable – times could well be (positively) skewed. Independence is unlikely to be reasonable – e.g. a competitor who is fast in one stage may well be fast in all three.	E1 E1 **[2]**	Discussion required. Accept any reasonable point. Accept "reasonable" provided an adequate explanation is given. Discussion required. Accept any reasonable point. This is independence between stages for a particular competitor, not between competitors.

Question			Answer	Marks	Guidance
4	(i)		H_0: The model for the number of callouts fits the data H_1: The model for the number of callouts does not fit the data.	B1 B1	Do not allow "Data fit the model" o.e for either hypothesis.

Obs'd frequency	145	79	22	6	3	0
Exp'd frequency	139.947	83.968	25.190	5.038	0.756	0.101

Question			Answer	Marks	Guidance
			Merge last 3 cells. Obs 9 Exp 5.895 $X^2 = 0.1824 + 0.2939 + 0.4040 + 1.6355$ $= 2.515(8)$ Refer to χ_2^2. Upper 5% point is 5.991. Not significant. Suggests it is reasonable to suppose that the model fits the data.	M1 M1 A1 M1 A1 A1 A1 **[9]**	 Calculation of X^2. Cao Require 3/4 sf; condone up to 6. Allow correct df (= cells – 2) from wrongly grouped table and ft. Otherwise, no ft if wrong. $P(X^2 > 2.5158) = 0.2842$. No ft from here if wrong. ft only c's test statistic. ft only c's test statistic. "Non-assertive" conclusion in words (+context). Do not allow "Data fit model" o.e.
4	(ii)		Mean = 5/3 $\therefore \lambda = 0.6$	B1 **[1]**	
4	(iii)		$F(t) = \int_0^t 0.6e^{-0.6x}dx$ $= \left[-e^{-0.6x}\right]_0^t$ $= \left(-e^{-0.6t} - (-e^0)\right) = 1 - e^{-0.6t}$	M1 A1 A1 **[3]**	Correct integral with limits (which may be implied subsequently). Allow use of "+ c" accompanied by a valid attempt to evaluate it. Correctly integrated. Limits used or c evaluated correctly. Accept unsimplified form. If final answer is given in terms of λ then allow max M1A1A0.
4	(iv)		$P(T > 1) = 1 - F(1)$ $= 1 - \left(1 - e^{-0.6}\right) = 0.5488$	M1 A1 **[2]**	ft c's F(t). cao Allow any exact form of the correct answer.
4	(v)		$F(m) = \frac{1}{2} \quad \therefore 1 - e^{-0.6m} = \frac{1}{2}$ $\therefore e^{-0.6m} = \frac{1}{2} \quad \therefore -0.6m = -\ln 2 \quad \therefore m = \frac{\ln 2}{0.6}$ $m = 1.155$ (days)	M1 M1 A1 **[3]**	Use of definition of median. Allow use of c's F(t). Convincing attempt to rearrange to "m = …", to include use of logs. Cao obtained only from the correct F(t). Must be evaluated. Require 2 to 4 sf; condone 5.