

## Mock Paper Mark Scheme

### Advanced Subsidiary/Advanced GCE General Certificate of Education

Subject **STATISTICS**

Paper No. **Mock S4**

Question number	Scheme	Marks
<b>1.</b>	<p><math>H_0: \mu = 100, H_1: \mu &gt; 100</math></p> $\bar{x} = \frac{710.9}{7} = 101.5571\dots; s^2 = \frac{72219.45 - \frac{(710.9)^2}{7}}{6}$ <p><math>s^2 = 3.746\dots</math> or <math>s = 1.9355</math></p> <p>test statistic <math>t = \frac{101.557 - 100}{\frac{1.936}{\sqrt{7}}} = \text{awrt } 2.13</math></p> <p><math>t_6</math> 5% 1-tail critical value = 1.943</p> <p>Significant result. Reject <math>H_0</math>, there is evidence that the mean weight is more than 100g.</p>	<p>B1</p> <p>B1; M1</p> <p>A1</p> <p>M1 A1</p> <p>B1      ✓</p> <p>A1      (8)</p> <p style="text-align: right;"><b>(8 marks)</b></p>
<b>2.</b>	<p><math>D = \text{dry—wet} \quad H_0: \mu_D = 0, H_1: \mu_D \neq 0</math></p> <p><math>d: 0.6, -1, -1.9, -1.4, -1.3, 0.5, -1.6, -0.6, -1.8</math></p> $\bar{d}: -\frac{8.5}{9} = -0.94, \quad s_d^2 = \frac{15.03 - 9 \times (\bar{d})^2}{8} = 0.87527\dots$ $t = \frac{-0.94}{\frac{s_d}{\sqrt{9}}} = \text{awrt } -3.03$ <p><math>t_8</math> 2-tail 1% critical value = 3.355</p> <p>Not significant – insufficient evidence of a difference between mean strength</p>	<p>B1</p> <p>M1</p> <p>A1, A1</p> <p>M1 A1</p> <p>B1</p> <p>A1      ✓ (8)</p> <p style="text-align: right;"><b>(8 marks)</b></p>

Question number	Scheme	Marks
3. (a)	$H_0: p = 0.3$ (or 0.7) $H_1: p < 0.3$ (or $> 0.7$ )	B1 (1)
(b)	Let $X$ = number who contract virus. Under $H_0$ $X \sim B(30, 0.3)$ $P(\text{Type I error}) = P(X < 6   p = 0.30) = P(X \leq 5) = 0.0766$	M1 A1 (2)
(c)	(i) Power = $P(Y \leq 5   Y \sim B(30, 0.2)) = 0.4275$	M1 A1
(d)	(ii) Power = $P(Y \leq 5   Y \sim B(30, 0.1)) = 0.9268$	A1 (3)
(e)	Let $C$ = number who contract virus. Under $H_0$ $C \sim B(50, 0.3)$ We require $c$ such that $P(C \leq c) \approx 0.05$ $P(C \leq 10) = 0.0789$ , $P(C \leq 9) = 0.0402$ $\therefore$ critical region is $C \leq 9$	M1 A1 (2)
(f)	Size = 0.0402	B1 (1)
(g)	(i) Power = $P(D \leq 9   D \sim B(50, 0.2)) = 0.4437$	B1
(h)	(ii) Power = $P(D \leq 9   D \sim B(50, 0.1)) = 0.9755$	B1 (2)
(i)	Advantage: second test is more powerful Disadvantage: second test involves greater sample size, $\therefore$ more expensive or takes longer	B1 B1 (2) <b>(13 marks)</b>

Question number	Scheme	Marks
4. (a)	$\bar{x} = \frac{847.89}{10} = 84.79; \quad s_x^2 = \frac{103712.6151 - (847.89)^2 / 10}{9}$	B1
	$s_x^2 = 3535.6522\dots$	B1
	<p>or <math>s_x = 59.461\dots</math></p>	
	<p style="text-align: right;">2.262</p>	B1
	<p>95% confidence interval for <math>\mu = 84.79 \pm 2.262 \times \frac{59.461}{\sqrt{10}} = (42.25, 127.33)</math></p> <p style="text-align: right;">accept (42.3, 127.3)</p>	M1, A1, A1 (6)
(b)	<p>95% confidence interval <math>\bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}}</math></p>	
	<p>so chief accountant requires <math>1.96 \frac{\sigma}{\sqrt{n}} &lt; 10</math></p>	M1 A1
(c)	<p>i.e. <math>\frac{\sigma^2}{n} &lt; \left(\frac{10}{1.96}\right)^2 = 26.0308\dots = 26.03</math> (2 d.p.)</p>	A1 cso (3)
	<p>Require the upper confidence limit of 98% confidence interval for <math>\sigma^2</math></p> <p><math>\chi_9^2 = 2.088</math>; i.e. <math>\frac{9s^2}{\sigma^2} &gt; 2.088, \Rightarrow \sigma^2 &lt; 15239.88\dots</math> awrt 15240</p>	B1; M1, A1 (3)
(d)	<p>Substitute into part (b), <math>n &gt; \frac{15240}{26.03} \Rightarrow n = 586</math></p>	M1, A1 (2)
		<b>(14 marks)</b>

Question number	Scheme	Marks
5. (a)	(i) $H_0: \sigma_C^2 = \sigma_N^2$ , $H_1: \sigma_C^2 > \sigma_N^2$	B1
	$\frac{s_C^2}{s_N^2} = \frac{5.7^2}{3.5^2} = 2.652\dots$ ; $F_{8,9}$ (5%) critical value = 3.23	M1; B1
	Not significant so do not reject $H_0$ – insufficient evidence that variance using conventional method is greater	A1      ✓ (4)
	(ii) $H_0: \mu_N = \mu_C$ , $H_1: \mu_N > \mu_C$	B1
	$s^2 = \frac{8 \times 5.7^2 + 9 \times 3.5^2}{17} = \frac{370.17}{17} = 21.774\dots$	M1
	Test statistic $t = \frac{82.3 - 78.2}{\sqrt{21.774\dots(\frac{1}{9} + \frac{1}{10})}} = 1.9122\dots$ awrt 1.91	M1 A1
	$t_{17}$ (5%) 1-tail critical value = 1.740	B1
	Significant – reject $H_0$ . There is evidence that new style leads to an increase in mean	A1      ✓ (6)
	(b) Assumed population of marks obtained were normally distributed	B1      (1)
	(c) Unbiased estimate of common variance is $s^2$ in (ii)	B1 M1 B1
$7.564 < \frac{17s^2}{\sigma^2} < 30.191$	A1	
$\sigma^2 > \frac{17 \times 21.774\dots}{30.191} = 12.3$ (1 d.p.)	A1	
$\sigma^2 < \frac{17 \times 21.774\dots}{7.564} = 48.9$ (1 d.p.)	A1      (5)	
		<b>(16 marks)</b>

Question number	Scheme	Marks
6. (a)	$X_1 \sim B(10, p) \therefore E(X_1) = 10p \Rightarrow E(R_1) = E\left(\frac{X_1}{10}\right) = \frac{10p}{10} = p$	B1 (1)
(b)	$X_2 \sim B(n, p) \therefore E(X_2) = np \Rightarrow E(R_2) = E\left(\frac{X_2}{n}\right) = \frac{np}{n} = p$	B1
	$E(Y) = E\left(\frac{1}{2}[R_1 + R_2]\right) = \frac{1}{2}[E(R_1) + E(R_2)] = \frac{1}{2}[p + p] = p$	B1 (2)
(c)	$\text{Var}(R_2) = \frac{1}{n^2} \text{Var}(X_2) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$	B1
	$\text{Var}(R_1) = \frac{p(1-p)}{10} \therefore \text{Var}(Y) = \frac{1}{4}[\text{Var}(R_1) + \text{Var}(R_2)],$ $= \frac{1}{4} \left[ \frac{p(1-p)}{10} + \frac{p(1-p)}{n} \right]$	M1 A1 (3)
(d)	<p>Since <math>\text{Var}(R_2) = \frac{p(1-p)}{n} \rightarrow 0</math> as <math>n \rightarrow \infty</math>, <math>\therefore R_2</math> is consistent</p>	M1, A1 (2)
(e)	$\text{Var}(R_1) = \frac{p(1-p)}{10} > \frac{p(1-p)}{20} = \text{Var}(R_2)$	
	$\text{Var}(Y) = \frac{p(1-p)}{4} \left[ \frac{1}{10} + \frac{1}{20} \right] = \frac{p(1-p)}{80} \times 3 < \text{Var}(R_2)$	M1
	<p>Since all 3 are unbiased, we select the one with minimum variance, i.e. <math>Y</math></p>	A1 (2)

Question number	Scheme	Marks
(f)	<p><math>X_1 + X_2 \sim B(n + 10, p)</math> so consider <math>\frac{X_1 + X_2}{n + 10}</math></p> <p><math>E\left(\frac{X_1 + X_2}{n + 10}\right) = \frac{(n + 10)p}{(n + 10)} = p</math> (show unbiased)</p> <p><math>\text{Var}\left(\frac{X_1 + X_2}{n + 10}\right) = \frac{p(1 - p)}{n + 10}</math> (find variance)</p> <p><math>\frac{p(1 - p)}{n + 10} &lt; \frac{p(1 - p)}{10} \therefore</math> always better than <math>R_1</math></p> <p>and both</p> <p><math>\frac{p(1 - p)}{n + 10} &lt; \frac{p(1 - p)}{n} \therefore</math> always better than <math>R_2</math></p> <p><math>\frac{p(1 - p)}{n + 10} &lt; \frac{p(1 - p)}{4} \left[ \frac{n + 10}{10n} \right]</math></p> <p><math>\Leftrightarrow 40n &lt; 100 + 20n + n^2</math></p> <p><math>\Leftrightarrow 0 &lt; 10^2 - 20n + n^2</math></p> <p><math>\Leftrightarrow 0 &lt; (10 - n)^2</math></p> <p style="text-align: right;">Show better than <math>Y</math> Use of <math>n = 20</math> acceptable</p> <p><math>\therefore \frac{X_1 + X_2}{n + 10}</math> is unbiased and always has smaller variance</p>	<p>B1</p> <p>M1</p> <p>M1</p> <p>A1</p> <p>M1</p> <p>A1 cso (6)</p> <p style="text-align: right;"><b>(16 marks)</b></p>