

June 2005
6686 Statistics S4
Mark Scheme

Question Number	Scheme	Marks
1.	<p>(a) $P(X > 19.023) = 0.025$ or $P(X < 19.023) = 0.975$ both $P(X > 2.088) = 0.990$ or $P(X < 2.088) = 0.010$</p> <p>$\therefore P(2.088 < X < 19.023) = 0.990 - 0.025$ or $0.975 - 0.010$ $= 0.965$</p> <p>(b) Upper Critical value of $F_{12,5} = 4.68$ lower Critical value of $F_{12,5} = \frac{1}{F_{5,12}}$ $= \frac{1}{3.11} = 0.3215...$ Answer 0.322</p>	<p>B1 M1 A1 (3) B1 M1 A1 (3)</p>
2.	<p>(a) $H_0: \sigma_1^2 = \sigma_2^2$; $H_1: \sigma_1^2 \neq \sigma_2^2$ both</p> <p>$\frac{S_1^2}{S_2^2} = \frac{14^2}{8^2} = 3.0625$ or $\frac{8^2}{14^2} = 0.32653...$ Answer 3.06 or 0.327</p> <p>C.V $F_{12,7} = 3.57$ cv: $F_{7,12} = \frac{1}{3.57} = 0.28011$</p> <p>Since 3.0625 is not in the Critical Region there is insufficient evidence to reject H_0. There is insufficient evidence of a difference in the variances of the lengths of the fence posts.</p> <p>(b) The distribution of the population of lengths of fence posts is normally distributed.</p>	<p>B1 M1 A1 B1 A1 (5) B1 (1)</p>

Question Number	Scheme	
3.	<p>Let x represent weight of flour</p> <p>$\sum x = 8055 \therefore \bar{x} = \underline{1006.875}$ AWRT 1006.9</p> <p>$\sum x^2 = 8110611 \therefore s^2 = \frac{1}{7} \left\{ 8110611 - \frac{8055^2}{8} \right\} = 33.26785\dots$ AWRT 33.7</p> <p style="border: 1px solid black; border-radius: 50%; padding: 5px; display: inline-block;">Allow from Calculator</p> <p>$\therefore s = 5.767828\dots$ or AWRT 5.77</p> <p>$H_0: \mu = 1010; H_1: \mu < 1010$ both</p> <p>CV: $k = 1.895$ BI</p> <p>$t = \frac{ 1006.875 - 1010 }{5.767828\dots/\sqrt{8}} = \pm 1.5324$ Ung $\frac{\bar{x} - \mu}{s/\sqrt{n}}$</p> <p style="text-align: right;">AWRT -1.53 AI</p> <p>Since -1.53 is not in the critical region ($k < -1.895$) there is insufficient evidence to reject H_0 and thus the mean weight of flour delivered by the machine is 1010g. AI ✓ (A)</p>	<p>BI</p> <p>MI</p> <p>AI</p> <p>BI</p> <p>BI</p> <p>MI</p> <p>AI</p> <p>AI ✓ (A)</p>

<p>4.</p>	<p>(a) The data were not collected in pairs.</p> <p>(b) Use data from twin lambs.</p> <p>(c) Age, weight, gender Any Two sensible factors.</p> <p>(d) $d = B - A$ $d: 2, 1.2, 1, 1.8, -1, 2.2, 2, -1.2, 1.1, 2.8$ $\Sigma d = 11.9; \Sigma d^2 = 30.01$ $\therefore \bar{d} = 1.19; s^2 = 1.761 (s = 1.327)$ $H_0: \delta = 0; H_1: \delta \neq 0$ Allow μ_D for δ both $t = \frac{1.19 - 0}{\sqrt{\frac{1.761}{10}}} = 2.83574 \dots$ $n = 9; CV: t = 2.262$ Since 2.8357... is in the critical region ($t > 2.262$) there is evidence to reject H_0. The (mean) weight gained by the lambs is different for each diet.</p> <p>(e) Diet B - it has the higher mean</p>	<p>B1</p> <p>B1 (1)</p> <p>B1; B1 (2)</p> <p>M1</p> <p>A1; A1</p> <p>B1</p> <p>M1</p> <p>A1</p> <p>B1</p> <p>A1 (8)</p> <p>B1 (1)</p>
	<p>(d) Using non-paired t-test. $H_0: \mu_A = \mu_B; H_1: \mu_A \neq \mu_B$ $t = \frac{\mu_A - \mu_B}{\sqrt{s_p^2 (\frac{1}{10} + \frac{1}{10})}} = -2.30$ AURT -2.30 $CV: t = 2.101$ Conclusion: Mean weight gained is different $\mu_A = 5.6 \mu_B = 6.79 S_p^2 = 1.342722 \dots$</p>	<p>B1</p> <p>B1</p> <p>B1</p> <p>B1 (4)</p>

<p>5.</p>	<p>(a) A Type I error occurs when H_0 is rejected when in fact it is true.</p> <p>(b) The size of a test is the probability of a Type I error</p> <p>(c) $X \sim B(8, 0.25)$ Can be implied</p> <p>Size = $P(X > 6) = 1 - P(X \leq 6 n=8, p=0.25)$</p> $= 1 - 0.9996$ $= \underline{\underline{0.0004}}$ <p>(d) Power = $P(X > 6 p=p, n=8)$</p> $= P(X=7) + P(X=8)$ $= \frac{8!}{7!1!} p^7(1-p) + p^8$ $= 8p^7 - 8p^8 + p^8$ $\underline{\underline{= 8p^7 - 7p^8}}$ <p>(e) Power when $p=0.3 = 8 \times 0.3^7 - 7 \times 0.3^8$</p> $= \underline{\underline{0.00129}}$ Answer 0.0013 <p>(f) $P(\text{Type II error}) = 1 - \text{Power}(0.3)$</p> $= \underline{\underline{0.99870 \dots}}$ <p>(g) <u>I</u>ncrease the number of trials <u>I</u>ncrease the critical region</p>	<p>BI</p> <p>BI (1)</p> <p>BI</p> <p>MI</p> <p>AI (3)</p> <p>MI</p> <p>AI</p> <p>AI (3)</p> <p>BI (1)</p> <p>MI</p> <p>AI (2)</p> <p>BI</p> <p>BI (2)</p>
-----------	---	---

6. (a) Confidence interval is given by

$$\bar{x} \pm t_{19} \times s/\sqrt{n}$$

ie:- $207.1 \pm 2.539 \times \sqrt{\frac{3.2}{20}}$

ie:- 207.1 ± 1.0156

ie:- $(206.084, 208.1156)$

2.539
Using $\bar{x} \pm t \times s/\sqrt{n}$
All correct

AWRT (206, 208)

BI
M1
A1
A1 (4)

(b)
$$s_p^2 = \frac{19 \times 3.2 + 9 \times 10.2173}{28}$$

$$= \underline{5.45557\dots}$$

10.2173
AWRT 10.2
AWRT 5.46

BI
M1
A1

Confidence interval is given by

$$\bar{x}_B - \bar{x}_G \pm t_{28} \times \sqrt{5.45557 \left(\frac{1}{20} + \frac{1}{10} \right)}$$

ie:- $(207.1 - 204.62) \pm 1.701 \sqrt{5.45557 \left(\frac{1}{20} + \frac{1}{10} \right)}$

1.701 BI

ie:- 2.48 ± 1.53875

Using $\bar{x} - \bar{y} \pm t \sqrt{s \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}$
All correct

M1
A1

ie:- $(0.94125, 4.01875)$

AWRT 0.941; 4.02

A1; A1 (8)

<p>7.</p>	<p>(a) $E(X) = np$, $E(Y) = mp$ both; can be implied</p> $E(p_1) = \frac{aE(X)}{n} + \frac{bE(Y)}{m} = p; \Rightarrow \frac{anp}{n} + \frac{bmp}{m} = p$ $\nRightarrow \underline{a+b=1} \quad \nexists$ <p>(b) $E(p_2) = \frac{1}{n+m} \{E(X) + E(Y)\}$</p> $= \frac{1}{n+m} \{np + mp\}$ $= \frac{1}{n+m} \cdot p(n+m) = p \Rightarrow \underline{p_2 \text{ is unbiased}}$ <p>(c) $\text{Var}(X) = np(1-p)$; $\text{Var}(Y) = mp(1-p)$ both; can be implied</p> $\text{Var}(p_1) = \frac{a^2 \text{Var}(X)}{n^2} + \frac{b^2 \text{Var}(Y)}{m^2}$ Use of $\text{Var}(ax) = a^2 \text{Var}(x)$ $= \frac{a^2 np(1-p)}{n^2} + \frac{b^2 mp(1-p)}{m^2}$ $= p(1-p) \left\{ \frac{a^2}{n} + \frac{b^2}{m} \right\}$ <p>(d) $\text{Var}(p_2) = \frac{1}{(n+m)^2} \{np(1-p) + mp(1-p)\}$</p> $= \frac{p(1-p)}{n+m}$ <p>(e) $\text{Var}(p_1) = 0.044p(1-p)$; $\text{Var}(p_2) = 0.0333p(1-p)$</p> <p>Use \leftarrow $\text{Var}(p_2) < \text{Var}(p_1)$</p>	<p>B1 M1 A1 A1 (4)</p> <p>M1 A1 A1 (3)</p> <p>B1 M1 A1 (3)</p> <p>M1 A1 A1 (3)</p> <p>B1; B1 B1; B1 (4) dep</p>
-----------	--	---

[Signature]
05/07/05