Further Hypothesis Tests

Questions

Q1.

The times, x seconds, taken by the competitors in the 100m freestyle events at a school swimming gala are recorded. The following statistics are obtained from the data.

	No. of competitors	Sample mean $\bar{x}$	$\sum x^2$
Girls	8	83.1	55 746
Boys	7	88.9	56 130

Following the gala, a mother claims that girls are faster swimmers than boys. Assuming that the times taken by the competitors are two independent random samples from normal distributions,

(a) test, at the 10% level of significance, whether or not the variances of the two distributions are the same. State your hypotheses clearly.

(7)

(b) Stating your hypotheses clearly, test the mother's claim. Use a 5% level of significance.

(6)

(Total for question = 13 marks)

Q2.

Yin grows two varieties of potato, plant A and plant B. A random sample of each variety of potato is taken and the yield, x kg, produced by each plant is measured. The following statistics are obtained from the data.

	Number of plants	$\sum x$	$\sum x^2$
A	25	194.7	1637.37
B	26	227.5	2031.19

(a) Stating your hypotheses clearly, test, at the 10% significance level, whether or not the variances of the yields of the two varieties of potato are the same.

(7)

(b) State an assumption you have made in order to carry out the test in part (a).

(1)

(Total for question = 8 marks)

Q3.

The concentration of an air pollutant is measured in micrograms/m^3

Samples of air were taken at two different sites and the concentration of this particular air pollutant was recorded.

For Site *A* the summary statistics are shown below.

	number of samples	s_A^2
Site *A*	13	6.39

For Site *B* there were 9 samples of air taken.

A test of the hypothesis $H_0: \sigma_A^2 = \sigma_B^2$ against the hypothesis $H_1: \sigma_A^2 \neq \sigma_B^2$ is carried out using a 2% level of significance.

(a) State a necessary assumption required to carry out the test.

(1)

Given that the assumption in part (a) holds,

(b) find the set of values of s_B^2 that would lead to the null hypothesis being rejected,

(4)

(c) find a 99% confidence interval for the variance of the concentration of the air pollutant at **Site *A***.

(3)

(Total for question = 8 marks)

Q4.

A nutritionist studied the levels of cholesterol, *X* mg/cm^3, of male students at a large college. She assumed that *X* was distributed N(μ, σ^2) and examined a random sample of 25 male students. Using this sample she obtained unbiased estimates of μ and σ^2 as $\hat{\mu}$ and $\hat{\sigma}^2$

A 95% confidence interval for μ was found to be (1.128, 2.232)

(a) Show that $\hat{\sigma}^2$ = 1.79 (correct to 3 significant figures)

(4)

(b) Obtain a 95% confidence interval for σ^2

(3)

(Total for question = 7 marks)

Mark Scheme – Further Hypothesis Tests

Q1.

Question	Scheme	Marks	AOs
(a)	H_0: $\sigma_G^2 = \sigma_B^2$, H_1: $\sigma_G^2 \neq \sigma_B^2$,	B1	2.5
	$s_B^2 = \frac{1}{6}(56130 - 7 \times 88.9^2) = \frac{807.53}{6} = 134.6$	M1 A1	2.1 1.1b
	$s_G^2 = \frac{1}{7}(55746 - 8 \times 83.1^2) = \frac{501.12}{7} = 71.58$	A1	1.1b
	$\frac{s_B^2}{s_G^2} = 1.880\ldots$	M1	3.4
	critical value $F_{6,7} = 3.87$	B1	1.1b
	not significant, variances can be treated as the same	A1 ft	2.2b
		(7)	
(b)	H_0: $\mu_B = \mu_G$, H_1: $\mu_B > \mu_G$	B1	2.5
	pooled estimate of variance $s^2 = \frac{6 \times 134.6 + 7 \times 71.58}{13} = 100.6653\ldots$	M1	3.1b
	test statistic $t = \frac{88.9 - 83.1}{s\sqrt{\frac{1}{7} + \frac{1}{8}}}$ = awrt 1.12	M1 A1	1.1b 1.1b
	critical value $t_{13}(5\%) = 1.771$	B1	1.1b
	Insufficient evidence to support mother's claim	A1 ft	2.2b
		(6)	
		(13 marks)	

Notes	
(a)	**B1:** Both hypotheses correct using the notation σ^2. Allow σ rather than σ^2. **M1:** Using a correct Method for either s_B^2 or s_G^2 May be implied by a correct value **A1:** awrt 135 **A1:** awrt 71.6 **M1:** Using the F-distribution as the model eg $\frac{s_B^2}{s_G^2}$ **B1:** awrt 3.87 **A1ft:** Drawing a correct inference following through their CV and value for $\frac{s_B^2}{s_G^2}$
(b)	**B1:** Both hypotheses correct using the notation μ. **M1:** For realising the need to find the pooled estimate for the test require from a correct interpretation of the question. **M1:** Correct method for test statistic $t = \frac{88.9 - 83.1}{\text{"their } s\text{"}\sqrt{\frac{1}{7} + \frac{1}{8}}}$ May be implied by a correct awrt 1.12 **A1:** awrt 1.12 **B1:** awrt 1.77 **A1ft** Drawing a correct inference following through their CV and test statistic

Q2.

Question	Scheme	Marks	AOs
(a)	$H_0 : \sigma_A^2 = \sigma_B^2$, $H_1 : \sigma_A^2 \neq \sigma_B^2$	B1	2.5
	$s_A^2 = \frac{1}{24}\left(1637.37 - 25\times\left(\frac{194.7}{25}\right)^2\right) = 5.0436$	M1 A1	2.1 1.1b
	$s_B^2 = \frac{1}{25}\left(2031.19 - 26\times\left(\frac{227.5}{26}\right)^2\right) = 1.6226$	A1	1.1b
	$\frac{s_A^2}{s_B^2} = 3.108...$	M1	3.4
	critical values upper tail $F_{24,25} = 1.96$	B1	1.1b
	There is evidence that the two **variances** are different.	A1ft	2.2b
		(7)	
(b)	The **yields** are normally distributed.	B1	1.2
		(1)	
		(8 marks)	

Notes:		
(a)	**B1**	both hypotheses correct using σ **or** σ^2
	M1	Using a correct method for either s_A^2 or s_B^2. May be implied by a correct value
	A1	awrt 5.04
	A1	awrt 1.62
	M1	Using the F-distribution as the model eg $\frac{s_A^2}{s_B^2}$ $\left(\text{allow } \frac{s_B^2}{s_A^2}[= 0.321...]\right)$
	B1	awrt 1.96 or 0.506 must match their method
	A1ft	Drawing a correct inference following through their CV and value for $\frac{s_B^2}{s_A^2}$ or $\frac{s_A^2}{s_B^2}$ Allow $\sigma_B^2 \neq \sigma_A^2$ Allow standard deviation instead of Var . Do not allow $\sigma_B^2 = \sigma_A^2$
(b)	**B1**	recalling the fact that the variable **yield** needs to be normally distributed

Q3.

Question	Scheme	Marks	AOs
(a)	The concentration of air pollutant (for each site) follows a normal distribution	B1	2.4
		(1)	
(b)	$F_{12,8\,(0.01)} = 5.67$ or $F_{8,12\,(0.01)} = 4.5(0)$	B1	3.3
	$\frac{6.39}{s_B^2} > '5.67'$ $\frac{s_B^2}{6.39} > '4.50'$	M1M1	1.1b 2.1
	$(0<)s_B^2 < 1.12698...$, $s_B^2 > 28.755$	A1	1.1b
		(4)	
(c)	$\chi^2_{12,0.005} = 28.3$ or $\chi^2_{12,0.995} = 3.074$	M1	3.3
	$\frac{12 \times 6.39}{"28.3"} < \sigma_A^2 < \frac{12 \times 6.39}{"3.074"}$	M1	1.1b
	$2.7095... < \sigma_A^2 < 24.94469...$	A1	1.1b
		(3)	
			(8 marks)

	Notes
(a)	**B1:** Correct modelling assumption. Allow air pollutant samples are independent. Samples of air pollutant are normally distributed is B0
(b)	**B1:** Setting up either *F* statistic **M1:** Either correct inequality (condone equation) **M1:** Understanding both correct inequalities are needed (condone equation) **A1:** Correct range of values with awrt 1.13 and awrt 28.8 $s_B^2 < 1.12698...$ and $s_B^2 > 28.755$ is A0
(c)	**M1:** For realising a χ^2 distribution must be used as a model with at least one value correct . Allow $\chi^2_{12,0.01} = 26.217$ or $\chi^2_{12,0.99} = 3.571$ for this mark. **M1:** Setting up 99% CI **A1:** Correct CI with awrt 2.71 and awrt 24.9

Q4.

Question	Scheme	Marks	AOs
(a)	95% CI for μ uses t value of **2.064**	B1	3.3
	$\frac{\hat{\sigma}}{\sqrt{25}} \times "2.064" = \frac{1}{2}(2.232 - 1.128)$ <u>or</u> $\frac{1}{2}(2.232 + 1.128) + "2.064" \times \frac{\hat{\sigma}}{\sqrt{25}} = 2.232$ (oe)	M1	2.1
	$\hat{\sigma} = \frac{2.76}{"2.064"}$ <u>or</u> 1.3372…	M1	1.1b
	$\hat{\sigma}^2 = 1.788...$[=1.79 (3sf)] *	A1*cso	1.1b
		(4)	
(b)	$12.401, < \frac{24 \times 1.79}{\sigma^2} <, 39.364$	B1 M1	1.1b 1.1a
	<u>**1.09**</u> $< \sigma^2 <$ <u>**3.46**</u>	A1	1.1b
		(3)	
			(7 marks)

Notes	
(a)	**B1**: Realising that the t-distribution must be used as a model and finding the correct value awrt 2.06 **M1**: Using the correct formula with a t-value, $\frac{\hat{\sigma}}{\sqrt{25}} \times "t \text{ value}" = \frac{1}{2}(2.232 - 1.128)$ or $\frac{1}{2}(2.232 + 1.128) + "t \text{ value}" \times \frac{\hat{\sigma}}{\sqrt{25}} = 2.232$ or $\frac{1}{2}(2.232 + 1.128) - "t \text{ value}" \times \frac{\hat{\sigma}}{\sqrt{25}} = 1.128$ **M1**: Rearranging one of these formula accurately to find a value of $\hat{\sigma}$ **A1cso***: A correct solution only using awrt 1.79
(b)	**B1**: awrt 12.4 or 39.4 May be implied by a correct confidence interval **M1**: $\frac{24 \times 1.79}{\sigma^2}$ May be implied by a correct confidence interval **A1**: awrt 1.09 and awrt 3.46