Surname	StemUp Model Solution	Centre Number	Candidate Number
First name(s)			2

GCE A LEVEL

S24-1305U50-1

1305U50-1

WEDNESDAY, 12 JUNE 2024 – AFTERNOON

FURTHER MATHEMATICS – A2 unit 5
FURTHER STATISTICS B

1 hour 45 minutes

For Examiner's use only		
Question	**Maximum Mark**	**Mark Awarded**
1	9	
2	9	
3	7	
4	11	
5	19	
6	6	
7	19	
Total	80	

1305U501
01

ADDITIONAL MATERIALS

In addition to this examination paper, you will need:

- a Formula Booklet;
- a calculator.
- statistical tables (RND/WJEC Publications).

INSTRUCTIONS TO CANDIDATES

Use black ink or black ball-point pen. Do not use gel pen or correction fluid.

You may use a pencil for graphs and diagrams only.

Write your name, centre number and candidate number in the spaces at the top of this page.

Answer **all** questions.

Write your answers in the spaces provided in this booklet. If you run out of space, use the additional page(s) at the back of the booklet, taking care to number the question(s) correctly.

INFORMATION FOR CANDIDATES

The maximum mark for this paper is 80.

The number of marks is given in brackets at the end of each question or part-question.

Sufficient working must be shown to demonstrate the **mathematical** method employed.

Answers without working may not gain full credit.

Unless the degree of accuracy is stated in the question, answers should be rounded appropriately.

You are reminded of the necessity for good English and orderly presentation in your answers.

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

Reminder: Sufficient working must be shown to demonstrate the **mathematical** method employed.

1. During practice sessions, a basketball coach makes his players run several 'line drills'.

(a) He records the times taken, in seconds, by one of his players to run the first 'line drill' on a random sample of 8 practice sessions. The results are shown below.

29·4 31·1 28·9 30·0 29·9 30·4 29·7 30·2

Assuming that these data come from a normal distribution with mean μ and variance 0·6, calculate a 95% confidence interval for μ. [5]

$\Sigma x = 239.6, \quad \bar{x} = \frac{239.6}{8} = 29.95$

$SE = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{0.6}{8}} = 0.27386...$

95% CI: $\bar{x} \pm z \times SE$ with $z = 1.96$

$= 29.95 \pm 1.96 \times \sqrt{\frac{0.6}{8}}$

$= 29.95 \pm 0.53677...$

$= 29.413...$ to $30.487...$

95% confidence interval for μ is $[29.41, 30.49]$

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

(b) State the two ways in which the method used to calculate the confidence interval in part (a) would change if the variance were unknown. [2]

1. Use the t-distribution with 7 degrees of freedom instead of the normal distribution, i.e. use $t = 2.365$ instead of $z = 1.96$.

2. Use the sample variance: replace $\sigma^2 = 0.6$ by the unbiased estimate s^2 of the variance calculated from the 8 times.

(c) During a practice session, a player recorded a mean time of 35·6 seconds for 'line drills'.

(i) Give a reason why this player may not be the same as the player in part (a).

(ii) Give a reason why this player could be the same as the player in part (a). [2]

(i) 35.6 s is far above the confidence interval [29.41, 30.49] and much higher than every time recorded by the player in part (a), so it may have been a different player.

(ii) In part (a) the times are for the first line drill only, but 35.6 s is the mean of all the line drills in a session. The player will get tired over several drills, so his later drills are slower and his mean time could be much higher.

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

2. In country A, the median daily caffeine intake per student who drinks coffee is 120 mg. A university professor who oversees a foreign exchange programme believes that students visiting from country B drink more coffee and therefore have a greater daily caffeine intake from coffee.

On a randomly chosen day, the caffeine intake, in mg, from coffee consumption by each of 15 randomly selected students from country B is given below.

136	149	202	0	110	0	100	180
0	187	0	0	138	197	115	

The professor suspects that the students with zero caffeine intake do not drink coffee, and decides to ignore those students and instead focus on the coffee-drinking students.

(a) Conduct an appropriate Wilcoxon test at a significance level as close to 5% as possible. State your conclusion in context. [8]

H_0: $\eta = 120$ H_1: $\eta > 120$, where η is the median daily caffeine intake (mg) from coffee of coffee-drinking students from country B.

Ignoring the 5 zeros, $n = 10$. Differences $d = x - 120$:

x	136	149	202	110	100	180	187	138	197	115
$x - 120$	16	29	82	−10	−20	60	67	18	77	−5
rank of $\lvert d \rvert$	3	6	10	2	5	7	8	4	9	1

$W_+ = 3 + 6 + 10 + 7 + 8 + 4 + 9 = 47$

$W_- = 2 + 5 + 1 = 8$ (check: $47 + 8 = 55 = \frac{1}{2} \times 10 \times 11$)

$n = 10$, one-tailed, as close to 5% as possible: critical value = 11

Critical region: $W_- \leq 11$ (equivalently $W_+ \geq 55 - 11 = 44$)

Since $8 < 11$ ($47 > 44$), the result is significant, so reject H_0.

There is sufficient evidence to suggest that the coffee-drinking students from country B have a median daily caffeine intake from coffee greater than 120 mg, i.e. they drink more coffee than the students from country A.

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

(b) State one limitation of this investigation. [1]

The 5 students with zero intake (a third of the sample) were ignored, so the conclusion only applies to coffee drinkers, not to all students from B.

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

3. Tony runs a pie stand that sells two types of pie outside a football ground. He wants to estimate the probability that a customer will buy a steak pie rather than a vegetable pie. He conducts a survey by randomly selecting customers and recording their choice of pie. When he feels he has enough data, he notes that 55 customers bought steak pies and 25 bought vegetable pies.

(a) Calculate an approximate 90% confidence interval for p, the probability that a randomly selected customer buys a steak pie. [6]

$\hat{p} = \frac{55}{80} = 0.6875$

$ESE = \sqrt{\hat{p}\,(1-\hat{p}) \div n} = \sqrt{\frac{0.6875 \times 0.3125}{80}}$

$= 0.051822...$

90% confidence: $z = 1.6449$

CI: $\hat{p} \pm z \times ESE = 0.6875 \pm 1.6449 \times 0.051822...$

$= 0.6875 \pm 0.085240...$

$= 0.60226...$ to $0.77274...$

Approximate 90% confidence interval for p is $[0.602, 0.773]$

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

(b) Suppose that Tony carries out 50 such surveys and calculates 90% confidence intervals for each survey. Determine the expected number of these confidence intervals that would contain the true value of p. [1]

Expected number = 0.9 × 50 = 45

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

4. The sports performance director at a university wishes to investigate whether there is a difference in the means of the specific gravities of blood of cyclists and runners. She models the distribution of specific gravity for cyclists as $N(\mu_X, 8^2)$ and for runners as $N(\mu_Y, 10^2)$.

(a) State suitable hypotheses for this investigation. [1]

H_0: $\mu_X = \mu_Y$ H_1: $\mu_X \neq \mu_Y$

The mean specific gravity of blood of a random sample of 40 cyclists from the university was 1063. The mean specific gravity of blood of a random sample of 40 runners from the same university was 1060.

(b) Calculate and interpret the *p*-value for the data. [6]

Under H_0: $\overline{X} - \overline{Y} \sim N(0, \sigma^2)$, observed difference = 1063 − 1060 = 3

$SE = \sqrt{\frac{8^2}{40} + \frac{10^2}{40}} = \sqrt{4.1} = 2.02484...$

$z = \frac{3 - 0}{2.02484...} = 1.4816$

$P(Z > 1.4816) = 0.0692$

Two-tailed: p-value = 2 × 0.0692 = 0.1384

Since 0.1384 > 0.05, the result is not significant at the 5% level,

so do not reject H_0.

There is insufficient evidence to suggest that there is a difference

between the mean specific gravity of blood of cyclists and of runners.

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

(c) Suppose now that both samples were of size n, instead of 40. Find the least value of n that would ensure that an observed difference of 3 in the mean specific gravities would be significant at the 1% level. [4]

SE of difference in means $= \sqrt{\frac{8^2}{n} + \frac{10^2}{n}} = \sqrt{\frac{164}{n}}$

Need $P(Z > \frac{3 - 0}{\sqrt{(164/n)}}) < 0.005$ (1% level, two-tailed)

$\frac{3}{\sqrt{(164/n)}} > 2.5758$

$\sqrt{n} > \frac{2.5758 \times \sqrt{164}}{3} = 10.9956...$

$n > 120.90...$

Least value of $n = 121$

1305U501 09

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

5. The probability density function of the continuous random variable X is given by

$$f(x) = \frac{3x^2}{\alpha^3} \qquad \text{for } 0 \leqslant x \leqslant \alpha$$

$$f(x) = 0 \qquad \text{otherwise.}$$

$\bar{X}$ is the mean of a random sample of n observations of X.

(a) (i) Show that $U = \frac{4\bar{X}}{3}$ is an unbiased estimator for α. [5]

(ii) If α is an integer, what is the smallest value of n that gives a rational value for the standard error of U? [9]

(i) $E(X) = \int_0^\alpha \frac{x \times 3x^2}{\alpha^3}\,dx = \int_0^\alpha \frac{3x^3}{\alpha^3}\,dx$

$= \left[\frac{3x^4}{4\alpha^3}\right]_0^\alpha$

$= \frac{3\alpha}{4}$

$E(\bar{X}) = E(X) = \frac{3\alpha}{4}$

$E(U) = \frac{4}{3}E(\bar{X}) = \frac{4}{3} \times \frac{3\alpha}{4} = \alpha$

Since $E(U) = \alpha$, U is an unbiased estimator for α.

(ii) $E(X^2) = \int_0^\alpha \frac{3x^4}{\alpha^3}\,dx = \left[\frac{3x^5}{5\alpha^3}\right]_0^\alpha = \frac{3\alpha^2}{5}$

$Var(X) = \frac{3\alpha^2}{5} - \left(\frac{3\alpha}{4}\right)^2 = \frac{3\alpha^2}{5} - \frac{9\alpha^2}{16} = \frac{3\alpha^2}{80}$

$Var(\bar{X}) = \frac{Var(X)}{n} = \frac{3\alpha^2}{80n}$

$Var(U) = \left(\frac{4}{3}\right)^2 Var(\bar{X}) = \frac{16}{9} \times \frac{3\alpha^2}{80n} = \frac{\alpha^2}{15n}$

$SE(U) = \sqrt{\frac{\alpha^2}{15n}} = \frac{\alpha}{\sqrt{(15n)}}$

α is an integer, so the SE is rational when $15n$ is a perfect square.

Smallest: $15n = 15^2 = 225 \Rightarrow n = 15$, giving $SE = \frac{\alpha}{15}$

Smallest value of n is 15.

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

(b) $\bar{X}_1$ and $\bar{X}_2$ are the means of independent random samples of X, each of size n.

The estimator $V = 4\bar{X}_1 - \frac{8}{3}\bar{X}_2$ is also an unbiased estimator for α.

(i) Show that $\frac{\text{Var}(U)}{\text{Var}(V)} = \frac{1}{13}$. [4]

(ii) Hence state, with a reason, which of U or V is the better estimator. [1]

(i) $\text{Var}(\bar{X}_1) = \text{Var}(\bar{X}_2) = \frac{3\alpha^2}{80n}$ (from (a)(ii))

$\text{Var}(V) = 4^2\,\text{Var}(\bar{X}_1) + (\frac{8}{3})^2\,\text{Var}(\bar{X}_2)$ (independent)

$= 16 \times \frac{3\alpha^2}{80n} + \frac{64}{9} \times \frac{3\alpha^2}{80n}$

$= \frac{3\alpha^2}{5n} + \frac{4\alpha^2}{15n} = \frac{9\alpha^2 + 4\alpha^2}{15n} = \frac{13\alpha^2}{15n}$

$\frac{\text{Var}(U)}{\text{Var}(V)} = \frac{\alpha^2}{15n} \times \frac{15n}{13\alpha^2} = \frac{1}{13}$ as required

(ii) U is the better estimator: both are unbiased, and since

$\text{Var}(U) \div \text{Var}(V) = \frac{1}{13} < 1$, U has the smaller variance.

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

6. Alana is a PhD student researching language acquisition. She gives one group of randomly selected participants, Group A, 4 minutes to memorise 40 words that are similar in meaning. She gives a different, randomly selected group of participants, Group B, 4 minutes to memorise 40 words that are different in meaning. Alana believes that the students in Group B will do better than the students in Group A. The following results are the number of words recalled on testing the students from the two groups.

Group A	32	8	24	16	10	20	22	18	23	21	26	14
Group B	30	29	11	25	38	36	28	12	17			

Conduct a Mann-Whitney U test at a significance level as close as possible to 5% to test Alana's belief. [6]

H_0: the median number of words recalled is the same for both groups

H_1: the median for Group B is greater than the median for Group A

Ranking all 21 scores together:

Group A	8	10	14	16	18	20	21	22	23	24	26	32
Rank	1	2	5	6	8	9	10	11	12	13	15	19
Group B	11	12	17	25	28	29	30	36	38			
Rank	3	4	7	14	16	17	18	20	21			

Sum of ranks: $R_A = 111$, $R_B = 120$ (total $231 = \frac{1}{2} \times 21 \times 22$)

$U_A = 111 - \frac{1}{2} \times 12 \times 13 = 111 - 78 = 33$

$U_B = 120 - \frac{1}{2} \times 9 \times 10 = 120 - 45 = 75$ $\quad (33 + 75 = 108 = 12 \times 9)$

$m = 9$, $n = 12$, one-tailed, as close to 5% as possible: critical value = 30

Critical region: $U \leq 30$ (equivalently $U \geq 108 - 30 = 78$)

Since $33 > 30$ ($75 < 78$), the result is not significant: do not reject H_0.

There is insufficient evidence to support Alana's belief that students in Group B (words different in meaning) recall more words than Group A.

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

7. A farmer uses many identical containers to store four different types of grain: wheat, corn, einkorn and emmer.

(a) The mass W, in kg, of wheat stored in each individual container is normally distributed with mean μ and standard deviation 0·6. Given that, for containers of wheat, 10% store less than 19 kg, find the value of μ. [3]

$P(W < 19) = 0.1$

$P(Z < \frac{19 - \mu}{0.6}) = 0.1$

$\frac{19 - \mu}{0.6} = -1.2816$

$\mu = 19 + 0.6 \times 1.2816 = 19.769...$

$\mu = 19.8$ (3 s.f.)

The mass X, in kg, of corn stored in each individual container is normally distributed with mean 20·1 and standard deviation 1·2.

(b) Find the probability that the mean mass of corn in a random sample of 8 containers of corn will be greater than 20 kg. [3]

$\overline{X} \sim N(20.1, \frac{1.2^2}{8}) = N(20.1, 0.18)$

$P(\overline{X} > 20) = P(Z > \frac{20 - 20.1}{\sqrt{0.18}})$

$= P(Z > -0.2357)$

$= 0.59317...$

$= 0.593$ (3 s.f.)

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

The mass Y, in kg, of einkorn stored in each individual container is normally distributed with mean 22·2 and standard deviation 1·5.

The farmer and his wife need to move two identical wheelbarrows, one of which is loaded with 3 containers of corn, and the other of which is loaded with 3 containers of einkorn. They agree that the farmer's wife will move the heavier wheelbarrow.

(c) Calculate the probability that the farmer's wife will move

(i) the einkorn,

(ii) the corn. [5]

Let $T = (Y_1 + Y_2 + Y_3) - (X_1 + X_2 + X_3)$

$E(T) = 3 \times 22.2 - 3 \times 20.1 = 66.6 - 60.3 = 6.3$

$Var(T) = 3 \times 1.5^2 + 3 \times 1.2^2 = 6.75 + 4.32 = 11.07$

$T \sim N(6.3, 11.07)$

(i) $P(\text{einkorn heavier}) = P(T > 0) = P\left(Z > \frac{0 - 6.3}{\sqrt{11.07}}\right)$

$= P(Z > -1.8935)$

$P(\text{wife moves the einkorn}) = 0.97085$

(ii) $P(\text{wife moves the corn}) = P(T < 0) = 1 - 0.97085 = 0.02915$

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

(d) The mass E, in kg, of emmer stored in each individual container is normally distributed with mean 10·5 and standard deviation σ. The farmer's son tries to calculate the probability that the mass of corn in a single container will be more than three times the mass of emmer in a single container. He obtains an answer of 0·35208.

(i) Find the value of σ that the farmer's son used.

(ii) Explain why the value of σ that he used is unreasonable. [8]

(i) Let $U = X - 3E$. $P(X > 3E) = P(U > 0) = 0.35208$

$E(U) = 20.1 - 3 \times 10.5 = -11.4$

$Var(U) = Var(X) + 9Var(E) = 1.2^2 + 9\sigma^2$

$P(Z > \frac{0 - (-11.4)}{\sigma_U}) = 0.35208$

$\frac{11.4}{\sigma_U} = 0.37971$

$\sigma_U = 30.0228...$

$\sigma_U^2 = 901.37... = 1.44 + 9\sigma^2$

$\sigma^2 = \frac{901.37... - 1.44}{9} = 99.99...$

$\sigma = 10$

(ii) $\sigma = 10$ is far too big: it is almost as large as the mean of 10.5.

It gives $P(E < 0) = P(Z < -1.05) = 0.147$, so about 15% of containers would hold a negative mass of emmer, which is impossible.

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

END OF PAPER

© WJEC CBAC Ltd.

stemup.co.uk

Question number	Additional page, if required. Write the question number(s) in the left-hand margin.

Examiner only

© WJEC CBAC Ltd.

stemup.co.uk

Question number	Additional page, if required. Write the question number(s) in the left-hand margin.

Examiner only

© WJEC CBAC Ltd.

stemup.co.uk

BLANK PAGE

PLEASE DO NOT WRITE ON THIS PAGE

© WJEC CBAC Ltd.

stemup.co.uk