AH

National
Qualifications
SPECIMEN ONLY

S803/77/12

**Statistics
Paper 2**

Date — Not applicable

Duration — 2 hours 45 minutes

Total marks — 90

Attempt ALL questions.

You may use a calculator.

To earn full marks you must show your working in your answers.

State the units for your answer where appropriate.

Write your answers clearly in the answer booklet provided. In the answer booklet you must clearly identify the question number you are attempting.

Use **blue** or **black** ink.

Before leaving the examination room you must give your answer booklet to the Invigilator; if you do not, you may lose all the marks for this paper.

You may refer to the Statistics Advanced Higher Statistical Formulae and Tables.

S8037712

Total marks — 90

Attempt ALL questions

1. The waiting times, to the nearest minute, for a sample of patients attending a hospital clinic were recorded as follows.

   2   2   3   6   7   7   9   10   13   14   29

   Calculate the upper fence for this data and comment on the waiting time of 29 minutes. **3**

2. With the aid of an example of each, define both discrete and continuous data. **4**

3. For three mutually exclusive events A, B and C
   $P(A) = 0\cdot6$ $P(B) = 0\cdot3$ $P(C) = 0\cdot1$.

   For another event S
   $P(S|A) = 0\cdot55$ $P(S|B) = 0\cdot25$ $P(S|C) = 0\cdot05$.

   Calculate

   (a) $P(S)$ **2**

   (b) $P(C|\bar{S})$. **3**

4. At the Vienna General Hospital in 1843 there were 5799 births in the two maternity clinics. Maternal mortality due to puerperal fever was of major concern and classification of the births yielded the following contingency table.

		Maternity clinic	
		No.1	No.2
Maternal outcome	Death from puerperal fever	274	164
	No death from puerperal fever	2786	2575

   (a) Carry out a formal test of association between maternal outcome and the clinic in which the birth took place. **6**

   (b) Interpret your conclusion in the context of the study. **1**

5. Widgets produced in work cell A are known to have tensile strengths with mean 74·3 N and standard deviation 3·2 N when the manufacturing process is operating in a stable manner. In an attempt to increase tensile strength, modifications were made to the process used in the cell. In order to monitor the modified process, random samples of four widgets were taken from the widgets produced each hour and their tensile strengths determined.

(a) Calculate three sigma limits for a control chart of the mean tensile strengths of the samples using the given historical parameters. **2**

The first ten consecutive hourly samples yielded the following means.

76·5　75·2　74·0　74·8　73·7　77·4　74·8　78·7　77·7　77·0

These are plotted on the control chart below with both one and two sigma limits shown in addition to the centre line and three sigma limits.

Control chart of mean tensile strength
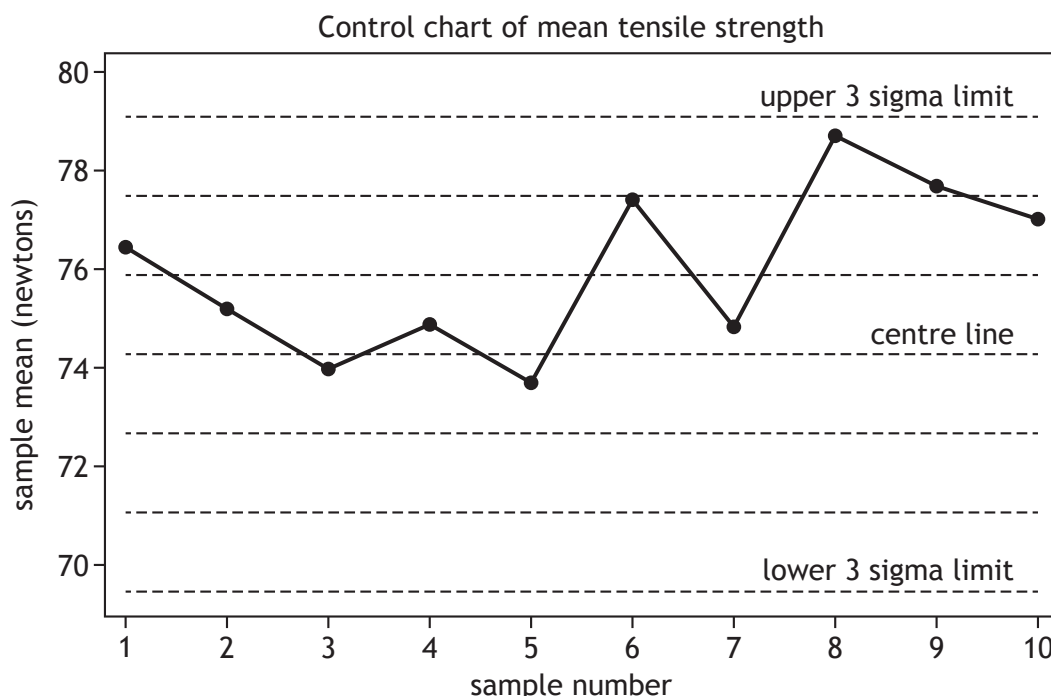
(b) Identify the feature of this control chart which indicates that the process is out of statistical control. State, giving a reason, whether or not you think the team on work cell A would be pleased with this lack of statistical control. **2**

(c) As a result of this analysis, suggest what work cell A could do to continue to monitor the tensile strength of the widgets produced. **2**

[Turn over

6. An unbiased regular octahedral (eight-faced) die is thrown.

Let the random variable X be the score shown on this die.

(a) Calculate E(X) and V(X). **2**

An unbiased cubical die is also thrown.

Let the random variable Y be the score shown on this die.

(b) Calculate E($3X - 2Y$) and V($3X - 2Y$). **4**

7. In order to meet current acoustic standards in the UK for separating floors between flats, the impact sound transmission (IST) should not exceed 62 dB. The IST results (dB) for a certain type of flooring can be adequately modelled by the N($53 \cdot 4$, $4 \cdot 7^2$) distribution.

(a) Calculate the proportion of such floors that would fail to meet the 62 dB standard. **3**

(b) Stating any assumption required, calculate the value to which the mean IST for this type of floor would have to be reduced in order to ensure that at most 1% of floors would fail to meet the 62 dB standard. **4**

8. A commuter travels to work by car using one of two routes, A and B. Their colleague claims that route A is generally faster than route B.

For three weeks the commuter selected the route to take by tossing a coin, using route A if the outcome was heads and route B otherwise.

Journey times (in minutes) were

Route A   46   37   53   50   42   43   49   37
Route B   54   63   44   49   47   55   47

Use a non-parametric test at the 5% level to investigate whether or not the colleague's claim is supported by the data. **6**

9. A competition is run at a local bridge club where, each week, players compete in pairs against other pairs. Scores are given as a percentage for each pair with the highest percentage being the winning score.

A random sample of nine scores obtained by a particular pair, during the first few months of this year's competition, yielded the following percentages.

$$53 \cdot 7 \quad 58 \cdot 9 \quad 56 \cdot 4 \quad 57 \cdot 1 \quad 62 \cdot 0 \quad 55 \cdot 3 \quad 59 \cdot 5 \quad 61 \cdot 1 \quad 57 \cdot 6$$

(a) Stating one assumption that needs to be made, calculate a 95% confidence interval for the population mean percentage score for this pair. **5**

(b) The mean percentage score for this pair in last year's competition was $55 \cdot 4$.

State, with a reason, whether or not you would agree with the opinion that their current scores yield evidence of an improvement on last year. **2**

10. It is known from experience that a machine used to fill soft drinks bottles of various sizes delivers 'fill volumes' that are normally distributed with standard deviation 0·1 ml.

In order to meet statutory requirements for fill volumes, it is desirable that the mean fill volume for 500 ml size bottles should be 500·30 ml. Shortly after the start of a production run of 500 ml bottles, a random sample of 10 bottles was taken and the fill volumes were

500·40   500·22   500·31   500·13   500·16   500·18   500·32   500·31   500·18   500·44

(a)   Use a statistical test to assess whether there is any evidence that the mean fill volume differs from 500·30 ml.   **5**

(b)   Under what circumstance would a t-test of the same hypotheses be required and how might this affect the conclusion?   **2**

11. The following statistics were available to a researcher.

$$\sum x = 2208 \quad \sum y = 1054 \cdot 8 \quad n = 18 \quad S_{xx} = 105904 \quad S_{xy} = 12288 \cdot 5 \quad S_{yy} = 2492 \cdot 2$$

Least squares regression line of y on x is $y = 44 \cdot 37 + 0 \cdot 1160x$.

The researcher wished to estimate some missing values of x from their corresponding values of y and decided to change the subject of the least squares regression line to x yielding $x = 8 \cdot 621y - 382 \cdot 5$.

(a)   The researcher's estimated value of x was 48·55. Explain why changing the subject, as above, is not an appropriate procedure for x on y regression.   **2**

(b)   Obtain the correct least squares regression line of x on y and estimate the value of x for $y = 50$.   **3**

12. In an investigation of whether taste preference changes with age, it was found that 56% of a random sample of 300 people under 40 years of age preferred the new brand of a low fat spread to the old one, whereas 48% of a random sample of 200 people over 40 showed the same preference.

Use a parametric hypothesis test to establish whether there is evidence of a difference in preference between the age groups.   **7**

[Turn over

13. A country hosting a major international soccer tournament produced special commemorative coins to be used by match referees and sold as souvenirs. Concern was expressed prior to the tournament that the coins were biased, yielding heads from 55% of tosses.

It was decided to select one of the coins at random and toss it 200 times. It was agreed to assess the coin biased, as suspected, if 110 or more heads were obtained and to assess the coin fair otherwise.

(a) If the coin is actually fair, use an appropriate approximation to estimate the probability that it is assessed biased in the way suspected. **5**

(b) Similarly, estimate the probability that a biased coin is assessed to be fair. **4**

(c) State a step that could be taken to reduce the risk of incorrect assessment of a coin as biased. **1**

14. The random variables F and G are the ages of the trees in two independent sustainable forests.

F has mean 10 and standard deviation 2.
G has mean 8 and standard deviation 1.

A random sample of size 100 is drawn from each forest and the sample means denoted by $\overline{F}$ and $\overline{G}$.

(a) Calculate the mean and standard deviation of $\overline{F} - \overline{G}$.

Hence find the approximate probability that the absolute difference between the two sample means is less than 1·5. **7**

(b) Explain why the probability above is only approximate and justify the use of this approximation in the context of this question. **3**

[END OF SPECIMEN QUESTION PAPER]

S803/77/12

**Statistics
Paper 2**

Marking Instructions

These marking instructions have been provided to show how SQA would mark this specimen question paper.

General marking principles for Advanced Higher Statistics

Always apply these general principles. Use them in conjunction with the detailed marking instructions, which identify the key features required in candidates' responses.

The marking instructions for each question are generally in two sections:

- *generic scheme — this indicates why each mark is awarded*
- *illustrative scheme — this covers methods which are commonly seen throughout the marking*

In general, you should use the illustrative scheme. Only use the generic scheme where a candidate has used a method not covered in the illustrative scheme.

(a) Always use positive marking. This means candidates accumulate marks for the demonstration of relevant skills, knowledge and understanding; marks are not deducted for errors or omissions.

(b) If you are uncertain how to assess a specific candidate response because it is not covered by the general marking principles or the detailed marking instructions, you must seek guidance from your team leader.

(c) One mark is available for each •. There are no half marks.

(d) If a candidate's response contains an error, all working subsequent to this error must still be marked. Only award marks if the level of difficulty in their working is similar to the level of difficulty in the illustrative scheme.

(e) Only award full marks where the solution contains appropriate working. A correct answer with no working receives no mark, unless specifically mentioned in the marking instructions.

(f) Candidates may use any mathematically correct method to answer questions, except in cases where a particular method is specified or excluded.

(g) If an error is trivial, casual or insignificant, for example $6 \times 6 = 12$, candidates lose the opportunity to gain a mark, except for instances such as the second example in point (h) below.

(h) If a candidate makes a transcription error (question paper to script or within script), they lose the opportunity to gain the next process mark, for example

This is a transcription error and so the mark is not awarded.

$$x^2 + 5x + 7 = 9x + 4$$
$$x - 4x + 3 = 0$$

This is no longer a solution of a quadratic equation, so the mark is not awarded.

$$x = 1$$

The following example is an exception to the above

This error is not treated as a transcription error, as the candidate deals with the intended quadratic equation. The candidate has been given the benefit of the doubt and all marks awarded.

$$x^2 + 5x + 7 = 9x + 4$$
$$x - 4x + 3 = 0$$
$$(x - 3)(x - 1) = 0$$
$$x = 1 \text{ or } 3$$

(i) **Horizontal/vertical marking**

If a question results in two pairs of solutions, apply the following technique, but only if indicated in the detailed marking instructions for the question.

Example:

$$\bullet^5 \qquad \bullet^6$$
$$\bullet^5 \quad x = 2 \quad x = -4$$
$$\bullet^6 \quad y = 5 \quad y = -7$$

Horizontal: $\bullet^5\ x = 2$ and $x = -4$ Vertical: $\bullet^5\ x = 2$ and $y = 5$
               $\bullet^6\ y = 5$ and $y = -7$ $\bullet^6\ x = -4$ and $y = -7$

You must choose whichever method benefits the candidate, **not** a combination of both.

(j) In final answers, candidates should simplify numerical values as far as possible unless specifically mentioned in the detailed marking instruction. For example

$\dfrac{15}{12}$ must be simplified to $\dfrac{5}{4}$ or $1\dfrac{1}{4}$ $\dfrac{43}{1}$ must be simplified to 43

$\dfrac{15}{0\cdot3}$ must be simplified to 50 $\dfrac{4/5}{3}$ must be simplified to $\dfrac{4}{15}$

$\sqrt{64}$ must be simplified to 8*

*The square root of perfect squares up to and including 100 must be known.

(k) Do not penalise candidates for any of the following, unless specifically mentioned in the detailed marking instructions:

- working subsequent to a correct answer
- correct working in the wrong part of a question
- legitimate variations in numerical answers/algebraic expressions, for example angles in degrees rounded to nearest degree
- omission of units
- bad form (bad form only becomes bad form if subsequent working is correct), for example

$\left(x^3 + 2x^2 + 3x + 2\right)\left(2x + 1\right)$ written as

$\left(x^3 + 2x^2 + 3x + 2\right) \times 2x + 1$

$= 2x^4 + 5x^3 + 8x^2 + 7x + 2$

gains full credit

- repeated error within a question, but not between questions or papers

(l) In any 'Show that…' question, where candidates have to arrive at a required result, the last mark is not awarded as a follow-through from a previous error, unless specified in the detailed marking instructions.

(m) You must check all working carefully, even where a fundamental misunderstanding is apparent early in a candidate's response. You may still be able to award marks later in the question so you must refer continually to the marking instructions. The appearance of the correct answer does not necessarily indicate that you can award all the available marks to a candidate.

(n) You should mark legible scored-out working that has not been replaced. However, if the scored-out working has been replaced, you must only mark the replacement working.

(o) If candidates make multiple attempts using the same strategy and do not identify their final answer, mark all attempts and award the lowest mark. If candidates try different valid strategies, apply the above rule to attempts within each strategy and then award the highest mark.

For example:

Strategy 1 attempt 1 is worth 3 marks.	Strategy 2 attempt 1 is worth 1 mark.
Strategy 1 attempt 2 is worth 4 marks.	Strategy 2 attempt 2 is worth 5 marks.
From the attempts using strategy 1, the resultant mark would be 3.	From the attempts using strategy 2, the resultant mark would be 1.

In this case, award 3 marks.

Marking instructions for each question

Question			Generic scheme	Illustrative scheme	Max mark
1.			•^{1} calculate quartiles •^{2} calculate upper fence •^{3} comment	•^{1} 3 and 13 •^{2} $13 + 1\cdot5 \times 10 = 28$ •^{3} 29 is an outlier	3
2.			•^{1} state definition •^{2} give an example •^{3} state definition •^{4} give an example	•^{1} discrete data usually involves counting •^{2} for example, the number of tomatoes on a plant •^{3} continuous data usually involves measuring •^{4} for example, the heights of pupils in a class	4
3.	(a)		•^{1} show strategy •^{2} calculate probability	•^{1} •^{2} $P(S) = 0\cdot33 + 0\cdot075 + 0\cdot005 = 0\cdot410$	2
	(b)		•^{3} show strategy •^{4} calculate probability •^{5} calculate probability	•^{3} tree diagram above •^{4} $P(\bar{S}) = 1 - 0\cdot41 = 0\cdot59$ •^{5} $P(C\,|\,\bar{S}) = \dfrac{0\cdot095}{0\cdot59} = 0\cdot161$	3

	Question		Generic scheme	Illustrative scheme	Max mark
4.	(a)		$\bullet^1$ state hypotheses	$\bullet^1$ H_0: there is no association between clinic and maternal outcome H_1: there is an association between clinic and maternal outcome	**6**
			$\bullet^2$ calculate E_i	$\bullet^2$ expected frequencies are <table><tr><td>231·12</td><td>206·88</td></tr><tr><td>2828·88</td><td>2532·12</td></tr></table>	
			$\bullet^3$ calculate value of chi-squared	$\bullet^3$ $x^2 = 18 \cdot 22$	
			$\bullet^4$ state critical value	$\bullet^4$ $\chi^2_{0 \cdot 001,1} = 10 \cdot 827$	
			$\bullet^5$ deal with H_0	$\bullet^5$ $18 \cdot 22 > 10 \cdot 827$ so we reject H_0 at the 0·1% level of significance	
			$\bullet^6$ write conclusion	$\bullet^6$ and conclude that there is very strong evidence of an association between clinic and maternal outcome	
	(b)		$\bullet^7$ continue conclusion	$\bullet^7$ and there is evidence of a higher death rate in Clinic 1	**1**
5.	(a)		$\bullet^1$ show strategy	$\bullet^1$ 3-σ limits are $\mu \pm 3 \dfrac{\sigma}{\sqrt{n}}$	**2**
			$\bullet^2$ calculate both limits	$\bullet^2$ giving 69·5 N and 79·1 N	
	(b)		$\bullet^3$ state WECO rule	$\bullet^3$ samples 8 and 9 are above 2-σ (WECO: 2 out of 3 consecutive points above 2-σ)	**2**
			$\bullet^4$ comment	$\bullet^4$ pleased with very high tensile strength	
	(c)		$\bullet^5$ offer idea	$\bullet^5$ start a new control chart with	**2**
			$\bullet^6$ suggest change	$\bullet^6$ the limits set a bit higher	

	Question		Generic scheme	Illustrative scheme	Max mark
6	(a)		$\bullet^1$ calculate $E(X)$ $\bullet^2$ calculate $V(X)$	$\bullet^1$ $X : U(8)$ and $E(X) = \dfrac{8+1}{2} = \dfrac{9}{2}$ $\bullet^2$ $V(X) = \dfrac{8^2 - 1}{12} = \dfrac{21}{4}$	2
	(b)		$\bullet^3$ calculate $E(Y)$ and $V(Y)$ $\bullet^4$ calculate $E(3X - 2Y)$ $\bullet^5$ square and add $\bullet^6$ calculate $V(3X - 2Y)$	$\bullet^3$ $E(Y) = \dfrac{k+1}{2} = \dfrac{7}{2}, V(Y) = \dfrac{k^2 - 1}{12} = \dfrac{35}{12}$ $\bullet^4$ $E(3X - 2Y) = 3 \times \dfrac{9}{2} - 2 \times \dfrac{7}{2} = \dfrac{13}{2}$ $\bullet^5$ $V(3X - 2Y) = 9 \times \dfrac{21}{4} + 4 \times \dfrac{35}{12}$ $\bullet^6$ $\dfrac{707}{12}$	4
7.	(a)		$\bullet^1$ show strategy $\bullet^2$ continue strategy $\bullet^3$ calculate probability	$\bullet^1$ $X \sim N(53\cdot4, \ 4\cdot7^2) \Rightarrow P(X > 62)$ $\bullet^2$ $= P\left(Z > \dfrac{62 - 53\cdot4}{4\cdot7}\right)$ $\bullet^3$ $= P(Z > 1\cdot83) = 0\cdot0336$	3
	(b)		$\bullet^4$ state assumption $\bullet^5$ calculate z $\bullet^6$ show strategy $\bullet^7$ calculate μ	$\bullet^4$ the variability remains unchanged $\bullet^5$ $z = 2\cdot33$ $\bullet^6$ $\dfrac{62 - \mu}{4\cdot7} = 2\cdot33$ $\bullet^7$ $\mu < 51\cdot05$	4

	Question		Generic scheme	Illustrative scheme	Max mark
8.			$\bullet^1$ state hypotheses	$\bullet$ ^{1} $H_o : \eta_A = \eta_B$ $\qquad H_1 : \eta_A < \eta_B$	6
			$\bullet^2$ find rank sum	$\bullet^2$ $W_B = 5 + 7 \cdot 5 + 7 \cdot 5 + 9 \cdot 5 + 13 + 14 + 15 = 71 \cdot 5$	
			$\bullet^3$ calculate test statistics	$\bullet^3$ $m(m + n - 1) - W_B = 40 \cdot 5$	
			$\bullet^4$ state critical value	$\bullet^4$ cv = 41	
			$\bullet^5$ deal with H_0	$\bullet^5$ $40 \cdot 5 < 41$ so we reject H_0 at the 5% level of significance	
			$\bullet^6$ write conclusion	$\bullet^6$ and conclude that there is evidence in support of the claim by the commuter's colleague	
9.	(a)		$\bullet^1$ show strategy	$\bullet^1$ an approximate 95% CI is given by $$\bar{x} \pm t_{n-1,0 \cdot 975} \frac{s}{\sqrt{n}}$$	5
			$\bullet^2$ calculate mean and standard deviation	$\bullet^2$ $\bar{x} = 57 \cdot 96$, $s = 2 \cdot 692$	
			$\bullet^3$ state critical value	$\bullet^3$ $t_{8,0 \cdot 975} = 2 \cdot 306$	
			$\bullet^4$ calculate interval	$\bullet^4$ 95% CI is $(55 \cdot 89, 60 \cdot 03)$	
			$\bullet^5$ state assumption	$\bullet^5$ scores are normally distributed	
	(b)		$\bullet^6$ offer a reason	$\bullet^6$ the whole CI is above 55% and so	2
			$\bullet^7$ comment	$\bullet^7$ there is evidence of an improvement in scores	

	Question		Generic scheme	Illustrative scheme	Max mark
10.	(a)		$\bullet^1$ state hypotheses	$\bullet^1$ $H_0 : \mu = 500 \cdot 30$ $H_1 : \mu \neq 500 \cdot 30$	5
			$\bullet^2$ calculate mean	$\bullet^2$ $\bar{x} = 500 \cdot 265$	
			$\bullet^3$ calculate z	$\bullet^3$ $z = \dfrac{x - \mu}{\sigma / \sqrt{n}} = \dfrac{500 \cdot 265 - 500 \cdot 3}{0 \cdot 1 / \sqrt{10}}$ $= -1 \cdot 11$	
			$\bullet^4$ deal with H_0	$\bullet^4$ $-1 \cdot 11 > -1 \cdot 96$ so we cannot reject H_0 at the 5% level of significance	
			$\bullet^5$ write conclusion	$\bullet^5$ and there is no evidence that the mean fill volume differs from $500 \cdot 30$ ml	
	(b)		$\bullet^6$ comment	$\bullet^6$ if the population variance was unknown	2
			$\bullet^7$ continue comment	$\bullet^7$ it may lead to evidence that the mean volume differs from $500 \cdot 30$ ml	
11.	(a)		$\bullet^1$ give an explanation	$\bullet^1$ regressing y on x assumes that the residuals are in y and that x is the independent variable	2
			$\bullet^2$ continue explanation	$\bullet^2$ changing the subject does not alter this and we require the assumption that the residuals are in x	
	(b)		$\bullet^3$ calculate b	$\bullet^3$ $b = \dfrac{S_{xy}}{S_{yy}} = 4 \cdot 931$	3
			$\bullet^4$ find correct least squares regression line	$\bullet^4$ $a = \bar{x} - b\bar{y} = -166 \cdot 3$ $\Leftrightarrow x = 4 \cdot 931y - 166 \cdot 3$	
			$\bullet^5$ calculate x	$\bullet^5$ $x = 4 \cdot 931 \times 50 - 166 \cdot 3 = 80 \cdot 25$	

Question			Generic scheme	Illustrative scheme	Max mark
12.			$\bullet^1$ state hypotheses	$\bullet^1$ $H_0 : p_1 = p_2$ $H_1 : p_1 \neq p_2$	7
			$\bullet^2$ decide on correct test statistic	$\bullet^2$ $z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$	
			$\bullet^3$ calculate $\hat{p}$	$\bullet^3$ $\hat{p} = \dfrac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = 0\cdot528$	
			$\bullet^4$ substitute correctly	$\bullet^4$ $z = \dfrac{0\cdot56 - 0\cdot48}{\sqrt{0\cdot528 \times 0\cdot472\left(\dfrac{1}{300} + \dfrac{1}{200}\right)}}$	
			$\bullet^5$ calculate z	$\bullet^5$ $= 1\cdot76$	
			$\bullet^6$ deal with H_0	$\bullet^6$ $1\cdot76 < 1\cdot96$ so we cannot reject H_0 at the 5% level of significance and	
			$\bullet^7$ write conclusion	$\bullet^7$ conclude that there is no evidence of a difference in preference between the two age groups	

Question			Generic scheme	Illustrative scheme	Max mark
13.	(a)		•¹ state distribution	•¹ $X : \text{B}(200, 0\cdot5)$	5
			•² state approximation	•² $Y : \text{N}(100, 50)$	
			•³ use continuity correction	•³ $\text{P}(X \geq 110) \approx \text{P}(Y > 109\cdot5)$	
			•⁴ show strategy	•⁴ $\text{P}\left(Z > \dfrac{109\cdot5 - 100}{\sqrt{50}}\right)$	
			•⁵ calculate probability	•⁵ $= 0\cdot0901$	
	(b)		•⁶ state distribution	•⁶ $\text{P}\left(X < 110 \mid X : \text{B}(200, 0\cdot55)\right)$	4
			•⁷ state approximation	•⁷ $\text{P}\left(Y < 109\cdot5 \mid Y : \text{N}(110, 49\cdot5)\right)$	
			•⁸ show strategy	•⁸ $\text{P}\left(Z < \dfrac{109\cdot5 - 110}{\sqrt{49\cdot5}}\right)$	
			•⁹ calculate probability	•⁹ $0\cdot4721$	
	(c)		•¹⁰ offer a suggestion	•¹⁰ the coin could be tossed more than 200 times	1

Question			Generic scheme	Illustrative scheme	Max mark
14.	(a)		•[1] state expectation	•[1] 2	7
			•[2] use $\dfrac{\sigma^2}{n}$	•[2] $\dfrac{4}{100}+\mathrm{K}$	
			•[3] calculate standard error	•[3] $\sqrt{\dfrac{4}{100}+\dfrac{1}{100}}=0\cdot2236$	
			•[4] appreciate there are two possible differences	•[4] $P\left(-1\cdot5<\overline{F}-\overline{G}<1\cdot5\right)$	
			•[5] invoke the CLT	•[5] $=P(....<Z<....)$	
			•[6] show strategy	•[6] $P\left(\dfrac{-1\cdot5-2}{0\cdot2236}<Z<\dfrac{1\cdot5-2}{0\cdot2236}\right)$	
			•[7] calculate probability	•[7] $0\cdot0127$	
	(b)		•[8] state CLT	•[8] the CLT states that the distribution of sample means is only approximately normal if the parent distributions are not themselves normal	3
			•[9] give a justification	•[9] the samples are large	
			•[10] continue justification	•[10] so the parent population distribution doesn't matter	

[END OF SPECIMEN MARKING INSTRUCTIONS)

page 12