National
Qualifications
ADDITIONAL SPECIMEN ONLY

X803/77/11

Statistics
Paper 1

Date — Not applicable

Duration — 1 hour

Total marks — 30

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.

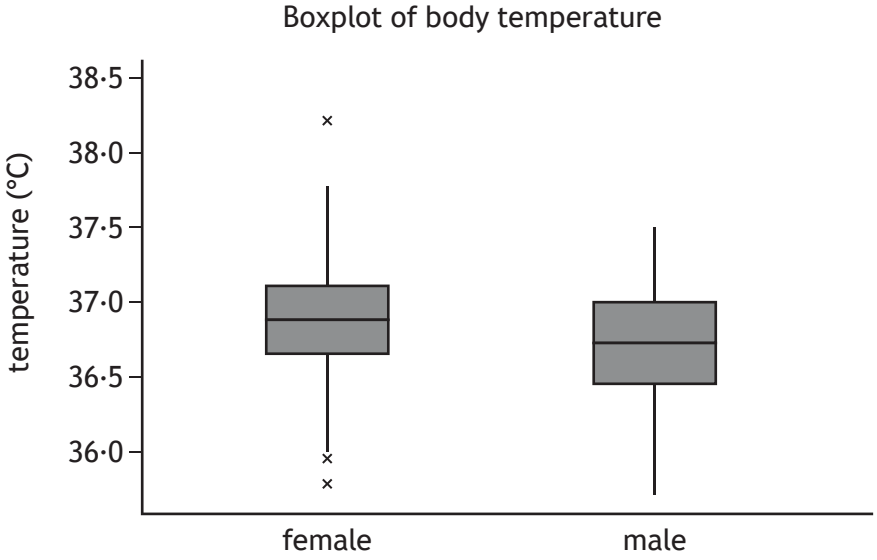
X8037711

Total marks — 30 **MARKS**

Attempt ALL questions

1. The average body temperature of a healthy adult human being is widely quoted as 37 °C. In the 1980s, a group of 130 volunteer adults who were known to have no underlying medical conditions which could affect body temperature, had their body temperature (°C) measured under standardised conditions.

A boxplot and summary statistics of the recorded body temperatures were obtained for the males and females separately using a statistical computing package. These are shown below.



Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Female Temp	65	36.885	0.413	35.778	36.667	36.889	37.111	38.222
Male Temp	65	36.725	0.388	35.722	36.444	36.722	37.000	37.500

(a) Comment briefly on whether the average body temperatures of males and females appear to be different from 37 °C, and whether there appears to be a difference in average body temperature between males and females. **3**

Assuming the body temperatures are normally distributed, a two-tailed hypothesis test was conducted, to compare the mean body temperatures in the wider populations of healthy adult males and females. Output from this test is given below.

```
Two-sample t for Female Temp vs Male Temp

               N    Mean   StDev  SE Mean
Female Temp   65  36.885   0.413    0.051
Male Temp     65  36.725   0.388    0.048

Difference = μ (Female Temp) - μ (Male Temp)
Estimate for difference: 0.1607

P-value: 0.0239
```

(b) Write down the hypotheses being tested and state the conclusion you would draw from this test. **3**

MARKS

1. (continued)

As both of the sample sizes are large, a z-test for a difference in population means could also have been conducted.

(c) Calculate the p-value for this test, and write down your conclusion. **4**

(d) Considering how the sample of 130 adults was selected, state a weakness in the method used and a resulting consequence. **2**

[Turn over

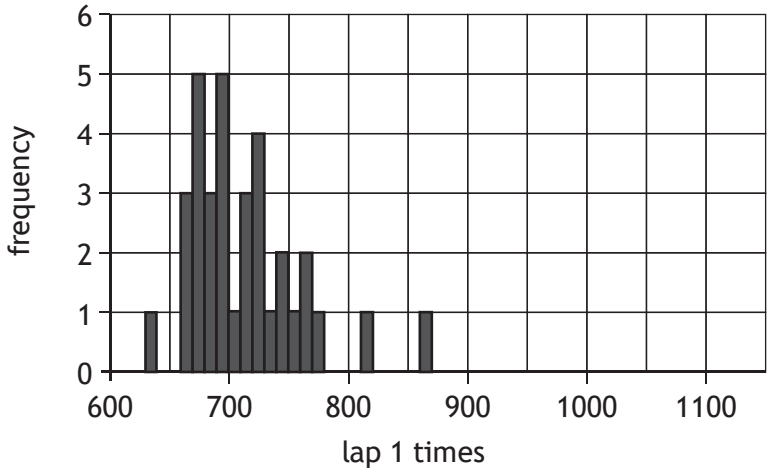
2. An extract from a draft report by a researcher is given below.

Read it and then answer the questions that follow.

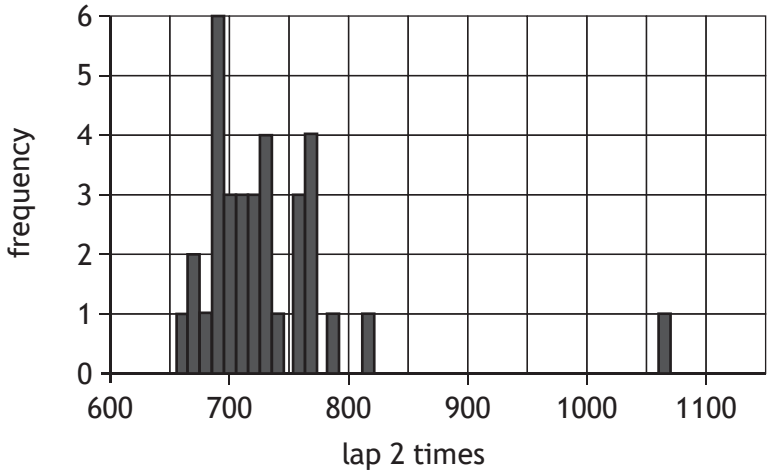
Just outside the centre of Edinburgh is Arthur's Seat, a 250 m high hill that has a road running up and around it, through Holyrood Park. Over the last year, depending on the daylight and amount of my time available, I rode either 2 or 3 complete laps of Arthur's Seat each week. Using a GPS tracking app on a smartphone, I recorded the times for cycling the uphill section of a lap on 34 separate days.

This investigation seeks to determine whether there is any difference in my times between the laps I completed.

I shall first look at all the data for laps 1 and 2 only, but I will observe lap 3 data, for the days that I have it, later on. Times, in seconds, for laps 1 and 2 are shown below.



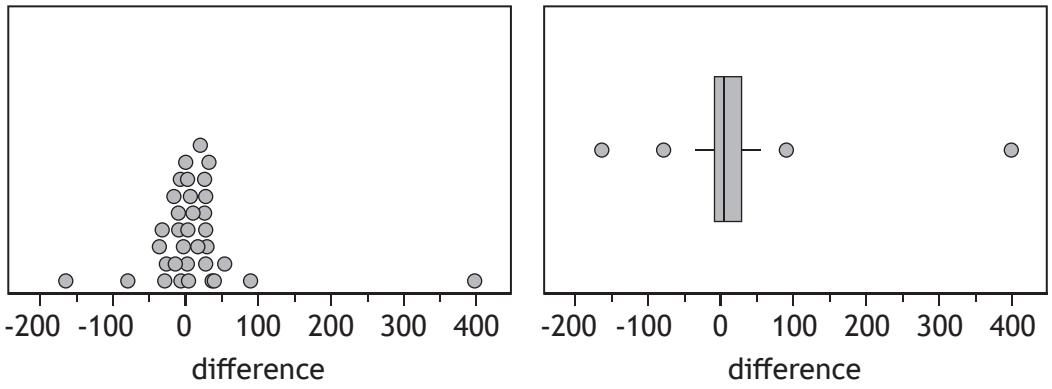
$\bar{x} = 713{\cdot}559, s = 45{\cdot}872$



$\bar{x} = 727{\cdot}795, s = 70{\cdot}486$

2. (continued)

There were some longer times due to mechanical failures and stoppages. I decided to calculate the paired difference between lap 1 and lap 2 times (lap 2 minus lap 1), to produce the following dotplot and boxplot.



This clearly highlighted 4 outliers that have come from when one lap time was considerably longer than the other.

I decided to remove the two most extreme outliers, leaving the two inner outliers as these represented a difference in lap times of under 100 seconds.

The summary data for this reduced data set of differences is given below.

Σx	249	Min	−78
Σx^2	30853	Q_1	−8.5
s	30.5411	Median	6
n	32	Q_3	29
		Max	89

I then performed a two-tailed paired t-test and a two-tailed Wilcoxon Signed Rank Test (using the normal approximation) on this data to establish if the average change in time was 0 seconds and I obtained p-values of 0·1595 and 0·1207 respectively.

Returning to the original data for lap times, and treating them as three independent samples, I calculated 95% confidence intervals for the mean time taken for each to be as follows.

lap 1 times	(697·554, 729·564)	$n = 34$
lap 2 times	(703·195, 752·394)	$n = 34$
lap 3 times	(709·328, 752·839)	$n = 12$

I conclude that there is no difference in my lap times, whether it is lap 1, lap 2 or lap 3. This surprised me as I would have expected to have recorded progressively slower times on each subsequent lap as I become more tired.

MARKS

2. (continued)

(a) (i) Calculate the upper and lower fences for the difference in lap times, using the summary data after line 17. **2**

(ii) Read lines 15 to 16.

Give a reason for the removal of the two data points. Referring to information in the report, give a further reason why this decision could be justified. **2**

In lines 18 to 20, two hypothesis tests are mentioned.

(b) (i) For each test, state the null and alternative hypotheses, the underlying assumptions, and whether these assumptions appear to be met, upon the removal of the two most extreme outliers. **5**

(ii) Given the researcher's expectation in the final sentence of the report, state how the alternative hypotheses could be improved. **1**

(c) (i) Calculate the sample mean of the lap 3 times, and show that the standard deviation is 34·2405. **4**

(ii) Comment on the three mean lap times in relation to the last sentence of the report. **1**

In lines 24 to 26, the researcher rather abruptly concludes that there is no difference in lap times, with no apparent rationale or reason for this conclusion.

(d) Write a conclusion that makes reference to the given p-values and confidence intervals that justifies arriving at the same decision. **2**

(e) A reviewer does not agree with the conclusion of this research.

They suggest that the study should be repeated with some changes.

State one change that would give more representative results. **1**

[END OF ADDITIONAL SPECIMEN QUESTION PAPER]

National
Qualifications
ADDITIONAL SPECIMEN ONLY

X803/77/11

Statistics
Paper 1

Marking Instructions

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

SQA
©

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.

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

$$x^2 + 5x + 7 = 9x + 4$$
$$x - 4x + 3 = 0$$
$$x = 1$$

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:

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

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

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

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

$\frac{15}{0{\cdot}3}$ must be simplified to 50 $\quad$ $\frac{4\!\!\not/5}{3}$ must be simplified to $\frac{4}{15}$

$\sqrt{64}$ 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

  $\left(x^3+2x^2+3x+2\right)(2x+1)$ 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.	(a)		•[1] comment	•[1] The sample mean and median values for both females and males are slightly below 37°C, so the evidence suggests that both population averages might be lower than claimed.	**3**
			•[2] comment	•[2] The distribution of male temperatures is lower than the distribution of female temperatures.	
			•[3] comment	•[3] All these differences are very small.	
	(b)		•[4] state hypotheses	•[4] $\begin{cases} H_0 : \mu_F = \mu_M \\ H_1 : \mu_F \neq \mu_M \end{cases}$	**3**
			•[5] state comparison	•[5] The p-value is less than 0·05 so reject H_0	
			•[6] state conclusion	•[6] and conclude that there is strong evidence that the population mean body temperature is different in the population of healthy adult females and males.	
	(c)		•[7] correct substitution	•[7] $z = \dfrac{36{\cdot}885 - 36{\cdot}725}{\sqrt{\dfrac{0{\cdot}413^2}{65} + \dfrac{0{\cdot}388^2}{65}}}$	**4**
			•[8] calculate z-value	•[8] $z = 2{\cdot}2764$	
			•[9] calculate p-value	•[9] $\begin{cases} p\text{-value} = 2 \times \mathrm{P}(Z > 2{\cdot}28) \\ \qquad\quad = 0{\cdot}0226 \end{cases}$	
			•[10] state conclusion	•[10] same conclusion as the two sample t-test in part 2(b)	
	(d)		•[11] state weakness	•[11] a non-random sampling method was used	**2**
			•[12] comment	•[12] statistical inference to the wider population of adults' temperatures is not valid	

Question			Generic scheme	Illustrative scheme	Max mark
2.	(a)	(i)	•[1] calculate IQR	•[1] $\text{IQR} = 29 - (-8{\cdot}5) = 37{\cdot}5$	**2**
			•[2] calculate fences	•[2] fences are $-8{\cdot}5 - 1{\cdot}5 \times 37{\cdot}5 = -64{\cdot}75$ and $29 + 1{\cdot}5 \times 37{\cdot}5 = 85{\cdot}25$	
		(ii)	•[3] appropriate comment	•[3] these two points are outside the fences	**2**
			•[4] appropriate comment	•[4] data points from mechanical failure or stoppages should not be included	
	(b)	(i)		*t*-test	**5**
			•[5] state hypotheses	•[5] H_0 : mean change in lap times is zero H_1 : mean change in lap times is not zero	
			•[6] state assumption	•[6] assumes parent population of changes to be normally distributed	
				Wilcoxon Paired Signed Rank Test	
			•[7] state hypothesis	•[7] H_0 : the median change in lap times is zero H_1 : the median change in lap times is not zero	
			•[8] state assumption	•[8] assumes parent population of changes to be symmetrical	
			•[9] state conclusion	•[9] both assumptions are supported by the dotplot (once the two most extreme outliers are removed)	
		(ii)	•[10] state appropriate comment	•[10] using one-tailed tests would be more appropriate	**1**

Question			Generic scheme	Illustrative scheme	Max mark
2.	(c)	(i)	•[11] calculate sample mean •[12] show strategy •[13] state critical value •[14] calculate s	•[11] $\overline{x} = \frac{709 \cdot 328 + 752 \cdot 839}{2}$ $= 731 \cdot 0835$ •[12] $\overline{x} + t\frac{s}{\sqrt{n}} = 752 \cdot 839$ •[13] $t_{11,0 \cdot 975} = 2 \cdot 201$ (from tables) •[14] $s = \left(\frac{21 \cdot 7555}{2 \cdot 201}\right)\sqrt{12}$ $= 34 \cdot 2405$	**4**
Notes:					
1. Alternatively, using the lower confidence limit could equally be used.					
		(ii)	•[15] appropriate comment	•[15] increasing mean lap times support the researcher's expectation	**1**
	(d)		•[16] write conclusion •[17] continue conclusion	•[16] both p-values were over 5%, so do not reject either H_0 •[17] all 3 confidence intervals have a lot of overlap, so one mean lap time is not obviously longer or shorter than the others	**2**
Notes:					
1. For mark •[17], the response must not refer to specific levels of significance arising from comparing the confidence intervals.					
	(e)		•[18] suggest change	•[18] for example, more riders; more laps; standard time of day for laps	**1**

[END OF ADDITIONAL SPECIMEN MARKING INSTRUCTIONS]