Please write clearly in block capitals.

Centre number | | | | | | Candidate number | | | | |

Surname ______________________________

Forename(s) ______________________________

Candidate signature ______________________________

A-level
STATISTICS

Unit Statistics 6

Monday 26 June 2017 Afternoon Time allowed: 1 hour 30 minutes

Materials
For this paper you must have:
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Instructions
- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- Write the question part reference (eg (a), (b)(i) etc) in the left-hand margin.
- You must answer each question in the space provided for that question. If you require extra space, use an AQA supplementary answer book; do **not** use the space provided for a different question.
- Do not write outside the box around each page.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work that you do not want to be marked.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 75.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.
- You do not necessarily need to use all the space provided.

For Examiner's Use	
Question	Mark
1	
2	
3	
4	
5	
TOTAL	

Do not write outside the box

Answer **all** questions.

Answer each question in the space provided for that question.

1 A scientist investigated the yield of coffee beans obtained from two species of coffee tree: *Robusta* and *Arabica.* He believed that *Robusta* trees would, on average, give a higher yield than *Arabica* trees.

The scientist planted each of ten plots of land with trees of both species and waited until all the trees were producing fruit.

(a) Two trees, one of each species, were selected randomly from the healthy trees on each plot. The total weight, in grams, of the coffee beans obtained from each of the selected trees was recorded and is given in **Table 1**.

Table 1

	Plot									
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
Robusta	623	654	733	704	822	669	980	892	805	762
Arabica	490	550	820	790	680	568	795	725	596	665

Making any necessary assumptions, carry out a paired t-test, using the 1% level of significance, to investigate whether the given data support the scientist's belief.

[8 marks]

(b) Beans from coffee trees are dried, roasted and mixed to produce many different blends of coffee.

Two tasters were asked to assign a score out of 100, where 100 indicates the best tasting coffee, to 9 randomly selected blends.

The scores assigned are given in **Table 2**.

Table 2

	Blend								
	R	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
Taster 1	78	53	85	88	76	84	60	65	86
Taster 2	73	57	78	96	82	84	62	61	94

(i) Carry out a Wilcoxon signed-rank test, using the 10% level of significance, to investigate whether, on average, the scores assigned by the two tasters differ.

[7 marks]

(ii) For a Wilcoxon signed-rank test carried out on 9 matched pairs, T, the sum of the ranks with positive differences, is calculated. Find the maximum possible value of T.

[2 marks]

Do not write outside the box

QUESTION PART REFERENCE	**Answer space for question 1**

QUESTION PART REFERENCE	Answer space for question 1

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 1

17

Turn over ►

2 A steel mill supplies sheet steel to a manufacturer that sets its machines assuming that the thickness of the sheet steel is 0.750 mm.

The steel mill's quality control manager, Roman, decides to set up control charts to monitor production so that the thickness of the sheet steel supplied is as close to 0.750 mm as possible.

Roman takes ten samples, each of 5 measurements, when production is thought to be satisfactory.

The sample means and ranges, both in millimetres, for these samples are given in **Table 3**.

Table 3

Sample	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Mean	0.748	0.762	0.751	0.758	0.757	0.754	0.749	0.760	0.766	0.741
Range	0.052	0.041	0.045	0.045	0.025	0.047	0.052	0.060	0.054	0.045

(a) Assuming that the measurements are normally distributed, use the mean value of the ranges given in the table to show that 0.020 mm is a suitable estimate, correct to three decimal places, for the value of the standard deviation of the process.

[3 marks]

(b) Using a standard deviation of 0.020 mm and a target value of 0.750 mm, calculate upper and lower warning (95%) and action (99.8%) control limits for charts for:

(i) means;

(ii) standard deviations.

[6 marks]

(c) Comment on the current state of the manufacturing process with regard to means only.

[1 mark]

(d) The measurements, in millimetres, for the next two samples, Sample 11 and Sample 12, are:

Sample 11: $0.748,\ 0.744,\ 0.798,\ 0.784,\ 0.788$

Sample 12: $0.726,\ 0.728,\ 0.788,\ 0.702,\ 0.802$

Referring to your calculated control limits in parts **(b)(i)** and **(b)(ii)**, state, with reasons, what action, if any, you would advise as a result of:

(i) Sample 11;
(ii) Sample 12.

[6 marks]

Do not write outside the box

QUESTION PART REFERENCE	Answer space for question 2

Do not write outside the box

QUESTION PART REFERENCE	Answer space for question 2

Do not write outside the box

QUESTION PART REFERENCE	**Answer space for question 2**

16

Turn over ►

3 A professional cycling team has been using the cycle-chain lubricant L1 for a number of years. The company that produces L1 claims that its new lubricant, L2, further reduces friction between chain links thus resulting in faster race times.

The team's engineers decide to carry out an experiment to investigate this claim. They also decide to test two competitor brand lubricants, L3 and L4.

The team has four cyclists all of whom help with the experiment by each completing four 5000 metre rides.

Before starting the experiment, the lead engineer, Heather, puts the lubricants in unlabelled bottles, so that she alone knows which lubricant is in which bottle.

For each ride, Pete, another engineer, applies a set quantity of one lubricant to a clean chain on a cycle. The ride is then timed over 5000 metres in an indoor velodrome.

(a) (i) Explain, in the context of the question, what is meant by the term 'double blind' experiment.

(ii) Explain why a double blind experiment is beneficial for investigating the company's claim.

[4 marks]

(b) For this experiment, Pete suggests that lubricants should be randomly allocated, for each ride, to each cyclist.

(i) State the name of Pete's suggested experimental design.

(ii) Name an alternative experimental design, preferable to Pete's design, that could be used for this experiment.

You should include an explanation regarding why this alternative design is preferable.

(iii) Construct a table that identifies how the four lubricants should be allocated to the cyclists and rides using the alternative design that you named in part **(b)(ii)**.

[5 marks]

(c) Name a statistical analysis that Heather could undertake on the race times recorded if the alternative design that you named in part **(b)(ii)** is used.

[1 mark]

QUESTION PART REFERENCE	**Answer space for question 3**

Do not write outside the box

QUESTION PART REFERENCE	Answer space for question 3

10

Turn over ►

4 In the United States, over one million baseballs are manufactured annually for use in league games. Surrey League purchases all of its baseballs in large batches from one manufacturer. Surrey League and the manufacturer decide to set up a mutually agreeable acceptance sampling scheme.

Surrey League requires that, when the proportion of non-conforming baseballs in a batch is 10% , the probability of batch acceptance is at most 12%

The manufacturer requires that, when the proportion of non-conforming baseballs in a batch is 1% , the probability of batch rejection is at most 2%

Surrey League and the manufacturer consider three alternative sampling schemes.

Scheme A

- Select a random sample of 20 baseballs from a batch.
- Accept the batch only if no non-conforming baseballs are found in the sample; otherwise reject the batch.

Scheme B

- Select a random sample of 50 baseballs from a batch.
- Accept the batch if fewer than three non-conforming baseballs are found in the sample; otherwise reject the batch.

Scheme C

- Select a random sample of 40 baseballs from a batch.
- Accept the batch if fewer than two non-conforming baseballs are found in the sample.
- Reject the batch if more than three non-conforming baseballs are found in the sample.
- If two or three non-conforming baseballs are found in the sample, select a further sample of 40 baseballs.
- Accept the batch if a total of three or fewer non-conforming baseballs are found in the 80 baseballs sampled; otherwise reject the batch.

(a) Show that **Scheme A** will satisfy neither Surrey League's nor the manufacturer's requirements.

[4 marks]

QUESTION PART REFERENCE	**Answer space for question 4(a)**

Do not write outside the box

QUESTION PART REFERENCE	Answer space for question 4(a)

Question 4 continues on the next page

Do not write outside the box

4 (b) **Scheme B**

- Select a random sample of 50 baseballs from a batch.
- Accept the batch if fewer than three non-conforming baseballs are found in the sample; otherwise reject the batch.

Table 4 below shows, for **Scheme B**, the probability of acceptance of a batch for various proportions of non-conforming baseballs.

(i) Complete **Table 4**

[3 marks]

(ii) Hence indicate why **Scheme B** satisfies both Surrey League's and the manufacturer's requirements.

[1 mark]

QUESTION PART REFERENCE

Answer space for question 4(b)

Table 4

Proportion of non-conforming baseballs (P)	0.01	0.02	0.03	0.04	0.05	0.10
Probability of acceptance		0.922	0.811	0.677	0.541	

QUESTION PART REFERENCE	**Answer space for question 4(b)**

Question 4 continues on the next page

Turn over ►

4 (c) **Scheme C**

- Select a random sample of 40 baseballs from a batch.
- Accept the batch if fewer than two non-conforming baseballs are found in the sample.
- Reject the batch if more than three non-conforming baseballs are found in the sample.
- If two or three non-conforming baseballs are found in the sample, select a further sample of 40 baseballs.
- Accept the batch if a total of three or fewer non-conforming baseballs are found in the 80 baseballs sampled; otherwise reject the batch.

Table 5 below shows, using **Scheme C**, the probability of acceptance of a batch for various proportions of non-conforming baseballs.

Complete **Table 5**.

[4 marks]

QUESTION PART REFERENCE

Answer space for question 4(c)

Table 5

Proportion of non-conforming baseballs (P)	0.01	0.02	0.03	0.04	0.05	0.10
Probability of acceptance		0.943	0.833	0.686	0.534	0.095

Do not write outside the box

QUESTION PART REFERENCE	Answer space for question 4(c)

Question 4 continues on the next page

Turn over ►

4(d) **Table 4** and **Table 5** are repeated below.

Table 4

Proportion of non-conforming baseballs (P)	0.01	0.02	0.03	0.04	0.05	0.10
Probability of acceptance		0.922	0.811	0.677	0.541	

Table 5

Proportion of non-conforming baseballs (P)	0.01	0.02	0.03	0.04	0.05	0.10
Probability of acceptance		0.943	0.833	0.686	0.534	0.095

Draw the operating characteristic for **Scheme B** and **Scheme C** on **Figure 1** opposite.

[4 marks]

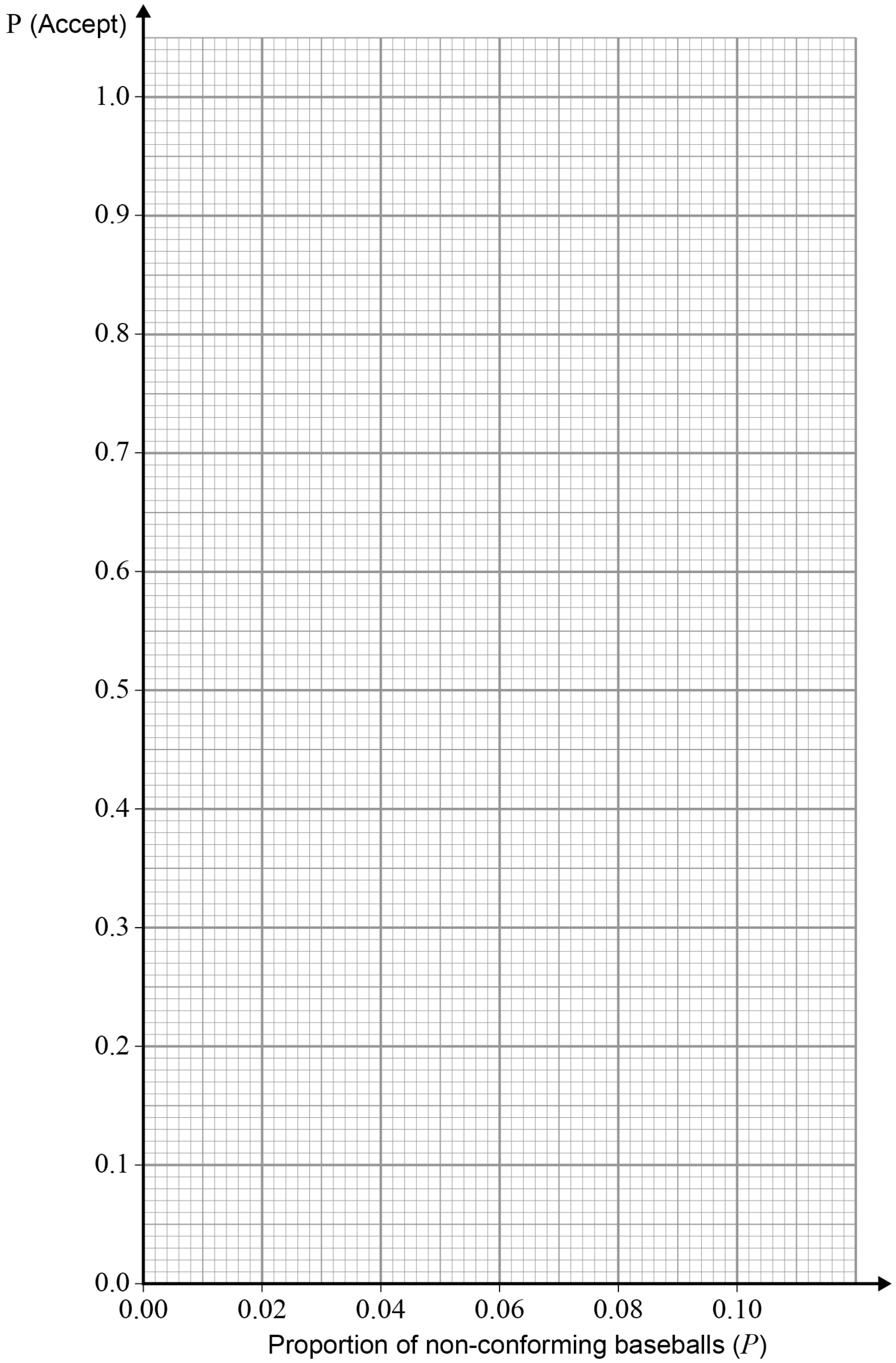
QUESTION PART REFERENCE	**Answer space for question 4(d)**

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 4(d)

Figure 1



Question 4 continues on the next page

4 (e) Surrey League decides that either **Scheme B** or **Scheme C** would be acceptable.

Which of **Scheme B** or **Scheme C** would be preferable to the manufacturer?

Give **two** reasons for your choice.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 4(e)

Do not write outside the box

QUESTION PART REFERENCE	Answer space for question 4(e)

18

Turn over ►

Do not write outside the box

5 Researchers set up a trial to investigate the effect of taking a vitamin supplement on a particular auto-immune disease.

Twenty-three adults, all of whom were at the same stage of the disease, were randomly allocated to receive one of three treatments:

- 'None' no vitamin supplement
- 'TR_1' a supplement of 750 International Units (IU) of the vitamin per day
- 'TR_2' a supplement of $10\ 500$ IU of the vitamin per day.

After six months of the trial, each adult had a blood test to measure the level of immune cells. A lower measurement indicated that the disease was progressing more slowly. **Table 6** gives the results of the blood test and the treatment allocated for each adult in the trial.

Table 6

Adult	Treatment allocated	Level of immune cells
A	TR_1	720
B	TR_2	650
C	None	820
D	None	940
E	TR_1	900
F	TR_1	790
G	TR_1	920
H	None	930
I	TR_2	710
J	None	880
K	None	860
L	TR_2	690
M	TR_2	710
N	TR_1	840
O	TR_1	870
P	None	790
Q	TR_1	810
R	TR_1	900
S	TR_2	620
T	TR_1	840
U	None	850
V	TR_2	700
W	TR_2	830

The sample of adults may be regarded as random and the level of immune cells, for each treatment, is normally distributed.

Investigate for a difference between the three treatments, with regard to the mean level of immune cells, for adults with the disease. Use the 1% level of significance.

[14 marks]

Do not write outside the box

QUESTION PART REFERENCE	Answer space for question 5

Turn over ►

Do not write outside the box

QUESTION PART REFERENCE	Answer space for question 5

14

END OF QUESTIONS