5. The heights of females from a country are normally distributed with

- a mean of 166.5 cm
- a standard deviation of 6.1 cm

Given that 1% of females from this country are shorter than k cm,

(a) find the value of k

(2)

(b) Find the proportion of females from this country with heights between 150 cm and 175 cm

(1)

A female, from this country, is chosen at random from those with heights between 150 cm and 175 cm

(c) Find the probability that her height is more than 160 cm

(4)

The heights of females from a different country are normally distributed with a standard deviation of 7.4 cm

Mia believes that the mean height of females from this country is less than 166.5 cm

Mia takes a random sample of 50 females from this country and finds the mean of her sample is 164.6 cm

(d) Carry out a suitable test to assess Mia's belief.
You should

- state your hypotheses clearly
- use a 5% level of significance

(4)

5. The lifetime, L hours, of a battery has a normal distribution with mean 18 hours and standard deviation 4 hours.

Alice's calculator requires 4 batteries and will stop working when any one battery reaches the end of its lifetime.

(a) Find the probability that a randomly selected battery will last for longer than 16 hours.
(1)

At the start of her exams Alice put 4 new batteries in her calculator.
She has used her calculator for 16 hours, but has another 4 hours of exams to sit.

(b) Find the probability that her calculator will not stop working for Alice's remaining exams.
(5)

Alice only has 2 new batteries so, after the first 16 hours of her exams, although her calculator is still working, she randomly selects 2 of the batteries from her calculator and replaces these with the 2 new batteries.

(c) Show that the probability that her calculator will not stop working for the remainder of her exams is 0.199 to 3 significant figures.
(3)

After her exams, Alice believed that the lifetime of the batteries was more than 18 hours. She took a random sample of 20 of these batteries and found that their mean lifetime was 19.2 hours.

(d) Stating your hypotheses clearly and using a 5% level of significance, test Alice's belief.
(5)

2.

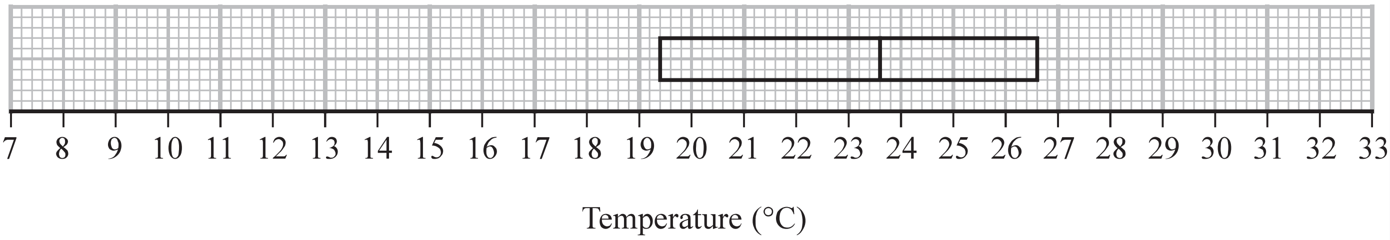

Figure 1

The partially completed box plot in Figure 1 shows the distribution of daily mean air temperatures using the data from the large data set for Beijing in 2015

An outlier is defined as a value
more than 1.5 × IQR below Q_1 or
more than 1.5 × IQR above Q_3

The three lowest air temperatures in the data set are 7.6 °C, 8.1 °C and 9.1 °C
The highest air temperature in the data set is 32.5 °C

(a) Complete the box plot in Figure 1 showing clearly any outliers.
(4)

(b) Using your knowledge of the large data set, suggest from which month the two outliers are likely to have come.
(1)

Using the data from the large data set, Simon produced the following summary statistics for the daily mean air temperature, x °C, for Beijing in 2015

$$n = 184 \qquad \sum x = 4153.6 \qquad S_{xx} = 4952.906$$

(c) Show that, to 3 significant figures, the standard deviation is 5.19 °C
(1)

Simon decides to model the air temperatures with the random variable

$$T \sim \mathrm{N}(22.6, 5.19^2)$$

(d) Using Simon's model, calculate the 10th to 90th interpercentile range.
(3)

Simon wants to model another variable from the large data set for Beijing using a normal distribution.

(e) State two variables from the large data set for Beijing that are **not** suitable to be modelled by a normal distribution. Give a reason for each answer.
(2)

5. A machine puts liquid into bottles of perfume. The amount of liquid put into each bottle, D ml, follows a normal distribution with mean 25 ml

Given that 15% of bottles contain less than 24.63 ml

(a) find, to 2 decimal places, the value of k such that $P(24.63 < D < k) = 0.45$

(5)

A random sample of 200 bottles is taken.

(b) Using a normal approximation, find the probability that fewer than half of these bottles contain between 24.63 ml and k ml

(3)

The machine is adjusted so that the standard deviation of the liquid put in the bottles is now 0.16 ml

Following the adjustments, Hannah believes that the mean amount of liquid put in each bottle is less than 25 ml

She takes a random sample of 20 bottles and finds the mean amount of liquid to be 24.94 ml

(c) Test Hannah's belief at the 5% level of significance.
You should state your hypotheses clearly.

(5)

5. A health centre claims that the time a doctor spends with a patient can be modelled by a normal distribution with a mean of 10 minutes and a standard deviation of 4 minutes.

(a) Using this model, find the probability that the time spent with a randomly selected patient is more than 15 minutes.

(1)

Some patients complain that the mean time the doctor spends with a patient is more than 10 minutes.

The receptionist takes a random sample of 20 patients and finds that the mean time the doctor spends with a patient is 11.5 minutes.

(b) Stating your hypotheses clearly and using a 5% significance level, test whether or not there is evidence to support the patients' complaint.

(4)

The health centre also claims that the time a dentist spends with a patient during a routine appointment, T minutes, can be modelled by the normal distribution where $T \sim N(5, 3.5^2)$

(c) Using this model,

(i) find the probability that a routine appointment with the dentist takes less than 2 minutes

(1)

(ii) find $P(T < 2 \mid T > 0)$

(3)

(iii) hence explain why this normal distribution may not be a good model for T.

(1)

The dentist believes that she cannot complete a routine appointment in less than 2 minutes.

She suggests that the health centre should use a refined model only including values of $T > 2$

(d) Find the median time for a routine appointment using this new model, giving your answer correct to one decimal place.

(5)

SECTION A: STATISTICS

Answer ALL questions. Write your answers in the spaces provided.

1. The number of hours of sunshine each day, y, for the month of July at Heathrow are summarised in the table below.

Hours	$0 \leqslant y < 5$	$5 \leqslant y < 8$	$8 \leqslant y < 11$	$11 \leqslant y < 12$	$12 \leqslant y < 14$
Frequency	12	6	8	3	2

A histogram was drawn to represent these data. The $8 \leqslant y < 11$ group was represented by a bar of width 1.5 cm and height 8 cm.

(a) Find the width and the height of the $0 \leqslant y < 5$ group.

(3)

(b) Use your calculator to estimate the mean and the standard deviation of the number of hours of sunshine each day, for the month of July at Heathrow.
Give your answers to 3 significant figures.

(3)

The mean and standard deviation for the number of hours of daily sunshine for the same month in Hurn are 5.98 hours and 4.12 hours respectably.
Thomas believes that the further south you are the more consistent should be the number of hours of daily sunshine.

(c) State, giving a reason, whether or not the calculations in part (b) support Thomas' belief.

(2)

(d) Estimate the number of days in July at Heathrow where the number of hours of sunshine is more than 1 standard deviation above the mean.

(2)

Helen models the number of hours of sunshine each day, for the month of July at Heathrow by $N(6.6, 3.7^2)$.

(e) Use Helen's model to predict the number of days in July at Heathrow when the number of hours of sunshine is more than 1 standard deviation above the mean.

(2)

(f) Use your answers to part (d) and part (e) to comment on the suitability of Helen's model.

(1)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

3. A machine cuts strips of metal to length L cm, where L is normally distributed with standard deviation 0.5 cm.

Strips with length either less than 49 cm or greater than 50.75 cm **cannot** be used.

Given that 2.5% of the cut lengths exceed 50.98 cm,

(a) find the probability that a randomly chosen strip of metal **can** be used.

(5)

Ten strips of metal are selected at random.

(b) Find the probability fewer than 4 of these strips **cannot** be used.

(2)

A second machine cuts strips of metal of length X cm, where X is normally distributed with standard deviation 0.6 cm

A random sample of 15 strips cut by this second machine was found to have a mean length of 50.4 cm

(c) Stating your hypotheses clearly and using a 1% level of significance, test whether or not the mean length of all the strips, cut by the second machine, is greater than 50.1 cm

(5)

5. A company sells seeds and claims that 55% of its pea seeds germinate.

(a) Write down a reason why the company should not justify their claim by testing all the pea seeds they produce.

(1)

A random selection of the pea seeds is planted in 10 trays with 24 seeds in each tray.

(b) Assuming that the company's claim is correct, calculate the probability that in at least half of the trays 15 or more of the seeds germinate.

(3)

(c) Write down two conditions under which the normal distribution may be used as an approximation to the binomial distribution.

(1)

A random sample of 240 pea seeds was planted and 150 of these seeds germinated.

(d) Assuming that the company's claim is correct, use a normal approximation to find the probability that at least 150 pea seeds germinate.

(3)

(e) Using your answer to part (d), comment on whether or not the proportion of the company's pea seeds that germinate is different from the company's claim of 55%

(1)

Leave blank

1. A shop sells rods of nominal length 200 cm. The rods are bought from a manufacturer who uses a machine to cut rods of length L cm, where $L \sim \mathrm{N}(\mu, 0.2^2)$

The value of μ is such that there is only a 5% chance that a rod, selected at random from those supplied to the shop, will have length less than 200 cm.

(a) Find the value of μ to one decimal place.

(3)

A customer buys a random sample of 8 of these rods.

(b) Find the probability that at least 3 of these rods will have length less than 200 cm.

(3)

Another customer buys a random sample of 60 of these rods.

(c) Using a suitable approximation, find the probability that more than 5 of these rods will have length less than 200 cm.

(3)

Leave blank

6. A fair 6-sided die is thrown n times. The number of sixes, X, is recorded. Using a normal approximation, $P(X < 50) = 0.0082$ correct to 4 decimal places.

Find the value of n.

(10)

Leave blank

1. A mobile phone company claims that each year 5% of its customers have their mobile phone stolen. An insurance company claims this percentage is higher. A random sample of 30 of the mobile phone company's customers is taken and 4 of them have had their mobile phone stolen during the last year.

(a) Test the insurance company's claim at the 10% level of significance. State your hypotheses clearly.

(6)

A new random sample of 90 customers is taken. A test is carried out using these 90 customers, to see if the percentage of customers who have had a mobile phone stolen in the last year is more than 5%

(b) Using a suitable approximation and a 10% level of significance, find the critical region for this test.

(4)

Leave blank

2. *Crispy-crisps* produces packets of crisps. During a promotion, a prize is placed in 25% of the packets. No more than 1 prize is placed in any packet. A box contains 6 packets of crisps.

(a) (i) Write down a suitable distribution to model the number of prizes found in a box.

(ii) Write down one assumption required for the model.

(2)

(b) Find the probability that in 2 randomly selected boxes, only 1 box contains exactly 1 prize.

(3)

(c) Find the probability that a randomly selected box contains at least 2 prizes.

(2)

Neha buys 80 boxes of crisps.

(d) Using a normal approximation, find the probability that no more than 30 of the boxes contain at least 2 prizes.

(5)

Leave blank

4. In a large population, past records show that 1 in 200 adults has a particular allergy.

In a random sample of 700 adults selected from the population, estimate

(a) (i) the mean number of adults with the allergy,

(ii) the standard deviation of the number of adults with the allergy.
Give your answer to 3 decimal places.

(3)

A doctor claims that the past records are out of date and the proportion of adults with the allergy is higher than the records indicate.

A random sample of 500 adults is taken from the population and 5 are found to have the allergy.

A test of the doctor's claim is to be carried out at the 5% level of significance.

(b) (i) State the hypotheses for this test.

(ii) Using a suitable approximation, carry out the test.

(6)

It is also claimed that 30% of those with the allergy take medication for it daily.

To test this claim, a random sample of n people with the allergy is taken. The random variable Y represents the number of people in the sample who take medication for the allergy daily.

A two-tailed test, at the 1% level of significance, is carried out to see if the proportion differs from 30%

The critical region for the test is $Y = 0$ or $Y \geqslant w$

(c) Find the smallest possible value of n and the corresponding value of w

(4)

Leave blank

6. Past information at a computer shop shows that 40% of customers buy insurance when they purchase a product. In a random sample of 30 customers, X buy insurance.

(a) Write down a suitable model for the distribution of X.

(1)

(b) State an assumption that has been made for the model in part (a) to be suitable.

(1)

The probability that fewer than r customers buy insurance is less than 0.05

(c) Find the largest possible value of r.

(2)

A second random sample, of 100 customers, is taken.

The probability that at least t of these customers buy insurance is 0.938, correct to 3 decimal places.

(d) Using a suitable approximation, find the value of t.

(6)

The shop now offers an extended warranty on all products. Following this, a random sample of 25 customers is taken and 6 of them buy insurance.

(e) Test, at the 10% level of significance, whether or not there is evidence that the proportion of customers who buy insurance has decreased. State your hypotheses clearly.

(5)

Leave blank

4. Pieces of ribbon are cut to length L cm where $L \sim \mathrm{N}(\mu, 0.5^2)$

(a) Given that 30% of the pieces of ribbon have length more than 100 cm, find the value of μ to the nearest 0.1 cm.

(3)

John selects 12 pieces of ribbon at random.

(b) Find the probability that fewer than 3 of these pieces of ribbon have length more than 100 cm.

(3)

Aditi selects 400 pieces of ribbon at random.

(c) Using a suitable approximation, find the probability that more than 127 of these pieces of ribbon will have length more than 100 cm.

(6)

Leave blank

6. A seed producer claims that 96% of its bean seeds germinate.

To test the producer's claim, a random sample of 75 bean seeds was planted and 66 of these seeds germinated.

Use a suitable approximation to test, at the 1% level of significance, whether or not the producer is overstating the probability of its bean seeds germinating. State your hypotheses clearly.

(7)

Leave blank

1. The probability of a leaf cutting successfully taking root is 0.05

Find the probability that, in a batch of 10 randomly selected leaf cuttings, the number taking root will be

(a) (i) exactly 1
(ii) more than 2

(5)

A second random sample of 160 leaf cuttings is selected.

(b) Using a suitable approximation, estimate the probability of at least 10 leaf cuttings taking root.

(3)

Leave blank

7. The random variable $Y \sim B(n, p)$.

Using a normal approximation the probability that Y is at least 65 is 0.2266 and the probability that Y is more than 52 is 0.8944

Find the value of n and the value of p.

(12)

Leave blank

3. Hei and Tang are designing some pieces of art. They collected a large number of sticks. The random variable L represents the length of a stick in centimetres and has a normal distribution with mean μ and standard deviation σ.

They sorted the sticks into lengths and painted them.

They found that 60% of the sticks were longer than 45 cm and these were painted red, whilst 15% of the sticks were shorter than 35 cm and these were painted blue. The remaining sticks were painted yellow.

(a) Show that μ and σ satisfy

$$45 + 0.2533\sigma = \mu$$

(2)

(b) Find a second equation in μ and σ.

(2)

(c) Hence find the value of μ and the value of σ.

(3)

(d) Find (i) $P(L > 35 \mid L < 45)$

(ii) $P(L < 45 \mid L > 35)$

(3)

Hei created her piece of art using a random selection of blue and yellow sticks.

Tang created his piece of art using a random selection of red and yellow sticks.

Hei and Tang each used the same number of sticks to create their piece of art.

George is viewing Hei's and Tang's pieces of art. He finds a yellow stick on the floor that has fallen from one of these pieces.

(e) With reference to your answers to part (d), state, giving a reason, whether the stick is more likely to have fallen from Hei's or Tang's piece of art.

(2)

Leave blank

1. The random variable $X \sim N(\mu, \sigma^2)$

Given that $P(X > \mu + a) = 0.35$ where a is a constant, find

(a) $P(X > \mu - a)$

(1)

(b) $P(\mu - a < X < \mu + a)$

(2)

(c) $P(X < \mu + a \mid X > \mu - a)$

(2)

Leave blank

3. At a school athletics day, the distances, in metres, achieved by students in the long jump are modelled by the normal distribution with mean 3.3 m and standard deviation 0.6 m

(a) Find an estimate for the proportion of students who jump less than 2.5 m

(3)

The long jump competition consists of 2 jumps. All the students can take part in the first jump and the 40% who jump the greatest distance in their first jump qualify for the second jump.

(b) Find an estimate for the minimum distance achieved in the first jump in order to qualify for the second jump.
Give your answer correct to 4 significant figures.

(3)

(c) Find an estimate for the median distance achieved in the first jump by those who qualify for the second jump.

(3)

The distance of the second jump is independent of the distance of the first jump and is modelled with the same normal distribution. Students who jump a distance greater than 4.1 m in their second jump receive a certificate.

At the start of the long jump competition, a student is selected at random.

(d) Find the probability that this student will receive a certificate.

(3)

Leave blank

7. A machine fills bottles with water. The volume of water delivered by the machine to a bottle is X ml where $X \sim \mathrm{N}(\mu, \sigma^2)$

One of these bottles of water is selected at random.

Given that $\mu = 503$ and $\sigma = 1.6$

(a) find

(i) $\mathrm{P}(X > 505)$
(ii) $\mathrm{P}(501 < X < 505)$

(5)

(b) Find w such that $\mathrm{P}(1006 - w < X < w) = 0.9426$

(3)

Following adjustments to the machine, the volume of water delivered by the machine to a bottle is such that $\mu = 503$ and $\sigma = q$

Given that $\mathrm{P}(X < r) = 0.01$ and $\mathrm{P}(X > r + 6) = 0.05$

(c) find the value of r and the value of q

(7)

Leave blank

5. Police measure the speed of cars passing a particular point on a motorway.
The random variable X is the speed of a car.

X is modelled by a normal distribution with mean 55 mph (miles per hour).

(a) Draw a sketch to illustrate the distribution of X. Label the mean on your sketch.
(2)

The speed limit on the motorway is 70 mph. Car drivers can choose to travel faster than the speed limit but risk being caught by the police.

The distribution of X has a standard deviation of 20 mph.

(b) Find the percentage of cars that are travelling faster than the speed limit.
(3)

The fastest 1% of car drivers will be banned from driving.

(c) Show that the lowest speed, correct to 3 significant figures, for a car driver to be banned is 102 mph. Show your working clearly.
(3)

Car drivers will just be given a caution if they are travelling at a speed m such that

$$\mathrm{P}(70 < X < m) = 0.1315$$

(d) Find the value of m. Show your working clearly.
(4)

Leave blank

7. One event at *Pentor* sports day is throwing a tennis ball. The distance a child throws a tennis ball is modelled by a normal distribution with mean 32 m and standard deviation 12 m. Any child who throws the tennis ball more than 50 m is awarded a gold certificate.

(a) Show that, to 3 significant figures, 6.68% of children are awarded a gold certificate.
(3)

A silver certificate is awarded to any child who throws the tennis ball more than d metres but less than 50 m.

Given that 19.1% of the children are awarded a silver certificate,

(b) find the value of d.
(4)

Three children are selected at random from those who take part in the throwing a tennis ball event.

(c) Find the probability that 1 is awarded a gold certificate and 2 are awarded silver certificates. Give your answer to 2 significant figures.
(3)

Leave blank

7. The weights, G, of a particular breed of gorilla are normally distributed with mean 180 kg and standard deviation 15 kg.

(a) Find the proportion of these gorillas whose weights exceed 174 kg.

(2)

(b) Find, to 1 decimal place, the value of k such that $P(k < G < 174) = 0.3196$

(4)

The weights, B, of a particular breed of buffalo are normally distributed with mean 216 kg and standard deviation 30 kg.

Given that $P(G > w) = P(B < w) = p$

(c) (i) find the value of w

(ii) find the value of p

(5)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

Leave blank

6. In a factory, a machine is used to fill bags of rice. The weights of bags of rice are modelled using a normal distribution with mean 250 g.

Only 1% of the bags of rice weigh more than 256 g.

(a) Write down the percentage of bags of rice with weights between 244 g and 256 g. **(1)**

(b) Find the standard deviation of the weights of the bags of rice. **(3)**

An inspection consists of selecting a bag of rice at random and checking if its weight is within 4 g of the mean. If the weight is more than 4 g away from the mean, then a second bag of rice is selected at random and checked. If the weight of each of the 2 bags of rice is more than 4 g away from the mean, then the machine is shut down.

(c) Find the probability that the machine is shut down following an inspection. **(4)**

Leave blank

5. Rosie keeps bees. The amount of honey, in kg, produced by a hive of Rosie's bees in a season, is modelled by a normal distribution with a mean of 22 kg and a standard deviation of 10 kg.

(a) Find the probability that a hive of Rosie's bees produces less than 18 kg of honey in a season.

(3)

The local bee keepers' club awards a certificate to every hive that produces more than 39 kg of honey in a season, and a medal to every hive that produces more than 50 kg in a season. Given that one of Rosie's bee hives is awarded a certificate

(b) find the probability that this hive is also awarded a medal.

(5)

Sam also keeps bees. The amount of honey, in kg, produced by a hive of Sam's bees in a season, is modelled by a normal distribution with mean μ kg and standard deviation σ kg. The probability that a hive of Sam's bees produces less than 28 kg of honey in a season is 0.8413

Only 20% of Sam's bee hives produce less than 18 kg of honey in a season.

(c) Find the value of μ and the value of σ. Give your answers to 2 decimal places.

(6)

Leave blank

7. The birth weights, W grams, of a particular breed of kitten are assumed to be normally distributed with mean 99 g and standard deviation 3.6 g

(a) Find $P(W > 92)$

(3)

(b) Find, to one decimal place, the value of k such that $P(W < k) = 3P(W > k)$

(4)

(c) Write down the name given to the value of k.

(1)

For a different breed of kitten, the birth weights are assumed to be normally distributed with mean 120 g

Given that the 20th percentile for this breed of kitten is 116 g

(d) find the standard deviation of the birth weight of this breed of kitten.

(3)

Leave blank

6. A manufacturer has a machine that fills bags with flour such that the weight of flour in a bag is normally distributed. A label states that each bag should contain 1 kg of flour.

(a) The machine is set so that the weight of flour in a bag has mean 1.04 kg and standard deviation 0.17 kg. Find the proportion of bags that weigh less than the stated weight of 1 kg.

(3)

The manufacturer wants to reduce the number of bags which contain less than the stated weight of 1 kg. At first she decides to adjust the mean but not the standard deviation so that only 5% of the bags filled are below the stated weight of 1 kg.

(b) Find the adjusted mean.

(3)

The manufacturer finds that a lot of the bags are overflowing with flour when the mean is adjusted, so decides to adjust the standard deviation instead to make the machine more accurate. The machine is set back to a mean of 1.04 kg. The manufacturer wants 1% of bags to be under 1 kg.

(c) Find the adjusted standard deviation. Give your answer to 3 significant figures.

(3)

Leave blank

5. Past records show that the proportion of customers buying organic vegetables from *Tesson* supermarket is 0.35

During a particular day, a random sample of 40 customers from *Tesson* supermarket was taken and 18 of them bought organic vegetables.

(a) Test, at the 5% level of significance, whether or not this provides evidence that the proportion of customers who bought organic vegetables has increased. State your hypotheses clearly.
(5)

The manager of *Tesson* supermarket claims that the proportion of customers buying organic eggs is different from the proportion of those buying organic vegetables. To test this claim the manager decides to take a random sample of 50 customers.

(b) Using a 5% level of significance, find the critical region to enable the *Tesson* supermarket manager to test her claim. The probability for each tail of the region should be as close as possible to 2.5%
(3)

During a particular day, a random sample of 50 customers from *Tesson* supermarket is taken and 8 of them bought organic eggs.

(c) Using your answer to part (b), state whether or not this sample supports the manager's claim. Use a 5% level of significance.
(1)

(d) State the actual significance level of this test.
(1)

The proportion of customers who buy organic fruit from *Tesson* supermarket is 0.2
During a particular day, a random sample of 200 customers from *Tesson* supermarket is taken. Using a suitable approximation, the probability that fewer than n of these customers bought organic fruit is 0.0465 correct to 4 decimal places.

(e) Find the value of n.
(6)

Leave blank

5. The time taken for a randomly selected person to complete a test is M minutes, where $M \sim \text{N}(14, \sigma^2)$

Given that 10% of people take less than 12 minutes to complete the test,

(a) find the value of σ

(3)

Graham selects 15 people at random.

(b) Find the probability that fewer than 2 of these people will take less than 12 minutes to complete the test.

(3)

Jovanna takes a random sample of n people.

Using a normal approximation, the probability that fewer than 9 of these n people will take less than 12 minutes to complete the test is 0.3085 to 4 decimal places.

(c) Find the value of n.

(8)

Leave blank

5. (a) State the conditions under which the normal distribution may be used as an approximation to the binomial distribution.
(2)

A company sells seeds and claims that 55% of its pea seeds germinate.

(b) Write down a reason why the company should not justify their claim by testing all the pea seeds they produce.
(1)

To test the company's claim, a random sample of 220 pea seeds was planted.

(c) State the hypotheses for a two-tailed test of the company's claim.
(1)

Given that 135 of the 220 pea seeds germinated,

(d) use a normal approximation to test, at the 5% level of significance, whether or not the company's claim is justified.
(7)