PMT

Paper Reference(s)
6684

Edexcel GCE

Statistics S2
(New Syllabus)
Advanced/Advanced Subsidiary
Tuesday 19 June 2001 – Morning
Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Answer Book (AB16)	Nil
Graph Paper (ASG2)	
Mathematical Formulae (Lilac)	

Candidates may use any calculator EXCEPT those with the facility for symbolic algebra, differentiation and/or integration. Thus candidates may NOT use calculators such as the Texas Instruments TI 89, TI 92, Casio CFX 9970G, Hewlett Packard HP 48G.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has seven questions. Pages 6, 7 and 8 are blank.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

N9091 83019/R6684 3/3/3/3/3/5600

1. The small village of *Tornep* has a preservation society which is campaigning for a new by-pass to be built. The society needs to measure

(i) the strength of opinion amongst the residents of *Tornep* for the scheme and

(ii) the flow of traffic through the village on weekdays.

The society wants to know whether to use a census or a sample survey for each of these measures.

(*a*) In each case suggest which they should use and specify a suitable sampling frame.
(4)

For the measurement of traffic flow through *Tornep*,

(*b*) suggest a suitable statistic and a possible statistical model for this statistic.
(2)

---

2. On a stretch of motorway accidents occur at a rate of 0.9 per month.

(*a*) Show that the probability of no accidents in the next month is 0.407, to 3 significant figures.
(1)

Find the probability of

(*b*) exactly 2 accidents in the next 6 month period,
(3)

(*c*) no accidents in exactly 2 of the next 4 months.
(3)

---

3. In a sack containing a large number of beads $\frac{1}{4}$ are coloured gold and the remainder are of different colours. A group of children use some of the beads in a craft lesson and do not replace them. Afterwards the teacher wishes to know whether or not the proportion of gold beads left in the sack has changed. He selects a random sample of 20 beads and finds that 2 of them are coloured gold.

Stating your hypotheses clearly test, at the 10% level of significance, whether or not there is evidence that the proportion of gold beads has changed.
(7)

PMT

4. A company always sends letters by second class post unless they are marked first class. Over a long period of time it has been established that 20% of letters to be posted are marked first class.

In a random selection of 10 letters to be posted, find the probability that the number marked first class is

(*a*) at least 3,
(2)

(*b*) fewer than 2.
(2)

One Monday morning there are only 12 first class stamps. Given that there are 70 letters to be posted that day,

(*c*) use a suitable approximation to find the probability that there are enough first class stamps.
(7)

(*d*) State an assumption about these 70 letters that is required in order to make the calculation in part (*c*) valid.
(1)

---

5. The maintenance department of a college receives requests for replacement light bulbs at a rate of 2 per week.

Find the probability that in a randomly chosen week the number of requests for replacement light bulbs is

(*a*) exactly 4,
(2)

(*b*) more than 5.
(2)

Three weeks before the end of term the maintenance department discovers that there are only 5 light bulbs left.

(*c*) Find the probability that the department can meet all requests for replacement light bulbs before the end of term.
(3)

The following term the principal of the college announces a package of new measures to reduce the amount of damage to college property. In the first 4 weeks following this announcement, 3 requests for replacement light bulbs are received.

(*d*) Stating your hypotheses clearly test, at the 5% level of significance, whether or not there is evidence that the rate of requests for replacement light bulbs has decreased.
(5)

---

6. The continuous random variable X has cumulative distribution function F(x) given by

$$F(x) = \begin{cases} 0, & x < 1 \\ \frac{1}{27}(-x^3 + 6x^2 - 5), & 1 \le x \le 4 \\ 1, & x > 4 \end{cases}$$

(*a*) Find the probability density function f(x).
(3)

(*b*) Find the mode of *X*.
(2)

(*c*) Sketch f(x) for all values of x.
(3)

(*d*) Find the mean μ of X.
(3)

(*e*) Show that F(μ) > 0.5.
(1)

(*f*) Show that the median of *X* lies between the mode and the mean.
(2)

PMT

7. In a computer game, a star moves across the screen, with constant speed, taking 1 s to travel from one side to the other. The player can stop the star by pressing a key. The object of the game is to stop the star in the middle of the screen by pressing the key exactly 0.5 s after the star first appears. Given that the player actually presses the key T s after the star first appears, a simple model of the game assumes that T is a continuous uniform random variable defined over the interval [0, 1].

(*a*) Write down $P(T < 0.2)$.

(1)

(*b*) Write down $E(T)$.

(1)

(*c*) Use integration to find $\mathrm{Var}(T)$.

(4)

A group of 20 children each play this game once.

(*d*) Find the probability that no more than 4 children stop the star in less than 0.2 s.

(3)

The children are allowed to practise.this game so that this continuous uniform model is no longer applicable.

(*e*) Explain how you would expect the mean and variance of T to change.

(2)

It is found that a more appropriate model of the game when played by experienced players assumes that T has a probability density function g(t) given by

$$g(t) = \begin{cases} 4t, & 0 \le t \le 0.5, \\ 4 - 4t, & 0.5 \le t \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

(*f*) Using this model show that $P(T < 0.2) = 0.08$.

(2)

A group of 75 experienced players each played this game once.

(*g*) Using a suitable approximation, find the probability that more than 7 of them stop the star in less than 0.2 s.

(4)

END

Paper Reference(s)

6684

Edexcel GCE

Statistics S2
(New Syllabus)
Advanced/Advanced Subsidiary

Wednesday 23 January 2002 – Afternoon

Time: 1 hour 30 minutes

Materials required for examination	**Items included with question papers**
Answer Book (AB16)	Nil
Graph Paper (ASG2)	
Mathematical Formulae (Lilac)	

Candidates may use any calculator EXCEPT those with the facility for symbolic algebra, differentiation and/or integration. Thus candidates may NOT use calculators such as the Texas Instruments TI 89, TI 92, Casio CFX 9970G, Hewlett Packard HP 48G.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 7 questions.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

PMT

1. Explain what you understand by

(*a*) a population,

(1)

(*b*) a statistic.

(2)

A questionnaire concerning attitudes to classes in a college was completed by a random sample of 50 students. The students gave the college a mean approval rating of 75%.

(*c*) Identify the population and the statistic in this situation.

(2)

(*d*) Explain what you understand by the sampling distribution of this statistic.

(2)

---

2. The number of houses sold per week by a firm of estate agents follows a Poisson distribution with mean 2.5. The firm appoints a new salesman and wants to find out whether or not house sales increase as a result. After the appointment of the salesman, the number of house sales in a randomly chosen 4-week period is 14.

Stating your hypotheses clearly test, at the 5% level of significance, whether or not the new salesman has increased house sales.

(7)

---

3. An airline knows that overall 3% of passengers do not turn up for flights. The airline decides to adopt a policy of selling more tickets than there are seats on a flight. For an aircraft with 196 seats, the airline sold 200 tickets for a particular flight.

(*a*) Write down a suitable model for the number of passengers who do not turn up for this flight after buying a ticket.

(2)

By using a suitable approximation, find the probability that

(*b*) more than 196 passengers turn up for this flight,

(3)

(*c*) there is at least one empty seat on this flight.

(2)

---

4. Jean catches a bus to work every morning. According to the timetable the bus is due at 8 a.m., but Jean knows that the bus can arrive at a random time between five minutes early and 9 minutes late. The random variable X represents the time, in minutes, after 7.55 a.m. when the bus arrives.

(*a*) Suggest a suitable model for the distribution of X and specify it fully.

(2)

(*b*) Calculate the mean time of arrival of the bus.

(3)

(*c*) Find the cumulative distribution function of X.

(4)

Jean will be late for work if the bus arrives after 8.05 a.m.

(*d*) Find the probability that Jean is late for work.

(2)

---

5. An Internet service provider has a large number of users regularly connecting to its computers. On average only 3 users every hour fail to connect to the Internet at their first attempt.

(*a*) Give 2 reasons why a Poisson distribution might be a suitable model for the number of failed connections every hour.

(2)

(*b*) Find the probability that in a randomly chosen hour

(i) all Internet users connect at their first attempt,

(2)

(ii) more than 4 users fail to connect at their first attempt.

(2)

(*c*) Write down the distribution of the number of users failing to connect at their first attempt in an 8-hour period.

(1)

(*d*) Using a suitable approximation, find the probability that 12 or more users fail to connect at their first attempt in a randomly chosen 8-hour period.

(6)

PMT

6. The owner of a small restaurant decides to change the menu. A trade magazine claims that 40% of all diners choose organic foods when eating away from home. On a randomly chosen day there are 20 diners eating in the restaurant.

(*a*) Assuming the claim made by the trade magazine to be correct, suggest a suitable model to describe the number of diners X who choose organic foods.
(2)

(*b*) Find $P(5 < X < 15)$.
(4)

(*c*) Find the mean and standard deviation of X.
(3)

The owner decides to survey her customers before finalising the new menu. She surveys 10 randomly chosen diners and finds 8 who prefer eating organic foods.

(*d*) Test, at the 5% level of significance, whether or not there is reason to believe that the proportion of diners in her restaurant who prefer to eat organic foods is higher than the trade magazine's claim. State your hypotheses clearly.
(5)

---

7. A continuous random variable X has cumulative distribution function $F(x)$ given by

$$F(x) = \begin{cases} 0, & x < 0, \\ kx^2 + 2kx, & 0 \le x \le 2, \\ 8k, & x > 2. \end{cases}$$

(*a*) Show that $k = \frac{1}{8}$.
(1)

(*b*) Find the median of X.
(3)

(*c*) Find the probability density function $f(x)$.
(3)

(*d*) Sketch $f(x)$ for all values of x.
(3)

(*e*) Write down the mode of X.
(1)

(*f*) Find $E(X)$.
(3)

(*g*) Comment on the skewness of this distribution.
(2)

END

Paper Reference(s)

6684

Edexcel GCE

Statistics S2

Advanced/Advanced Subsidiary

Friday 14 June 2002 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Answer Book (AB16)	Nil
Graph Paper (ASG2)	
Mathematical Formulae (Lilac)	

Candidates may use any calculator EXCEPT those with the facility for symbolic algebra, differentiation and/or integration. Thus candidates may NOT use calculators such as the Texas Instruments TI 89, TI 92, Casio CFX 9970G, Hewlett Packard HP 48G.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has seven questions.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

N10638

PMT

1. The manager of a leisure club is considering a change to the club rules. The club has a large membership and the manager wants to take the views of the members into consideration before deciding whether or not to make the change.

(*a*) Explain briefly why the manager might prefer to use a sample survey rather than a census to obtain the views.
(2)

(*b*) Suggest a suitable sampling frame.
(1)

(*c*) Identify the sampling units.
(1)

---

2. A random sample $X_1, X_2, ..., X_n$ is taken from a finite population. A statistic Y is based on this sample.

(*a*) Explain what you understand by the statistic Y.
(2)

(*b*) Give an example of a statistic.
(1)

(*c*) Explain what you understand by the sampling distribution of Y.
(2)

---

3. The continuous random variable R is uniformly distributed on the interval $\alpha \leq R \leq \beta$. Given that $E(R) = 3$ and $\text{Var}(R) = \frac{25}{3}$, find

(*a*) the value of α and the value of β,
(7)

(*b*) $P(R < 6.6)$.
(2)

---

4. Past records show that 20% of customers who buy crisps from a large supermarket buy them in single packets. During a particular day a random sample of 25 customers who had bought crisps was taken and 2 of them had bought them in single packets.

(*a*) Use these data to test, at the 5% level of significance, whether or not the percentage of customers who bought crisps in single packets that day was lower than usual. State your hypotheses clearly.
(6)

At the same supermarket, the manager thinks that the probability of a customer buying a bumper pack of crisps is 0.03. To test whether or not this hypothesis is true the manager decides to take a random sample of 300 customers.

(*b*) Stating your hypotheses clearly, find the critical region to enable the manager to test whether or not there is evidence that the probability is different from 0.03. The probability for each tail of the region should be as close as possible to 2.5%.
(6)

(*c*) Write down the significance level of this test.
(1)

---

5. A garden centre sells canes of nominal length 150 cm. The canes are bought from a supplier who uses a machine to cut canes of length L where $L \sim N(\mu, 0.3^2)$.

(*a*) Find the value of μ, to the nearest 0.1 cm, such that there is only a 5% chance that a cane supplied to the garden centre will have length less than 150 cm.
(4)

A customer buys 10 of these canes from the garden centre.

(*b*) Find the probability that at most 2 of the canes have length less than 150 cm.
(3)

Another customer buys 500 canes.

(*c*) Using a suitable approximation, find the probability that fewer than 35 of the canes will have length less than 150 cm.
(6)

PMT

6. From past records, a manufacturer of twine knows that faults occur in the twine at random and at a rate of 1.5 per 25 m.

(*a*) Find the probability that in a randomly chosen 25 m length of twine there will be exactly 4 faults.
(2)

The twine is usually sold in balls of length 100 m. A customer buys three balls of twine.

(*b*) Find the probability that only one of them will have fewer than 6 faults.
(6)

As a special order a ball of twine containing 500 m is produced.

(*c*) Using a suitable approximation, find the probability that it will contain between 23 and 33 faults inclusive.
(6)

---

7. The continuous random variable X has probability density function

$$f(x) = \begin{cases} \dfrac{x}{15}, & 0 \le x \le 2, \\ \dfrac{2}{15}, & 2 < x < 7, \\ \dfrac{4}{9} - \dfrac{2x}{45}, & 7 \le x \le 10, \\ 0, & \text{otherwise.} \end{cases}$$

(*a*) Sketch f(x) for all values of x.
(3)

(*b*) (i) Find expressions for the cumulative distribution function, F(x), for $0 \le x \le 2$ and for $7 \le x \le 10$.

(ii) Show that for $2 < x < 7$, $F(x) = \dfrac{2x}{15} - \dfrac{2}{15}$.

(iii) Specify F(x) for $x < 0$ and for $x > 10$.
(8)

(*c*) Find $P(X \le 8.2)$.
(2)

(*d*) Find, to 3 significant figures, E(X).
(4)

END

Paper Reference(s)
6684

Edexcel GCE
Statistics S2
Advanced/Advanced Subsidiary
Friday 24 January 2003 – Morning
Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Answer Book (AB16)	Nil
Graph Paper (ASG2)	
Mathematical Formulae (Lilac)	

Candidates may use any calculator EXCEPT those with the facility for symbolic algebra, differentiation and/or integration. Thus candidates may NOT use calculators such as the Texas Instruments TI 89, TI 92, Casio CFX 9970G, Hewlett Packard HP 48G.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has six questions.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

N10634A

PMT

1. An engineer measures, to the nearest cm, the lengths of metal rods.

(*a*) Suggest a suitable model to represent the difference between the true lengths and the measured lengths.
(2)

(*b*) Find the probability that for a randomly chosen rod the measured length will be within 0.2 cm of the true length.
(2)

Two rods are chosen at random.

(*c*) Find the probability that for both rods the measured lengths will be within 0.2 cm of their true lengths.
(2)

---

2. A single observation x is to be taken from a Poisson distribution with parameter λ. This observation is to be used to test $H_0 : \lambda = 7$ against $H_1 : \lambda \neq 7$.

(*a*) Using a 5% significance level, find the critical region for this test assuming that the probability of rejection in either tail is as close as possible to 2.5%.
(5)

(*b*) Write down the significance level of this test.
(1)

The actual value of x obtained was 5.

(*c*) State a conclusion that can be drawn based on this value.
(2)

---

3. A botanist suggests that the number of a particular variety of weed growing in a meadow can be modelled by a Poisson distribution.

(*a*) Write down two conditions that must apply for this model to be applicable.
(2)

Assuming this model and a mean of 0.7 weeds per m^2, find

(*b*) the probability that in a randomly chosen plot of size 4 m^2 there will be fewer than 3 of these weeds.
(4)

(*c*) Using a suitable approximation, find the probability that in a plot of 100 m^2 there will be more than 66 of these weeds.
(6)

---

4. The continuous random variable X has cumulative distribution function

$$\mathrm{F}(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{3}x^2(4 - x^2), & 0 \le x \le 1, \\ 1 & x > 1. \end{cases}$$

(*a*) Find $\mathrm{P}(X > 0.7)$.
(2)

(*b*) Find the probability density function $\mathrm{f}(x)$ of X.
(2)

(*c*) Calculate $\mathrm{E}(X)$ and show that, to 3 decimal places, $\mathrm{Var}(X) = 0.057$.
(6)

One measure of skewness is

$$\frac{\text{Mean} - \text{Mode}}{\text{Standard deviation}}.$$

(*d*) Evaluate the skewness of the distribution of X.
(4)

---

5. A farmer noticed that some of the eggs laid by his hens had double yolks. He estimated the probability of this happening to be 0.05. Eggs are packed in boxes of 12.

Find the probability that in a box, the number of eggs with double yolks will be

(*a*) exactly one,
(3)

(*b*) more than three.
(2)

A customer bought three boxes.

(*c*) Find the probability that only 2 of the boxes contained exactly 1 egg with a double yolk.
(3)

The farmer delivered 10 boxes to a local shop.

(*d*) Using a suitable approximation, find the probability that the delivery contained at least 9 eggs with double yolks.
(4)

The weight of an individual egg can be modelled by a normal distribution with mean 65 g and standard deviation 2.4 g.

(*e*) Find the probability that a randomly chosen egg weighs more than 68 g.
(3)

PMT

6. A magazine has a large number of subscribers who each pay a membership fee that is due on January 1st each year. Not all subscribers pay their fee by the due date. Based on correspondence from the subscribers, the editor of the magazine believes that 40% of subscribers wish to change the name of the magazine. Before making this change the editor decides to carry out a sample survey to obtain the opinions of the subscribers. He uses only those members who have paid their fee on time.

(*a*) Define the population associated with the magazine.
(1)

(*b*) Suggest a suitable sampling frame for the survey.
(1)

(*c*) Identify the sampling units.
(1)

(*d*) Give one advantage and one disadvantage that would have resulted from the editor using a census rather than a sample survey.
(2)

As a pilot study the editor took a random sample of 25 subscribers.

(*e*) Assuming that the editor's belief is correct, find the probability that exactly 10 of these subscribers agreed with changing the name.
(3)

In fact only 6 subscribers agreed to the name being changed.

(*f*) Stating your hypotheses clearly test, at the 5% level of significance, whether or not the percentage agreeing to the change is less that the editor believes.
(5)

The full survey is to be carried out using 200 randomly chosen subscribers.

(*g*) Again assuming the editor's belief to be correct and using a suitable approximation, find the probability that in this sample there will be least 71 but fewer than 83 subscribers who agree to the name being changed.
(7)

END

Paper Reference(s)
6684

Edexcel GCE

Statistics S2
Advanced/Advanced Subsidiary
Tuesday 17 June 2003 – Afternoon
Time: 1 hour 30 minutes

Materials required for examination	**Items included with question papers**
Answer Book (AB16)	Nil
Graph Paper (ASG2)	
Mathematical Formulae (Lilac)	

Candidates may use any calculator EXCEPT those with the facility for symbolic algebra, differentiation and/or integration. Thus candidates may NOT use calculators such as the Texas Instruments TI 89, TI 92, Casio CFX 9970G, Hewlett Packard HP 48G.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has seven questions.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

N13899A

PMT

1. Explain briefly what you understand by

(*a*) a statistic,
(2)

(*b*) a sampling distribution.
(2)

---

2. (*a*) Write down the condition needed to approximate a Poisson distribution by a Normal distribution.
(1)

The random variable $Y \sim \text{Po}(30)$.

(*b*) Estimate $P(Y > 28)$.
(6)

---

3. In a town, 30% of residents listen to the local radio station. Four residents are chosen at random.

(*a*) State the distribution of the random variable X, the number of these four residents that listen to local radio.
(2)

(*b*) On graph paper, draw the probability distribution of X.
(3)

(*c*) Write down the most likely number of these four residents that listen to the local radio station.
(1)

(*d*) Find $E(X)$ and Var (X).
(3)

---

4. (*a*) Write down the conditions under which the binomial distribution may be a suitable model to use in statistical work.
(4)

A six-sided die is biased. When the die is thrown the number 5 is twice as likely to appear as any other number. All the other faces are equally likely to appear. The die is thrown repeatedly.

Find the probability that

(*b*) (i) the first 5 will occur on the sixth throw,

(ii) in the first eight throws there will be exactly three 5s.
(8)

---

5. A drinks machine dispenses lemonade into cups. It is electronically controlled to cut off the flow of lemonade randomly between 180 ml and 200 ml. The random variable X is the volume of lemonade dispensed into a cup.

(*a*) Specify the probability density function of X and sketch its graph.
(4)

(*b*) Find the probability that the machine dispenses

(i) less than 183 ml,

(ii) exactly 183 ml.
(3)

(*c*) Calculate the inter-quartile range of X.
(1)

(*d*) Determine the value of x such that $P(X \geq x) = 2P(X \leq x)$.
(3)

(*e*) Interpret in words your value of x.
(2)

---

6. A doctor expects to see, on average, 1 patient per week with a particular disease.

(*a*) Suggest a suitable model for the distribution of the number of times per week that the doctor sees a patient with the disease. Give a reason for your answer.
(3)

(*b*) Using your model, find the probability that the doctor sees more than 3 patients with the disease in a 4 week period.
(4)

The doctor decides to send information to his patients to try to reduce the number of patients he sees with the disease. In the first 6 weeks after the information is sent out, the doctor sees 2 patients with the disease.

(*c*) Test, at the 5% level of significance, whether or not there is reason to believe that sending the information has reduced the number of times the doctor sees patients with the disease. State your hypotheses clearly.
(6)

Medical research into the nature of the disease discovers that it can be passed from one patient to another.

(*d*) Explain whether or not this research supports your choice of model. Give a reason for your answer.
(2)

PMT

7. A continuous random variable X has probability density function f(x) where

$$f(x) = \begin{cases} k(x^2 + 2x + 1) & -1 \le x \le 0, \\ 0, & \text{otherwise} \end{cases}$$

where k is a positive integer.

(a) Show that $k = 3$.

(4)

Find

(b) E(X),

(4)

(c) the cumulative distribution function F(x),

(4)

(d) P($-0.3 < X < 0.3$).

(3)

END

Paper Reference(s)

6684

Edexcel GCE

Statistics S2

Advanced/Advanced Subsidiary

Friday 23 January 2004 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Answer Book (AB16)	Nil
Graph Paper (ASG2)	
Mathematical Formulae (Lilac)	

Candidates may use any calculator EXCEPT those with the facility for symbolic algebra, differentiation and/or integration. Thus candidates may NOT use calculators such as the Texas Instruments TI 89, TI 92, Casio CFX 9970G, Hewlett Packard HP 48G.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet ‘Mathematical Formulae and Statistical Tables’ is provided.
Full marks may be obtained for answers to ALL questions.
This paper has seven questions.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

N13319A

PMT

1. A large dental practice wishes to investigate the level of satisfaction of its patients.

(*a*) Suggest a suitable sampling frame for the investigation. **(1)**

(*b*) Identify the sampling units. **(1)**

(*c*) State one advantage and one disadvantage of using a sample survey rather than a census. **(2)**

(*d*) Suggest a problem that might arise with the sampling frame when selecting patients. **(1)**

---

2. The random variable R has the binomial distribution B(12, 0.35).

(*a*) Find $P(R \geq 4)$. **(2)**

The random variable S has the Poisson distribution with mean 2.71.

(*b*) Find $P(S \leq 1)$. **(3)**

The random variable T has the normal distribution $N(25, 5^2)$.

(*c*) Find $P(T \leq 18)$. **(2)**

---

3. The discrete random variable X is distributed $B(n, p)$.

(*a*) Write down the value of p that will give the most accurate estimate when approximating the binomial distribution by a normal distribution. **(1)**

(*b*) Give a reason to support your value. **(1)**

(*c*) Given that $n = 200$ and $p = 0.48$, find $P(90 \leq X < 105)$. **(7)**

---

4. (*a*) Write down two conditions needed to be able to approximate the binomial distribution by the Poisson distribution. **(2)**

A researcher has suggested that 1 in 150 people is likely to catch a particular virus.

Assuming that a person catching the virus is independent of any other person catching it,

(*b*) find the probability that in a random sample of 12 people, exactly 2 of them catch the virus. **(4)**

(*c*) Estimate the probability that in a random sample of 1200 people fewer than 7 catch the virus. **(4)**

---

5. Vehicles pass a particular point on a road at a rate of 51 vehicles per hour.

(*a*) Give two reasons to support the use of the Poisson distribution as a suitable model for the number of vehicles passing this point. **(2)**

Find the probability that in any randomly selected 10 minute interval

(*b*) exactly 6 cars pass this point, **(3)**

(*c*) at least 9 cars pass this point. **(2)**

After the introduction of a roundabout some distance away from this point it is suggested that the number of vehicles passing it has decreased. During a randomly selected 10 minute interval 4 vehicles pass the point.

(*d*) Test, at the 5% level of significance, whether or not there is evidence to support the suggestion that the number of vehicles has decreased. State your hypotheses clearly. **(6)**

PMT

6. From past records a manufacturer of ceramic plant pots knows that 20% of them will have defects. To monitor the production process, a random sample of 25 pots is checked each day and the number of pots with defects is recorded.

(*a*) Find the critical regions for a two-tailed test of the hypothesis that the probability that a plant pot has defects is 0.20. The probability of rejection in either tail should be as close as possible to 2.5%.
(5)

(*b*) Write down the significance level of the above test.
(1)

A garden centre sells these plant pots at a rate of 10 per week. In an attempt to increase sales, the price was reduced over a six-week period. During this period a total of 74 pots was sold.

(*c*) Using a 5% level of significance, test whether or not there is evidence that the rate of sales per week has increased during this six-week period.
(7)

---

7. The continuous random variable X has probability density function

$$f(x) = \begin{cases} kx(5-x), & 0 \le x \le 4, \\ 0, & \text{otherwise,} \end{cases}$$

where k is a constant.

(*a*) Show that $k = \dfrac{3}{56}$.
(3)

(*b*) Find the cumulative distribution function F(x) for all values of x.
(4)

(*c*) Evaluate E(X).
(3)

(*d*) Find the modal value of X.
(3)

(*e*) Verify that the median value of X lies between 2.3 and 2.5.
(3)

(*f*) Comment on the skewness of X. Justify your answer.
(2)

END

Paper Reference(s)
6684

Edexcel GCE
Statistics S2
Advanced/Advanced Subsidiary
Wednesday 23 June 2004 – Morning
Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Answer Book (AB16)	Nil
Graph Paper (ASG2)	
Mathematical Formulae (Lilac)	

Candidates may use any calculator EXCEPT those with the facility for symbolic algebra, differentiation and/or integration. Thus candidates may NOT use calculators such as the Texas Instruments TI 89, TI 92, Casio CFX 9970G, Hewlett Packard HP 48G.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has seven questions.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

N17023A

PMT

1. Explain briefly what you understand by

(*a*) a sampling frame,
(1)

(*b*) a statistic.
(2)

---

2. The continuous random variable X is uniformly distributed over the interval $[-1, 4]$.

Find

(*a*) $P(X < 2.7)$,
(1)

(*b*) $E(X)$,
(2)

(*c*) $Var(X)$.
(2)

---

3. Brad planted 25 seeds in his greenhouse. He has read in a gardening book that the probability of one of these seeds germinating is 0.25. Ten of Brad's seeds germinated. He claimed that the gardening book had underestimated this probability. Test, at the 5% level of significance, Brad's claim. State your hypotheses clearly.
(7)

---

4. (*a*) State two conditions under which a random variable can be modelled by a binomial distribution.
(2)

In the production of a certain electronic component it is found that 10% are defective.

The component is produced in batches of 20.

(*b*) Write down a suitable model for the distribution of defective components in a batch.
(1)

Find the probability that a batch contains

(*c*) no defective components,
(2)

(*d*) more than 6 defective components.
(2)

(*e*) Find the mean and the variance of the defective components in a batch.
(2)

A supplier buys 100 components. The supplier will receive a refund if there are more than 15 defective components.

(*f*) Using a suitable approximation, find the probability that the supplier will receive a refund.
(4)

---

5. (*a*) Explain what you understand by a critical region of a test statistic.
(2)

The number of breakdowns per day in a large fleet of hire cars has a Poisson distribution with mean $\frac{1}{7}$.

(*b*) Find the probability that on a particular day there are fewer than 2 breakdowns.
(3)

(*c*) Find the probability that during a 14-day period there are at most 4 breakdowns.
(3)

The cars are maintained at a garage. The garage introduced a weekly check to try to decrease the number of cars that break down. In a randomly selected 28-day period after the checks are introduced, only 1 hire car broke down.

(*d*) Test, at the 5% level of significance, whether or not the mean number of breakdowns has decreased. State your hypotheses clearly.
(7)

PMT

6. Minor defects occur in a particular make of carpet at a mean rate of 0.05 per m^2.

(*a*) Suggest a suitable model for the distribution of the number of defects in this make of carpet. Give a reason for your answer.

A carpet fitter has a contract to fit this carpet in a small hotel. The hotel foyer requires 30 m^2 of this carpet. Find the probability that the foyer carpet contains

(*b*) exactly 2 defects,
(3)

(*c*) more than 5 defects.
(3)

The carpet fitter orders a total of 355 m^2 of the carpet for the whole hotel.

(*d*) Using a suitable approximation, find the probability that this total area of carpet contains 22 or more defects.
(6)

---

7. A random variable X has probability density function given by

$$\mathrm{f}(x) = \begin{cases} \frac{1}{3}, & 0 \le x \le 1, \\ \frac{8x^3}{45}, & 1 \le x \le 2, \\ 0, & \text{otherwise.} \end{cases}$$

(*a*) Calculate the mean of X.
(5)

(*b*) Specify fully the cumulative distribution function $\mathrm{F}(x)$.
(7)

(*c*) Find the median of X.
(3)

(*d*) Comment on the skewness of the distribution of X.
(2)

END

Paper Reference(s)

6684

Edexcel GCE

Statistics S2

Advanced/Advanced Subsidiary

Tuesday 25 January 2005 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Answer Book (AB16)	Nil
Graph Paper (ASG2)	
Mathematical Formulae	

Candidates may use any calculator EXCEPT those with the facility for symbolic algebra, differentiation and/or integration. Thus candidates may NOT use calculators such as the Texas Instruments TI 89, TI 92, Casio CFX 9970G, Hewlett Packard HP 48G.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has seven questions.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

N16742A

PMT

1. The random variables R, S and T are distributed as follows

$$R \sim \mathrm{B}(15, 0.3), \quad S \sim \mathrm{Po}(7.5), \quad T \sim \mathrm{N}(8, 2^2).$$

Find

(*a*) $\mathrm{P}(R = 5)$,
(2)

(*b*) $\mathrm{P}(S = 5)$,
(1)

(*c*) $\mathrm{P}(T = 5)$.
(1)

---

2. (*a*) Explain what you understand by (i) a population and (ii) a sampling frame.
(2)

The population and the sampling frame may not be the same.

(*b*) Explain why this might be the case.
(1)

(*c*) Give an example, justifying your choices, to illustrate when you might use

(i) a census,

(ii) a sample.
(4)

---

3. A rod of length $2l$ was broken into 2 parts. The point at which the rod broke is equally likely to be anywhere along the rod. The length of the **shorter** piece of rod is represented by the random variable X.

(*a*) Write down the name of the probability density function of X, and specify it fully.
(3)

(*b*) Find $\mathrm{P}(X < \frac{1}{3}l)$.
(2)

(*c*) Write down the value of $\mathrm{E}(X)$.
(1)

Two identical rods of length $2l$ are broken.

(*d*) Find the probability that both of the shorter pieces are of length less than $\frac{1}{3}l$.
(2)

---

4. In an experiment, there are 250 trials and each trial results in a success or a failure.

(*a*) Write down two other conditions needed to make this into a binomial experiment.
(2)

It is claimed that 10% of students can tell the difference between two brands of baked beans. In a random sample of 250 students, 40 of them were able to distinguish the difference between the two brands.

(*b*) Using a normal approximation, test at the 1% level of significance whether or not the claim is justified. Use a one-tailed test.
(6)

(*c*) Comment on the acceptability of the assumptions you needed to carry out the test.
(2)

---

5. From company records, a manager knows that the probability that a defective article is produced by a particular production line is 0.032.

A random sample of 10 articles is selected from the production line.

(*a*) Find the probability that exactly 2 of them are defective.
(3)

On another occasion, a random sample of 100 articles is taken.

(*b*) Using a suitable approximation, find the probability that fewer than 4 of them are defective.
(4)

At a later date, a random sample of 1000 is taken.

(*c*) Using a suitable approximation, find the probability that more than 42 are defective.
(6)

PMT

6. Over a long period of time, accidents happened on a stretch of road at random at a rate of 3 per month.

Find the probability that

(*a*) in a randomly chosen month, more than 4 accidents occurred,
(3)

(*b*) in a three-month period, more than 4 accidents occurred.
(2)

At a later date, a speed restriction was introduced on this stretch of road. During a randomly chosen month only one accident occurred.

(*c*) Test, at the 5% level of significance, whether or not there is evidence to support the claim that this speed restriction reduced the mean number of road accidents occurring per month.
(4)

The speed restriction was kept on this road. Over a two-year period, 55 accidents occurred.

(*d*) Test, at the 5% level of significance, whether or not there is now evidence that this speed restriction reduced the mean number of road accidents occurring per month.
(7)

---

7. The random variable X has probability density function

$$f(x) = \begin{cases} k(-x^2 + 5x - 4), & 1 \leq x \leq 4, \\ 0, & \text{otherwise.} \end{cases}$$

(*a*) Show that $k = \frac{2}{9}$.
(3)

Find

(*b*) $E(X)$,
(3)

(*c*) the mode of X.
(2)

(*d*) the cumulative distribution function $F(x)$ for all x.
(5)

(*e*) Evaluate $P(X \leq 2.5)$,
(2)

(*f*) Deduce the value of the median and comment on the shape of the distribution.
(2)

END

Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced/Advanced Subsidiary

Wednesday 22 June 2005 – Afternoon

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Lilac or Green)	Nil

Candidates may use any calculator EXCEPT those with the facility for symbolic algebra, differentiation and/or integration. Thus candidates may NOT use calculators such as the Texas Instruments TI 89, TI 92, Casio CFX 9970G, Hewlett Packard HP 48G.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 7 questions.
The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

1. It is estimated that 4% of people have green eyes. In a random sample of size n, the expected number of people with green eyes is 5.

(*a*) Calculate the value of n. **(3)**

The expected number of people with green eyes in a second random sample is 3.

(*b*) Find the standard deviation of the number of people with green eyes in this second sample. **(4)**

---

2. The continuous random variable X is uniformly distributed over the interval [2, 6].

(*a*) Write down the probability density function f(x). **(2)**

Find

(*b*) E(X), **(1)**

(*c*) Var(X), **(2)**

(*d*) the cumulative distribution function of X, for all x, **(4)**

(*e*) P($2.3 < X < 3.4$). **(2)**

---

3. The random variable X is the number of misprints per page in the first draft of a novel.

(*a*) State two conditions under which a Poisson distribution is a suitable model for X. **(2)**

The number of misprints per page has a Poisson distribution with mean 2.5. Find the probability that

(*b*) a randomly chosen page has no misprints, **(2)**

(*c*) the total number of misprints on 2 randomly chosen pages is more than 7. **(3)**

The first chapter contains 20 pages.

(*d*) Using a suitable approximation find, to 2 decimal places, the probability that the chapter will contain less than 40 misprints. **(7)**

---

4. Explain what you understand by

(*a*) a sampling unit, **(1)**

(*b*) a sampling frame, **(1)**

(*c*) a sampling distribution. **(2)**

---

5. In a manufacturing process, 2% of the articles produced are defective. A batch of 200 articles is selected.

(*a*) Giving a justification for your choice, use a suitable approximation to estimate the probability that there are exactly 5 defective articles. **(5)**

(*b*) Estimate the probability that there are less than 5 defective articles. **(2)**

---

6. A continuous random variable X has probability density function f(x) where

$$\mathrm{f}(x) = \begin{cases} k(4x - x^3), & 0 \le x \le 2, \\ 0, & \text{otherwise,} \end{cases}$$

where k is a positive constant.

(*a*) Show that $k = \frac{1}{4}$. **(4)**

Find

(*b*) E(X), **(3)**

(*c*) the mode of X, **(3)**

(*d*) the median of X. **(4)**

(*e*) Comment on the skewness of the distribution. **(2)**

(*f*) Sketch f(x). **(2)**

---

7. A drugs company claims that 75% of patients suffering from depression recover when treated with a new drug.

A random sample of 10 patients with depression is taken from a doctor's records.

(*a*) Write down a suitable distribution to model the number of patients in this sample who recover when treated with the new drug.
(2)

Given that the claim is correct,

(*b*) find the probability that the treatment will be successful for exactly 6 patients.
(2)

The doctor believes that the claim is incorrect and the percentage who will recover is lower. From her records she took a random sample of 20 patients who had been treated with the new drug. She found that 13 had recovered.

(*c*) Stating your hypotheses clearly, test, at the 5% level of significance, the doctor's belief.
(6)

(*d*) From a sample of size 20, find the greatest number of patients who need to recover from the test in part (*c*) to be significant at the 1% level.
(4)

TOTAL FOR PAPER: 75 MARKS

END

Paper Reference(s)
6684/01

Edexcel GCE

Statistics S2

Advanced/Advanced Subsidiary

Monday 16 January 2006 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Lilac)	Nil

Candidates may use any calculator EXCEPT those with the facility for symbolic algebra, differentiation and/or integration. Thus candidates may NOT use calculators such as the Texas Instruments TI 89, TI 92, Casio CFX 9970G, Hewlett Packard HP 48G.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
There are 7 questions on this paper. The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

N20909A

1. A fair coin is tossed 4 times.

Find the probability that

(*a*) an equal number of head and tails occur **(2)**

(*b*) all the outcomes are the same, **(3)**

(*c*) the first tail occurs on the third throw. **(2)**

---

2. Accidents on a particular stretch of motorway occur at an average rate of 1.5 per week.

(*a*) Write down a suitable model to represent the number of accidents per week on this stretch of motorway. **(1)**

Find the probability that

(*b*) there will be 2 accidents in the same week, **(2)**

(*c*) there is at least one accident per week for 3 consecutive weeks, **(3)**

(*d*) there are more than 4 accidents in a 2 week period. **(2)**

---

3. The random variable X is uniformly distributed over the interval $[-1, 5]$.

(*a*) Sketch the probability density function f(x) of X. **(3)**

Find

(*b*) E(X), **(1)**

(*c*) Var(X), **(2)**

(*d*) $P(-0.3 < X < 3.3)$. **(2)**

---

4. The random variable $X \sim B(150, 0.02)$.

Use a suitable approximation to estimate $P(X > 7)$. **(4)**

---

5. A continuous random variable X has probability density function f(x) where

$$f(x) = \begin{cases} kx(x-2), & 2 \le x \le 3, \\ 0, & \text{otherwise,} \end{cases}$$

where k is a positive constant.

(*a*) Show that $k = \frac{3}{4}$. **(4)**

Find

(*b*) E(X), **(3)**

(c) the cumulative distribution function F(x). **(6)**

(d) Show that the median value of X lies between 2.70 and 2.75. **(2)**

---

6. A bag contains a large number of coins. Half of them are 1p coins, one third are 2p coins and the remainder are 5p coins.

(*a*) Find the mean and variance of the value of the coins. **(4)**

A random sample of 2 coins is chosen from the bag.

(*b*) List all the possible samples that can be drawn. **(3)**

(*c*) Find the sampling distribution of the mean value of these samples. **(6)**

PMT

7. A teacher thinks that 20% of the pupils in a school read the Deano comic regularly.

He chooses 20 pupils at random and finds 9 of them read the Deano.

(*a*) (i) Test, at the 5% level of significance, whether or not there is evidence that the percentage of pupils that read the Deano is different from 20%. State your hypotheses clearly.

(ii) State all the possible numbers of pupils that read the Deano from a sample of size 20 that will make the test in part (*a*)(i) significant at the 5% level.
(9)

The teacher takes another 4 random samples of size 20 and they contain 1, 3, 1 and 4 pupils that read the Deano.

(*b*) By combining all 5 samples and using a suitable approximation test, at the 5% level of significance, whether or not this provides evidence that the percentage of pupils in the school that read the Deano is different from 20%.
(8)

(*c*) Comment on your results for the tests in part (*a*) and part (*b*).
(2)

TOTAL FOR PAPER: 75 MARKS

END

Paper Reference(s)
6684/01

Edexcel GCE

Statistics S2

Advanced/Advanced Subsidiary

Wednesday 8 June 2006 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Lilac or Green)	Nil

Candidates may use any calculator EXCEPT those with the facility for symbolic algebra, differentiation and/or integration. Thus candidates may NOT use calculators such as the Texas Instruments TI 89, TI 92, Casio CFX 9970G, Hewlett Packard HP 48G.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet ‘Mathematical Formulae and Statistical Tables’ is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 7 questions.
The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

N22338A

PMT

1. Before introducing a new rule, the secretary of a golf club decided to find out how members might react to this rule.

(*a*) Explain why the secretary decided to take a random sample of club members rather than ask all the members.
(1)

(*b*) Suggest a suitable sampling frame.
(1)

(*c*) Identify the sampling units.
(1)

---

2. The continuous random variable L represents the error, in mm, made when a machine cuts rods to a target length. The distribution of L is continuous uniform over the interval [–4.0, 4.0].

Find

(*a*) $\text{P}(L < -2.6)$,
(1)

(*b*) $\text{P}(L < -3.0 \text{ or } L > 3.0)$.
(2)

A random sample of 20 rods cut by the machine was checked.

(*c*) Find the probability that more than half of them were within 3.0 mm of the target length.
(4)

---

3. An estate agent sells properties at a mean rate of 7 per week.

(*a*) Suggest a suitable model to represent the number of properties sold in a randomly chosen week. Give two reasons to support your model.
(3)

(*b*) Find the probability that in any randomly chosen week the estate agent sells exactly 5 properties.
(2)

(*c*) Using a suitable approximation find the probability that during a 24 week period the estate agent sells more than 181 properties.
(6)

---

4. Breakdowns occur on a particular machine at random at a mean rate of 1.25 per week.

(*a*) Find the probability that fewer than 3 breakdowns occurred in a randomly chosen week.
(4)

Over a 4 week period the machine was monitored. During this time there were 11 breakdowns.

(*b*) Test, at the 5% level of significance, whether or not there is evidence that the rate of breakdowns has changed over this period. State your hypotheses clearly.
(7)

---

5. A manufacturer produces large quantities of coloured mugs. It is known from previous records that 6% of the production will be green.

A random sample of 10 mugs was taken from the production line.

(*a*) Define a suitable distribution to model the number of green mugs in this sample.
(1)

(*b*) Find the probability that there were exactly 3 green mugs in the sample.
(3)

A random sample of 125 mugs was taken.

(*c*) Find the probability that there were between 10 and 13 (inclusive) green mugs in this sample, using

(i) a Poisson approximation,
(3)

(ii) a Normal approximation.
(6)

---

6. The continuous random variable X has probability density function

$$f(x) = \begin{cases} \dfrac{1+x}{k}, & 1 \le x \le 4, \\ 0, & \text{otherwise.} \end{cases}$$

(*a*) Show that $k = \dfrac{21}{2}$.
(3)

(*b*) Specify fully the cumulative distribution function of X.
(5)

(*c*) Calculate E(X).
(3)

(*d*) Find the value of the median.
(3)

(*e*) Write down the mode.
(1)

(*f*) Explain why the distribution is negatively skewed.
(1)

---

7. It is known from past records that 1 in 5 bowls produced in a pottery have minor defects. To monitor production a random sample of 25 bowls was taken and the number of such bowls with defects was recorded.

(*a*) Using a 5% level of significance, find critical regions for a two-tailed test of the hypothesis that 1 in 5 bowls have defects. The probability of rejecting, in either tail, should be as close to 2.5% as possible.
(6)

(*b*) State the actual significance level of the above test.
(1)

At a later date, a random sample of 20 bowls was taken and 2 of them were found to have defects.

(*c*) Test, at the 10% level of significance, whether or not there is evidence that the proportion of bowls with defects has decreased. State your hypotheses clearly.
(7)

TOTAL FOR PAPER: 75 MARKS

END

Paper Reference(s)
6684/01

Edexcel GCE

Statistics S2

Advanced/Advanced Subsidiary

Tuesday 16 January 2007 – Morning

Time: 1 hour 30 minutes

Materials required for examination	**Items included with question papers**
Mathematical Formulae (Lilac)	Nil

Candidates may use any calculator EXCEPT those with the facility for symbolic algebra, differentiation and/or integration. Thus candidates may NOT use calculators such as the Texas Instruments TI 89, TI 92, Casio CFX 9970G, Hewlett Packard HP 48G.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
There are 7 questions on this paper. The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

N23556A

1. (*a*) Define a statistic. **(2)**

A random sample $X_1, X_2, ..., X_n$ is taken from a population with unknown mean μ.

(*b*) For each of the following state whether or not it is a statistic.

(i) $\frac{X_1 + X_4}{2}$, **(1)**

(ii) $\frac{\sum X^2}{n} - \mu^2$. **(1)**

---

2. The random variable *J* has a Poisson distribution with mean 4.

(*a*) Find $P(J \geq 10)$. **(2)**

The random variable *K* has a binomial distribution with parameters $n = 25$, $p = 0.27$.

(*b*) Find $P(K \leq 1)$. **(3)**

---

3. For a particular type of plant 45% have white flowers and the remainder have coloured flowers. Gardenmania sells plants in batches of 12. A batch is selected at random.

Calculate the probability this batch contains

(*a*) exactly 5 plants with white flowers, **(3)**

(*b*) more plants with white flowers than coloured ones. **(2)**

Gardenmania takes a random sample of 10 batched of plants.

(*c*) Find the probability that exactly 3 of these batches contain more plants with white flowers than coloured ones. **(3)**

Due to an increasing demand for these plants by large companies, Gardenmania decides to sell them in batches of 50.

(*d*) Use a suitable approximation to calculate the probability that a batch of 50 plants contains more than 25 plants with white flowers. **(7)**

---

4. (*a*) State the condition under which the normal distribution may be used as an approximation to the Poisson distribution. **(1)**

(*b*) Explain why a continuity correction must be incorporated when using the normal distribution as an approximation to the Poisson distribution. **(1)**

A company has yachts that can only be hired for a week at a time. All hiring starts on a Saturday. During the winter the mean number of yachts hired per week is 5.

(*c*) Calculate the probability that fewer than 3 yachts are hired on a particular Saturday in winter. **(2)**

During the summer the mean number of yachts hired per week increases to 25. The company has only 30 yachts for hire.

(*d*) Using a suitable approximation find the probability that the demand for yachts cannot be met on a particular Saturday in summer. **(6)**

In the summer there are 16 Saturdays on which a yacht can be hired.

(*e*) Estimate the number of Saturdays in the summer that the company will not be able to meet the demand for yachts. **(2)**

---

5. The continuous random variable *X* is uniformly distributed over the interval $\alpha < x < \beta$.

(*a*) Write down the probability density function of *X*, for all *x*. **(2)**

(*b*) Given that $E(X) = 2$ and $P(X < 3) = \frac{5}{8}$, find the value of α and the value of β. **(4)**

A gardener has wire cutters and a piece of wire 150 cm long which has a ring attached at one end. The gardener cuts the wire, at a randomly chosen point, into 2 pieces. The length, in cm, of the piece of wire with the ring on it is represented by the random variable *X*. Find

(*c*) E(*X*), **(1)**

(*d*) the standard deviation of *X*, **(2)**

(*e*) the probability that the shorter piece of wire is at most 30 cm long. **(3)**

PMT

6. Past records from a large supermarket show that 20% of people who buy chocolate bars buy the family size bar. On one particular day a random sample of 30 people was taken from those that had bought chocolate bars and 2 of them were found to have bought a family size bar.

(*a*) Test, at the 5% significance level, whether or not the proportion p of people who bought a family size bar of chocolate that day had decreased. State your hypotheses clearly.
(6)

The manager of the supermarket thinks that the probability of a person buying a gigantic chocolate bar is only 0.02. To test whether this hypothesis is true the manager decides to take a random sample of 200 people who bought chocolate bars.

(*b*) Find the critical region that would enable the manager to test whether or not there is evidence that the probability is different from 0.02. The probability of each tail should be as close to 2.5% as possible.
(6)

(*c*) Write down the significance level of this test.
(1)

---

7. The continuous random variable X has cumulative distribution function

$$\mathrm{F}(x) = \begin{cases} 0, & x < 0, \\ 2x^2 - x^3, & 0 \le x \le 1, \\ 1, & x > 1. \end{cases}$$

(*a*) Find $\mathrm{P}(X > 0.3)$.
(2)

(*b*) Verify that the median value of X lies between $x = 0.59$ and $x = 0.60$.
(3)

(*c*) Find the probability density function $\mathrm{f}(x)$.
(2)

(*d*) Evaluate $\mathrm{E}(X)$.
(3)

(*e*) Find the mode of X.
(2)

(*f*) Comment on the skewness of X. Justify your answer.
(2)

TOTAL FOR PAPER: 75 MARKS

END

Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced

Tuesday 15 January 2008 – Morning

Time: 1 hour 30 minutes

Materials required for examination	**Items included with question papers**
Mathematical Formulae (Green)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
There are 8 questions on this paper. The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

N29278A

1. (*a*) Explain what you understand by a census.

(1)

Each cooker produced at GT Engineering is stamped with a unique serial number. GT Engineering produces cookers in batches of 2000. Before selling them, they test a random sample of 5 to see what electric current overload they will take before breaking down.

(*b*) Give one reason, other than to save time and cost, why a sample is taken rather than a census.

(1)

(*c*) Suggest a suitable sampling frame from which to obtain this sample.

(1)

(*d*) Identify the sampling units.

(1)

---

2. The probability of a bolt being faulty is 0.3. Find the probability that in a random sample of 20 bolts there are

(*a*) exactly 2 faulty bolts,

(2)

(*b*) more than 3 faulty bolts.

(2)

These bolts are sold in bags of 20. John buys 10 bags.

(*c*) Find the probability that exactly 6 of these bags contain more than 3 faulty bolts.

(3)

---

3. (*a*) State two conditions under which a Poisson distribution is a suitable model to use in statistical work.

(2)

The number of cars passing an observation point in a 10 minute interval is modelled by a Poisson distribution with mean 1.

(*b*) Find the probability that in a randomly chosen 60 minute period there will be

(i) exactly 4 cars passing the observation point,

(ii) at least 5 cars passing the observation point.

(5)

The number of other vehicles, other than cars, passing the observation point in a 60 minute interval is modelled by a Poisson distribution with mean 12.

(*c*) Find the probability that exactly 1 vehicle, of any type, passes the observation point in a 10 minute period.

(4)

---

4. The continuous random variable Y has cumulative distribution function $\mathrm{F}(y)$ given by

$$\mathrm{F}(y) = \begin{cases} 0 & y < 1 \\ k(y^4 + y^2 - 2) & 1 \le y \le 2 \\ 1 & y > 2 \end{cases}$$

(*a*) Show that $k = \frac{1}{18}$.

(2)

(*b*) Find $\mathrm{P}(Y > 1.5)$.

(2)

(*c*) Specify fully the probability density function $\mathrm{f}(y)$.

(3)

---

5. Dhriti grows tomatoes. Over a period of time, she has found that there is a probability 0.3 of a ripe tomato having a diameter greater than 4 cm. She decides to try a new fertiliser. In a random sample of 40 ripe tomatoes, 18 have a diameter greater than 4 cm. Dhriti claims that the new fertiliser has increased the probability of a ripe tomato being greater than 4 cm in diameter.

Test Dhriti's claim at the 5% level of significance. State your hypotheses clearly.

(7)

---

6. The probability that a sunflower plant grows over 1.5 metres high is 0.25. A random sample of 40 sunflower plants is taken and each sunflower plant is measured and its height recorded.

(*a*) Find the probability that the number of sunflower plants over 1.5 m high is between 8 and 13 (inclusive) using

(i) a Poisson approximation,

(ii) a Normal approximation.

(10)

(*b*) Write down which of the approximations used in part (*a*) is the most accurate estimate of the probability. You must give a reason for your answer.

(2)

PMT

7. (*a*) Explain what you understand by

(i) a hypothesis test,

(ii) a critical region.

(3)

During term time, incoming calls to a school are thought to occur at a rate of 0.45 per minute. To test this, the number of calls during a random 20 minute interval, is recorded.

(*b*) Find the critical region for a two-tailed test of the hypothesis that the number of incoming calls occurs at a rate of 0.45 per 1 minute interval. The probability in each tail should be as close to 2.5% as possible.

(5)

(*c*) Write down the actual significance level of the above test.

(1)

In the school holidays, 1 call occurs in a 10 minute interval.

(*d*) Test, at the 5% level of significance, whether or not there is evidence that the rate of incoming calls is less during the school holidays than in term time.

(5)

---

8. The continuous random variable X has probability density function $\mathrm{f}(x)$ given by

$$\mathrm{f}(x) = \begin{cases} 2(x-2) & 2 \le x \le 3 \\ 0 & \text{otherwise} \end{cases}$$

(*a*) Sketch $\mathrm{f}(x)$ for all values of x.

(3)

(*b*) Write down the mode of X.

(1)

Find

(*c*) $\mathrm{E}(X)$,

(3)

(*d*) the median of X.

(4)

(*e*) Comment on the skewness of this distribution. Give a reason for your answer.

(2)

TOTAL FOR PAPER: 75 MARKS

END

Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced Level

Friday 23 May 2008 – Morning

Time: 1 hour 30 minutes

Materials required for examination	**Items included with question papers**
Mathematical Formulae (Lilac or Green)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulas stored in them.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 7 questions.
The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner.
Answers without working may gain no credit.

H32581A

PMT

1. Jean regularly takes a break from work to go to the post office. The amount of time Jean waits in the queue to be served at the post office has a continuous uniform distribution between 0 and 10 minutes.

(*a*) Find the mean and variance of the time Jean spends in the post office queue.
(3)

(*b*) Find the probability that Jean does not have to wait more than 2 minutes.
(2)

Jean visits the post office 5 times.

(*c*) Find the probability that she never has to wait more than 2 minutes.
(2)

Jean is in the queue when she receives a message that she must return to work for an urgent meeting. She can only wait in the queue for a further 3 minutes.

Given that Jean has already been queuing for 5 minutes,

(*d*) find the probability that she must leave the post office queue without being served.
(3)

---

2. In a large college 58% of students are female and 42% are male. A random sample of 100 students is chosen from the college. Using a suitable approximation find the probability that more than half the sample are female.
(7)

---

3. A test statistic has a Poisson distribution with parameter λ.

Given that

$$\mathrm{H_0}: \lambda = 9, \ \mathrm{H_1}: \lambda \neq 9,$$

(*a*) find the critical region for the test statistic such that the probability in each tail is as close as possible to 2.5%.
(3)

(*b*) State the probability of incorrectly rejecting $\mathrm{H_0}$ using this critical region.
(2)

---

4. Each cell of a certain animal contains 11 000 genes. It is known that each gene has a probability 0.0 005 of being damaged.

A cell is chosen at random.

(*a*) Suggest a suitable model for the distribution of the number of damaged genes in the cell.
(2)

(*b*) Find the mean and variance of the number of damaged genes in the cell.
(2)

(*c*) Using a suitable approximation, find the probability that there are at most 2 damaged genes in the cell.
(4)

---

5. Sue throws a fair coin 15 times and records the number of times it shows a head.

(*a*) State the distribution to model the number of times the coin shows a head.
(2)

Find the probability that Sue records

(*b*) exactly 8 heads,
(2)

(*c*) at least 4 heads.
(2)

Sue has a different coin which she believes is biased in favour of heads. She throws the coin 15 times and obtains 13 heads.

(*d*) Test Sue's belief at the 1% level of significance. State your hypotheses clearly.
(6)

PMT

6. A call centre agent handles telephone calls at a rate of 18 per hour.

(*a*) Give two reasons to support the use of a Poisson distribution as a suitable model for the number of calls per hour handled by the agent.
(2)

(*b*) Find the probability that in any randomly selected 15 minute interval the agent handles

(i) exactly 5 calls,

(ii) more than 8 calls.
(5)

The agent received some training to increase the number of calls handled per hour. During a randomly selected 30 minute interval after the training the agent handles 14 calls.

(*c*) Test, at the 5% level of significance, whether or not there is evidence to support the suggestion that the rate at which the agent handles calls has increased. State your hypotheses clearly.
(6)

---

7. A random variable X has probability density function given by

$$f(x) = \begin{cases} \frac{1}{2}x & 0 \le x < 1 \\ kx^3 & 1 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant.

(*a*) Show that $k = \frac{1}{5}$.
(4)

(*b*) Calculate the mean of X.
(4)

(*c*) Specify fully the cumulative distribution function F(x).
(7)

(*d*) Find the median of X.
(3)

(*e*) Comment on the skewness of the distribution of X.
(2)

TOTAL FOR PAPER: 75 MARKS

END

Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced

Wednesday 21 January 2009 – Afternoon

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Green)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
There are 7 questions on this paper. The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

N30996A

1. A botanist is studying the distribution of daisies in a field. The field is divided into a number of equal sized squares. The mean number of daisies per square is assumed to be 3. The daisies are distributed randomly throughout the field.

Find the probability that, in a randomly chosen square there will be

(*a*) more than 2 daisies,
(3)

(*b*) either 5 or 6 daisies.
(2)

The botanist decides to count the number of daisies, x, in each of 80 randomly selected squares within the field. The results are summarised below

$$\sum x = 295 \qquad \sum x^2 = 1386$$

(*c*) Calculate the mean and the variance of the number of daisies per square for the 80 squares. Give your answers to 2 decimal places.
(3)

(*d*) Explain how the answers from part (*c*) support the choice of a Poisson distribution as a model.
(1)

(*e*) Using your mean from part (*c*), estimate the probability that exactly 4 daisies will be found in a randomly selected square.
(2)

---

2. The continuous random variable X is uniformly distributed over the interval $[-2, 7]$.

(*a*) Write down fully the probability density function $f(x)$ of X.
(2)

(*b*) Sketch the probability density function $f(x)$ of X.
(2)

Find

(*c*) $E(X^2)$,
(3)

(*d*) $P(-0.2 < X < 0.6)$.
(2)

---

3. A single observation x is to be taken from a Binomial distribution $B(20, p)$.

This observation is used to test $H_0 : p = 0.3$ against $H_1 : p \neq 0.3$.

(*a*) Using a 5% level of significance, find the critical region for this test. The probability of rejecting either tail should be as close as possible to 2.5%.
(3)

(*b*) State the actual significance level of this test.
(2)

The actual value of x obtained is 3.

(*c*) State a conclusion that can be drawn based on this value, giving a reason for your answer.
(2)

---

4. The length of a telephone call made to a company is denoted by the continuous random variable T. It is modelled by the probability density function

$$f(t) = \begin{cases} kt, & 0 \leq t \leq 10 \\ 0, & \text{otherwise.} \end{cases}$$

(*a*) Show that the value of k is $\frac{1}{50}$.
(3)

(*b*) Find $P(T > 6)$.
(2)

(*c*) Calculate an exact value for $E(T)$ and for $\mathrm{Var}(T)$.
(5)

(*d*) Write down the mode of the distribution of T.
(1)

It is suggested that the probability density function, $f(t)$, is not a good model for T.

(*e*) Sketch the graph of a more suitable probability density function for T.
(1)

---

5. A factory produces components of which 1% are defective. The components are packed in boxes of 10. A box is selected at random.

(*a*) Find the probability that the box contains exactly one defective component.
(2)

(*b*) Find the probability that there are at least 2 defective components in the box.
(3)

(*c*) Using a suitable approximation, find the probability that a batch of 250 components contains between 1 and 4 (inclusive) defective components.
(4)

---

6. A web server is visited on weekdays, at a rate of 7 visits per minute. In a random one minute on a Saturday the web server is visited 10 times.

(*a*) (i) Test, at the 10% level of significance, whether or not there is evidence that the rate of visits is greater on a Saturday than on weekdays. State your hypotheses clearly.

(ii) State the minimum number of visits required to obtain a significant result.
(7)

(*b*) State an assumption that has been made about the visits to the server.
(1)

In a random two minute period on a Saturday the web server is visited 20 times.

(*c*) Using a suitable approximation, test at the 10% level of significance, whether or not the rate of visits is greater on a Saturday.
(6)

7. A random variable X has probability density function given by

$$f(x) = \begin{cases} -\frac{2}{9}x + \frac{8}{9}, & 1 \leq x \leq 4 \\ 0, & \text{otherwise.} \end{cases}$$

(*a*) Show that the cumulative distribution function F(x) can be written in the form $ax^2 + bx + c$, for $1 \leq x \leq 4$ where a, b and c are constants.
(3)

(*b*) Define fully the cumulative distribution function F(x).
(2)

(*c*) Show that the upper quartile of X is 2.5 and find the lower quartile.
(6)

Given that the median of X is 1.88,

(*d*) describe the skewness of the distribution. Give a reason for your answer.
(2)

TOTAL FOR PAPER: 75 MARKS

END

PMT

Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced Level

Monday 1 June 2009 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Orange or Green)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulas stored in them.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 8 questions.
The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

M34280A

1. A bag contains a large number of counters of which 15% are coloured red. A random sample of 30 counters is selected and the number of red counters is recorded.

(*a*) Find the probability of no more than 6 red counters in this sample.

(2)

A second random sample of 30 counters is selected and the number of red counters is recorded.

(*b*) Using a Poisson approximation, estimate the probability that the total number of red counters in the combined sample of size 60 is less than 13.

(3)

---

2. An effect of a certain disease is that a small number of the red blood cells are deformed. Emily has this disease and the deformed blood cells occur randomly at a rate of 2.5 per ml of her blood. Following a course of treatment, a random sample of 2 ml of Emily's blood is found to contain only 1 deformed red blood cell.

Stating your hypotheses clearly and using a 5% level of significance, test whether or not there has been a decrease in the number of deformed red blood cells in Emily's blood.

(6)

---

3. A random sample $X_1, X_2, \ldots X_n$ is taken from a population with unknown mean μ and unknown variance σ^2. A statistic Y is based on this sample.

(*a*) Explain what you understand by the statistic Y.

(2)

(*b*) Explain what you understand by the sampling distribution of Y.

(1)

(*c*) State, giving a reason which of the following is **not** a statistic based on this sample.

(i) $\sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{n}$ (ii) $\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2$ (iii) $\sum_{i=1}^{n} X_i^2$

(2)

---

4. Past records suggest that 30% of customers who buy baked beans from a large supermarket buy them in single tins. A new manager questions whether or not there has been a change in the proportion of customers who buy baked beans in single tins. A random sample of 20 customers who had bought baked beans was taken.

(*a*) Using a 10% level of significance, find the critical region for a two-tailed test to answer the manager's question. You should state the probability of rejection in each tail which should be less than 0.05.

(5)

(*b*) Write down the actual significance level of a test based on your critical region from part (*a*).

(1)

The manager found that 11 customers from the sample of 20 had bought baked beans in single tins.

(*c*) Comment on this finding in the light of your critical region found in part (*a*).

(2)

---

5. An administrator makes errors in her typing randomly at a rate of 3 errors every 1000 words.

(*a*) In a document of 2000 words find the probability that the administrator makes 4 or more errors.

(3)

The administrator is given an 8000 word report to type and she is told that the report will only be accepted if there are 20 or fewer errors.

(*b*) Use a suitable approximation to calculate the probability that the report is accepted.

(7)

---

6. The three independent random variables A, B and C each has a continuous uniform distribution over the interval [0, 5].

(*a*) Find $P(A > 3)$.
(1)

(*b*) Find the probability that A, B and C are all greater than 3.
(2)

The random variable Y represents the maximum value of A, B and C.

The cumulative distribution function of Y is

$$F(y) = \begin{cases} 0, & y < 0 \\ \dfrac{y^3}{125}, & 0 \le y \le 5 \\ 1, & y > 5 \end{cases}$$

(*c*) Find the probability density function of Y.
(2)

(*d*) Sketch the probability density function of Y.
(2)

(*e*) Write down the mode of Y.
(1)

(*f*) Find $E(Y)$.
(3)

(*g*) Find $P(Y > 3)$.
(2)

7.

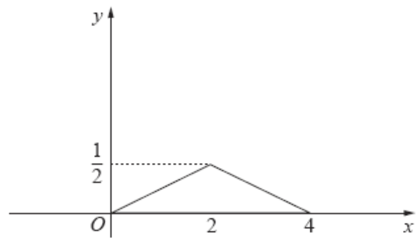

Figure 1

Figure 1 shows a sketch of the probability density function $f(x)$ of the random variable X. The part of the sketch from $x = 0$ to $x = 4$ consists of an isosceles triangle with maximum at (2, 0.5).

(*a*) Write down $E(X)$.
(1)

The probability density function $f(x)$ can be written in the following form.

$$f(x) = \begin{cases} ax & 0 \le x < 2 \\ b - ax & 2 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

(*b*) Find the values of the constants a and b.
(2)

(*c*) Show that σ, the standard deviation of X, is 0.816 to 3 decimal places.
(7)

(*d*) Find the lower quartile of X.
(3)

(*e*) State, giving a reason, whether $P(2 - \sigma < X < 2 + \sigma)$ is more or less than 0.5
(2)

PMT

8. A cloth manufacturer knows that faults occur randomly in the production process at a rate of 2 every 15 metres.

(*a*) Find the probability of exactly 4 faults in a 15 metre length of cloth.
(2)

(*b*) Find the probability of more than 10 faults in 60 metres of cloth.
(3)

A retailer buys a large amount of this cloth and sells it in pieces of length x metres. He chooses x so that the probability of no faults in a piece is 0.80.

(*c*) Write down an equation for x and show that $x = 1.7$ to 2 significant figures.
(4)

The retailer sells 1200 of these pieces of cloth. He makes a profit of 60p on each piece of cloth that does not contain a fault but a loss of £1.50 on any pieces that do contain faults.

(*d*) Find the retailer's expected profit.
(4)

TOTAL FOR PAPER: 75 MARKS

END

Paper Reference(s)
6684/01

Edexcel GCE

Statistics S2

Advanced Level

Tuesday 19 January 2010 – Morning

Time: 1 hour 30 minutes

Materials required for examination	**Items included with question papers**
Mathematical Formulae (Pink or Green)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulas stored in them.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 7 questions.
The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

M35712A

PMT

1. A manufacturer supplies DVD players to retailers in batches of 20. It has 5% of the players returned because they are faulty.

(*a*) Write down a suitable model for the distribution of the number of faulty DVD players in a batch.
(2)

Find the probability that a batch contains

(*b*) no faulty DVD players,
(2)

(*c*) more than 4 faulty DVD players.
(2)

(*d*) Find the mean and variance of the number of faulty DVD players in a batch.
(2)

---

2. A continuous random variable X has cumulative distribution function

$$\mathrm{F}(x) = \begin{cases} 0, & x < -2 \\ \dfrac{x+2}{6}, & -2 \le x \le 4 \\ 1, & x > 4 \end{cases}$$

(*a*) Find $\mathrm{P}(X < 0)$.
(2)

(*b*) Find the probability density function $\mathrm{f}(x)$ of X.
(3)

(*c*) Write down the name of the distribution of X.
(1)

(*d*) Find the mean and the variance of X.
(3)

(*e*) Write down the value of $\mathrm{P}(X = 1)$.
(1)

---

3. A robot is programmed to build cars on a production line. The robot breaks down at random at a rate of once every 20 hours.

(*a*) Find the probability that it will work continuously for 5 hours without a breakdown.
(3)

Find the probability that, in an 8 hour period,

(*b*) the robot will break down at least once,
(3)

(*c*) there are exactly 2 breakdowns.
(2)

In a particular 8 hour period, the robot broke down twice.

(*d*) Write down the probability that the robot will break down in the following 8 hour period. Give a reason for your answer.
(2)

---

4. The continuous random variable X has probability density function $\mathrm{f}(x)$ given by

$$\mathrm{f}(x) = \begin{cases} k(x^2 - 2x + 2), & 0 < x \le 3, \\ 3k, & 3 < x \le 4, \\ 0, & \text{otherwise.} \end{cases}$$

where k is a constant.

(*a*) Show that $k = \dfrac{1}{9}$.
(4)

(*b*) Find the cumulative distribution function $\mathrm{F}(x)$.
(6)

(*c*) Find the mean of X.
(3)

(*d*) Show that the median of X lies between $x = 2.6$ and $x = 2.7$.
(4)

PMT

5. A café serves breakfast every morning. Customers arrive for breakfast at random at a rate of 1 every 6 minutes.

Find the probability that

(*a*) fewer than 9 customers arrive for breakfast on a Monday morning between 10 a.m. and 11 a.m.
(3)

The café serves breakfast every day between 8 a.m. and 12 noon.

(*b*) Using a suitable approximation, estimate the probability that more than 50 customers arrive for breakfast next Tuesday.
(6)

---

6. (*a*) Define the critical region of a test statistic.
(2)

A discrete random variable X has a Binomial distribution B(30, p). A single observation is used to test $H_0 : p = 0.3$ against $H_1 : p \neq 0.3$

(*b*) Using a 1% level of significance find the critical region of this test. You should state the probability of rejection in each tail which should be as close as possible to 0.005.
(5)

(*c*) Write down the actual significance level of the test.
(1)

The value of the observation was found to be 15.

(*d*) Comment on this finding in light of your critical region.
(2)

---

7. A bag contains a large number of coins. It contains only 1p and 2p coins in the ratio 1:3.

(*a*) Find the mean μ and the variance σ^2 of the values of this population of coins.
(3)

A random sample of size 3 is taken from the bag.

(*b*) List all the possible samples.
(2)

(*c*) Find the sampling distribution of the mean value of the samples.
(6)

TOTAL FOR PAPER: 75 MARKS

END

Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced Level

Wednesday 9 June 2010 – Afternoon

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Pink)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulas stored in them.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 7 questions.
The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

H35396A

PMT

1. Explain what you understand by

(*a*) a population,
(1)

(*b*) a statistic.
(1)

A researcher took a sample of 100 voters from a certain town and asked them who they would vote for in an election. The proportion who said they would vote for Dr Smith was 35%.

(*c*) State the population and the statistic in this case.
(2)

(*d*) Explain what you understand by the sampling distribution of this statistic.
(1)

---

2. Bhim and Joe play each other at badminton and for each game, independently of all others, the probability that Bhim loses is 0.2.

Find the probability that, in 9 games, Bhim loses

(*a*) exactly 3 of the games,
(3)

(*b*) fewer than half of the games.
(2)

Bhim attends coaching sessions for 2 months. After completing the coaching, the probability that he loses each game, independently of all others, is 0.05.

Bhim and Joe agree to play a further 60 games.

(*c*) Calculate the mean and variance for the number of these 60 games that Bhim loses.
(2)

(*d*) Using a suitable approximation calculate the probability that Bhim loses more than 4 games.
(3)

---

3. A rectangle has a perimeter of 20 cm. The length, X cm, of one side of this rectangle is uniformly distributed between 1 cm and 7 cm.

Find the probability that the length of the longer side of the rectangle is more than 6 cm long.
(5)

---

4. The lifetime, X, in tens of hours, of a battery has a cumulative distribution function $\mathrm{F}(x)$ given by

$$\mathrm{F}(x) = \begin{cases} 0 & x < 1 \\ \frac{4}{9}(x^2 + 2x - 3) & 1 \le x \le 1.5 \\ 1 & x > 1.5 \end{cases}$$

(*a*) Find the median of X, giving your answer to 3 significant figures.
(3)

(*b*) Find, in full, the probability density function of the random variable X.
(3)

(*c*) Find $\mathrm{P}(X \ge 1.2)$
(2)

A camping lantern runs on 4 batteries, all of which must be working. Four new batteries are put into the lantern.

(*d*) Find the probability that the lantern will still be working after 12 hours.
(2)

---

5. A company has a large number of regular users logging onto its website. On average 4 users every hour fail to connect to the company's website at their first attempt.

(*a*) Explain why the Poisson distribution may be a suitable model in this case.
(1)

Find the probability that, in a randomly chosen **2 hour** period,

(*b*) (i) all users connect at their first attempt,

(ii) at least 4 users fail to connect at their first attempt.
(5)

The company suffered from a virus infecting its computer system. During this infection it was found that the number of users failing to connect at their first attempt, over a 12 hour period, was 60.

(*c*) Using a suitable approximation, test whether or not the mean number of users per hour who failed to connect at their first attempt had increased. Use a 5% level of significance and state your hypotheses clearly.
(9)

PMT

6. A company claims that a quarter of the bolts sent to them are faulty. To test this claim the number of faulty bolts in a random sample of 50 is recorded.

(*a*) Give two reasons why a binomial distribution may be a suitable model for the number of faulty bolts in the sample.

(2)

(*b*) Using a 5% significance level, find the critical region for a two-tailed test of the hypothesis that the probability of a bolt being faulty is $\frac{1}{4}$. The probability of rejection in either tail should be as close as possible to 0.025.

(3)

(*c*) Find the actual significance level of this test.

(2)

In the sample of 50 the actual number of faulty bolts was 8.

(*d*) Comment on the company's claim in the light of this value. Justify your answer.

(2)

The machine making the bolts was reset and another sample of 50 bolts was taken. Only 5 were found to be faulty.

(*e*) Test at the 1% level of significance whether or not the probability of a faulty bolt has decreased. State your hypotheses clearly.

(6)

7. The random variable Y has probability density function f(y) given by

$$f(y) = \begin{cases} ky(a-y) & 0 \leq y \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

where k and a are positive constants.

(*a*) (i) Explain why $a \geq 3$.

(ii) Show that $k = \dfrac{2}{9(a-2)}$.

(6)

Given that $E(Y) = 1.75$,

(*b*) show that $a = 4$ and write down the value of k.

(6)

For these values of a and k,

(*c*) sketch the probability density function,

(2)

(*d*) write down the mode of Y.

(1)

TOTAL FOR PAPER: 75 MARKS

END

PMT

Paper Reference(s)
6684/01

Edexcel GCE

Statistics S2

Advanced Level

Friday 14 January 2011 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Pink)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulas stored in them.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 7 questions.
The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

1. A disease occurs in 3% of a population.

(*a*) State any assumptions that are required to model the number of people with the disease in a random sample of size n as a binomial distribution.
(2)

(*b*) Using this model, find the probability of exactly 2 people having the disease in a random sample of 10 people.
(3)

(*c*) Find the mean and variance of the number of people with the disease in a random sample of 100 people.
(2)

A doctor tests a random sample of 100 patients for the disease. He decides to offer all patients a vaccination to protect them from the disease if more than 5 of the sample have the disease.

(*d*) Using a suitable approximation, find the probability that the doctor will offer all patients a vaccination.
(3)

---

2. A student takes a multiple choice test. The test is made up of 10 questions each with 5 possible answers. The student gets 4 questions correct. Her teacher claims she was guessing the answers. Using a one tailed test, at the 5% level of significance, test whether or not there is evidence to reject the teacher's claim.

State your hypotheses clearly.
(6)

---

3. The continuous random variable X is uniformly distributed over the interval $[-1, 3]$.
Find

(*a*) $\mathrm{E}(X)$
(1)

(*b*) $\mathrm{Var}(X)$
(2)

(*c*) $\mathrm{E}(X^2)$
(2)

(*d*) $\mathrm{P}(X < 1.4)$
(1)

A total of 40 observations of X are made.

(*e*) Find the probability that at least 10 of these observations are negative.
(5)

---

4. Richard regularly travels to work on a ferry. Over a long period of time, Richard has found that the ferry is late on average 2 times every week. The company buys a new ferry to improve the service. In the 4-week period after the new ferry is launched, Richard finds the ferry is late 3 times and claims the service has improved. Assuming that the number of times the ferry is late has a Poisson distribution, test Richard's claim at the 5% level of significance. State your hypotheses clearly.

(6)

---

5. A continuous random variable X has the probability density function f(x) shown in Figure 1.

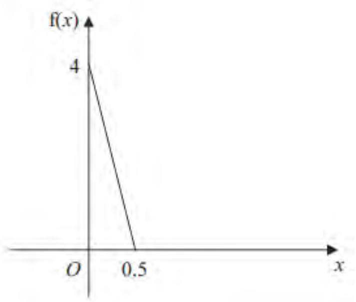

Figure 1

(a) Show that $f(x) = 4 - 8x$ for $0 \leq x \leq 0.5$ and specify f(x) for all real values of x.

(4)

(b) Find the cumulative distribution function F(x).

(4)

(c) Find the median of X.

(3)

(d) Write down the mode of X.

(1)

(e) State, with a reason, the skewness of X.

(1)

---

6. Cars arrive at a motorway toll booth at an average rate of 150 per hour.

(a) Suggest a suitable distribution to model the number of cars arriving at the toll booth, X, per minute.

(2)

(b) State clearly any assumptions you have made by suggesting this model.

(2)

Using your model,

(c) find the probability that in any given minute

(i) no cars arrive,

(ii) more than 3 cars arrive.

(3)

(d) In any given 4 minute period, find m such that $P(X > m) = 0.0487$

(3)

(e) Using a suitable approximation find the probability that fewer than 15 cars arrive in any given 10 minute period.

(6)

---

7. The queuing time in minutes, X, of a customer at a post office is modelled by the probability density function

$$f(x) = \begin{cases} kx(81 - x^2) & 0 \leq x \leq 9 \\ 0 & \text{otherwise} \end{cases}$$

(a) Show that $k = \dfrac{4}{6561}$.

(3)

Using integration, find

(b) the mean queuing time of a customer,

(4)

(c) the probability that a customer will queue for more than 5 minutes.

(3)

Three independent customers shop at the post office.

(d) Find the probability that at least 2 of the customers queue for more than 5 minutes.

(3)

TOTAL FOR PAPER: 75 MARKS

END

PMT

Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced Level

Thursday 26 May 2011 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Pink)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulas stored in them.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 7 questions.
The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

P38165A

1. A factory produces components. Each component has a unique identity number and it is assumed that 2% of the components are faulty. On a particular day, a quality control manager wishes to take a random sample of 50 components.

(*a*) Identify a sampling frame.

(1)

The statistic F represents the number of faulty components in the random sample of size 50.

(*b*) Specify the sampling distribution of F.

(2)

---

2. A traffic officer monitors the rate at which vehicles pass a fixed point on a motorway. When the rate exceeds 36 vehicles per minute he must switch on some speed restrictions to improve traffic flow.

(*a*) Suggest a suitable model to describe the number of vehicles passing the fixed point in a 15 s interval.

(1)

The traffic officer records 12 vehicles passing the fixed point in a 15 s interval.

(*b*) Stating your hypotheses clearly, and using a 5% level of significance, test whether or not the traffic officer has sufficient evidence to switch on the speed restrictions.

(6)

(*c*) Using a 5% level of significance, determine the smallest number of vehicles the traffic officer must observe in a 10 s interval in order to have sufficient evidence to switch on the speed restrictions.

(3)

3.

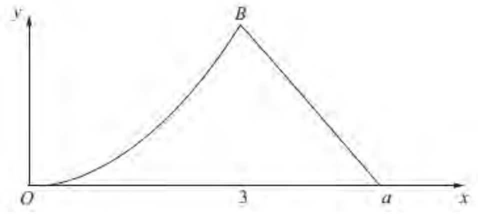

Figure 1

Figure 1 shows a sketch of the probability density function f(x) of the random variable X.

For $0 \leq x \leq 3$, f(x) is represented by a curve OB with equation $f(x) = kx^2$, where k is a constant.

For $3 \leq x \leq a$, where a is a constant, f(x) is represented by a straight line passing through B and the point (a, 0).

For all other values of x, f(x) = 0.

Given that the mode of X = the median of X, find

(a) the mode,
(1)

(b) the value of k,
(4)

(c) the value of a.
(3)

Without calculating E(X) and with reference to the skewness of the distribution

(d) state, giving your reason, whether $E(X) < 3$, $E(X) = 3$ or $E(X) > 3$.
(2)

4. In a game, players select sticks at random from a box containing a large number of sticks of different lengths. The length, in cm, of a randomly chosen stick has a continuous uniform distribution over the interval [7, 10].

A stick is selected at random from the box.

(a) Find the probability that the stick is shorter than 9.5 cm.
(2)

To win a bag of sweets, a player must select 3 sticks and wins if the length of the longest stick is more than 9.5 cm.

(b) Find the probability of winning a bag of sweets.
(2)

To win a soft toy, a player must select 6 sticks and wins the toy if more than four of the sticks are shorter than 7.6 cm.

(c) Find the probability of winning a soft toy.
(4)

5. Defects occur at random in planks of wood with a constant rate of 0.5 per 10 cm length. Jim buys a plank of length 100 cm.

(a) Find the probability that Jim's plank contains at most 3 defects.
(2)

Shivani buys 6 planks each of length 100 cm.

(b) Find the probability that fewer than 2 of Shivani's planks contain at most 3 defects.
(5)

(c) Using a suitable approximation, estimate the probability that the total number of defects on Shivani's 6 planks is less than 18.
(6)

6. A shopkeeper knows, from past records, that 15% of customers buy an item from the display next to the till. After a refurbishment of the shop, he takes a random sample of 30 customers and finds that only 1 customer has bought an item from the display next to the till.

(*a*) Stating your hypotheses clearly, and using a 5% level of significance, test whether or not there has been a change in the proportion of customers buying an item from the display next to the till.

(6)

During the refurbishment a new sandwich display was installed. Before the refurbishment 20% of customers bought sandwiches. The shopkeeper claims that the proportion of customers buying sandwiches has now increased. He selects a random sample of 120 customers and finds that 31 of them have bought sandwiches.

(*b*) Using a suitable approximation and stating your hypotheses clearly, test the shopkeeper's claim. Use a 10% level of significance.

(8)

7. The continuous random variable X has probability density function given by

$$\mathrm{f}(x) = \begin{cases} \dfrac{3}{32}(x-1)(5-x) & 1 \leq x \leq 5, \\ 0 & \text{otherwise.} \end{cases}$$

(*a*) Sketch $\mathrm{f}(x)$ showing clearly the points where it meets the x-axis.

(2)

(*b*) Write down the value of the mean, μ, of X.

(1)

(*c*) Show that $\mathrm{E}(X^2) = 9.8$.

(4)

(*d*) Find the standard deviation, σ, of X.

(2)

The cumulative distribution function of X is given by

$$\mathrm{F}(x) = \begin{cases} 0 & x < 1 \\ \dfrac{1}{32}(a - 15x + 9x^2 - x^3) & 1 \leq x \leq 5 \\ 1 & x > 5 \end{cases}$$

where a is a constant.

(*e*) Find the value of a.

(2)

(*f*) Show that the lower quartile of X, q_1, lies between 2.29 and 2.31.

(3)

(*g*) Hence find the upper quartile of X, giving your answer to 1 decimal place.

(1)

(*h*) Find, to 2 decimal places, the value of k so that

$$\mathrm{P}(\mu - k\sigma < X < \mu + k\sigma) = 0.5.$$

(2)

TOTAL FOR PAPER: 75 MARKS

END

PMT

Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced Level

Tuesday 17 January 2012 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Pink)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulas stored in them.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 7 questions.
The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

P40700A

1. The time in minutes that Elaine takes to checkout at her local supermarket follows a continuous uniform distribution defined over the interval [3, 9].

Find

(*a*) Elaine's expected checkout time,
(1)

(*b*) the variance of the time taken to checkout at the supermarket,
(2)

(*c*) the probability that Elaine will take more than 7 minutes to checkout.
(2)

Given that Elaine has already spent 4 minutes at the checkout,

(*d*) find the probability that she will take a total of less than 6 minutes to checkout.
(3)

---

2. David claims that the weather forecasts produced by local radio are no better than those achieved by tossing a fair coin and predicting rain if a head is obtained or no rain if a tail is obtained. He records the weather for 30 randomly selected days. The local radio forecast is correct on 21 of these days.

Test David's claim at the 5% level of significance.

State your hypotheses clearly.
(7)

---

3. The probability of a telesales representative making a sale on a customer call is 0.15.

Find the probability that

(*a*) no sales are made in 10 calls,
(2)

(*b*) more than 3 sales are made in 20 calls.
(2)

Representatives are required to achieve a mean of at least 5 sales each day.

(*c*) Find the least number of calls each day a representative should make to achieve this requirement.
(2)

(*d*) Calculate the least number of calls that need to be made by a representative for the probability of at least 1 sale to exceed 0.95.
(3)

PMT

4. A website receives hits at a rate of 300 per hour.

(*a*) State a distribution that is suitable to model the number of hits obtained during a 1 minute interval.
(1)

(*b*) State two reasons for your answer to part (*a*).
(2)

Find the probability of

(*c*) 10 hits in a given minute,
(3)

(*d*) at least 15 hits in 2 minutes.
(3)

The website will go down if there are more than 70 hits in 10 minutes.

(*e*) Using a suitable approximation, find the probability that the website will go down in a particular 10 minute interval.
(7)

---

5. The probability of an electrical component being defective is 0.075.
The component is supplied in boxes of 120.

(*a*) Using a suitable approximation, estimate the probability that there are more than 3 defective components in a box.
(5)

A retailer buys 2 boxes of components.

(*b*) Estimate the probability that there are at least 4 defective components in each box.
(2)

---

6. A random variable X has probability density function given by

$$\mathrm{f}(x) = \begin{cases} \dfrac{1}{2}, & 0 \le x < 1, \\ x - \dfrac{1}{2}, & 1 \le x \le k, \\ 0 & \text{otherwise,} \end{cases}$$

where k is a positive constant.

(*a*) Sketch the graph of $\mathrm{f}(x)$.
(2)

(*b*) Show that $k = \dfrac{1}{2}(1 + \sqrt{5})$.
(4)

(*c*) Define fully the cumulative distribution function $\mathrm{F}(x)$.
(6)

(*d*) Find $\mathrm{P}(0.5 < X < 1.5)$.
(2)

(*e*) Write down the median of X and the mode of X.
(2)

(*f*) Describe the skewness of the distribution of X. Give a reason for your answer.
(2)

PMT

7. (*a*) Explain briefly what you understand by

(i) a critical region of a test statistic,

(ii) the level of significance of a hypothesis test.

(2)

(*b*) An estate agent has been selling houses at a rate of 8 per month. She believes that the rate of sales will decrease in the next month.

(i) Using a 5% level of significance, find the critical region for a one tailed test of the hypothesis that the rate of sales will decrease from 8 per month.

(ii) Write down the actual significance level of the test in part (*b*)(i).

(3)

The estate agent is surprised to find that she actually sold 13 houses in the next month. She now claims that this is evidence of an increase in the rate of sales per month.

(*c*) Test the estate agent's claim at the 5% level of significance. State your hypotheses clearly.

(5)

TOTAL FOR PAPER: 75 MARKS

END

Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced Level

Thursday 24 May 2012 – Morning

Time: 1 hour 30 minutes

Materials required for examination	**Items included with question papers**
Mathematical Formulae (Pink)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulas stored in them.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 8 questions.
The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

P40106A

PMT

1. A manufacturer produces sweets of length L mm where L has a continuous uniform distribution with range [15, 30].

(*a*) Find the probability that a randomly selected sweet has length greater than 24 mm.
(2)

These sweets are randomly packed in bags of 20 sweets.

(*b*) Find the probability that a randomly selected bag will contain at least 8 sweets with length greater than 24 mm.
(3)

(*c*) Find the probability that 2 randomly selected bags will both contain at least 8 sweets with length greater than 24 mm.
(2)

---

2. A test statistic has a distribution B(25, p).

Given that

$$H_0 : p = 0.5, \quad H_1 : p \neq 0.5,$$

(*a*) find the critical region for the test statistic such that the probability in each tail is as close as possible to 2.5%.
(3)

(*b*) State the probability of incorrectly rejecting H_0 using this critical region.
(2)

---

3. (*a*) Write down the two conditions needed to approximate the binomial distribution by the Poisson distribution.
(2)

A machine which manufactures bolts is known to produce 3% defective bolts. The machine breaks down and a new machine is installed. A random sample of 200 bolts is taken from those produced by the new machine and 12 bolts are defective.

(*b*) Using a suitable approximation, test at the 5% level of significance whether or not the proportion of defective bolts is higher with the new machine than with the old machine. State your hypotheses clearly.
(7)

---

4. The number of houses sold by an estate agent follows a Poisson distribution, with a mean of 2 per week.

(*a*) Find the probability that in the next four weeks the estate agent sells

(i) exactly 3 houses,

(ii) more than 5 houses.
(5)

The estate agent monitors sales in periods of 4 weeks.

(*b*) Find the probability that in the next twelve of those 4 week periods there are exactly nine periods in which more than 5 houses are sold.
(3)

The estate agent will receive a bonus if he sells more than 25 houses in the next 10 weeks.

(*c*) Use a suitable approximation to estimate the probability that the estate agent receives a bonus.
(6)

---

5. The queuing time, X minutes, of a customer at a till of a supermarket has probability density function

$$f(x) = \begin{cases} \frac{3}{32}x(k-x) & 0 \le x \le k, \\ 0 & \text{otherwise.} \end{cases}$$

(*a*) Show that the value of k is 4.
(4)

(*b*) Write down the value of $\mathrm{E}(X)$.
(1)

(*c*) Calculate $\mathrm{Var}(X)$.
(4)

(*d*) Find the probability that a randomly chosen customer's queuing time will differ from the mean by at least half a minute.
(3)

---

6. A bag contains a large number of balls.

65% are numbered 1

35% are numbered 2

A random sample of 3 balls is taken from the bag.

Find the sampling distribution for the range of the numbers on the 3 selected balls.
(6)

---

7. The continuous random variable X has probability density function $f(x)$ given by

$$f(x) = \begin{cases} \frac{x^2}{45} & 0 \le x \le 3, \\ \frac{1}{5} & 3 < x < 4, \\ \frac{1}{3} - \frac{x}{30} & 4 \le x \le 10, \\ 0 & \text{otherwise.} \end{cases}$$

(*a*) Sketch $f(x)$ for $0 \le x \le 10$.
(4)

(*b*) Find the cumulative distribution function $F(x)$ for all values of x.
(8)

(*c*) Find $P(X \le 8)$.
(2)

---

8. In a large restaurant an average of 3 out of every 5 customers ask for water with their meal.

A random sample of 10 customers is selected.

(*a*) Find the probability that

(i) exactly 6 ask for water with their meal,

(ii) less than 9 ask for water with their meal.
(5)

A second random sample of 50 customers is selected.

(*b*) Find the smallest value of n such that

$$P(X < n) \ge 0.9,$$

where the random variable X represents the number of these customers who ask for water.
(3)

TOTAL FOR PAPER: 75 MARKS

END

PMT

Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced Level

Friday 18 January 2013 – Afternoon

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Pink)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulas stored in them.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 7 questions.
The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

P41482A

1. (*a*) Write down the conditions under which the Poisson distribution can be used as an approximation to the binomial distribution.
(2)

The probability of any one letter being delivered to the wrong house is 0.01.
On a randomly selected day Peter delivers 1000 letters.

(*b*) Using a Poisson approximation, find the probability that Peter delivers at least 4 letters to the wrong house.

Give your answer to 4 decimal places.
(3)

---

2. In a village, power cuts occur randomly at a rate of 3 per year.

(*a*) Find the probability that in any given year there will be

(i) exactly 7 power cuts,

(ii) at least 4 power cuts.
(5)

(*b*) Use a suitable approximation to find the probability that in the next 10 years the number of power cuts will be less than 20.
(6)

---

3. A random variable X has the distribution B(12, p).

(*a*) Given that $p = 0.25$, find

(i) $P(X < 5)$,

(ii) $P(X \geq 7)$.
(3)

(*b*) Given that $P(X = 0) = 0.05$, find the value of p to 3 decimal places.
(3)

(*c*) Given that the variance of X is 1.92, find the possible values of p.
(4)

PMT

4. The continuous random variable X is uniformly distributed over the interval $[-4, 6]$.

(*a*) Write down the mean of X.
(1)

(*b*) Find $P(X \leq 2.4)$.
(2)

(*c*) Find $P(-3 < X - 5 < 3)$.
(2)

The continuous random variable Y is uniformly distributed over the interval $[a, 4a]$.

(*d*) Use integration to show that $E(Y^2) = 7a^2$.
(4)

(*e*) Find $\text{Var}(Y)$.
(2)

(*f*) Given that $P(X < \frac{8}{3}) = P(Y < \frac{8}{3})$, find the value of a.
(3)

---

5. The continuous random variable T is used to model the number of days, t, a mosquito survives after hatching.

The probability that the mosquito survives for more than t days is

$$\frac{225}{(t+15)^2}, \quad t \geq 0.$$

(*a*) Show that the cumulative distribution function of T is given by

$$F(t) = \begin{cases} 1 - \dfrac{225}{(t+15)^2}, & t \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

(1)

(*b*) Find the probability that a randomly selected mosquito will die within 3 days of hatching.
(2)

(*c*) Given that a mosquito survives for 3 days, find the probability that it will survive for at least 5 more days.
(3)

A large number of mosquitoes hatch on the same day.

(*d*) Find the number of days after which only 10% of these mosquitoes are expected to survive.
(4)

---

6. (*a*) Explain what you understand by a hypothesis.
(1)

(*b*) Explain what you understand by a critical region.
(2)

Mrs George claims that 45% of voters would vote for her.

In an opinion poll of 20 randomly selected voters it was found that 5 would vote for her.

(*c*) Test at the 5% level of significance whether or not the opinion poll provides evidence to support Mrs George's claim.
(4)

In a second opinion poll of n randomly selected people it was found that no one would vote for Mrs George.

(*d*) Using a 1% level of significance, find the smallest value of n for which the hypothesis $H_0 : p = 0.45$ will be rejected in favour of $H_1 : p < 0.45$.
(3)

---

7. The continuous random variable X has the following probability density function

$$f(x) = \begin{cases} a + bx, & 0 \leq x \leq 5, \\ 0, & \text{otherwise.} \end{cases}$$

where a and b are constants.

(*a*) Show that $10a + 25b = 2$.
(4)

Given that $E(X) = \frac{35}{12}$,

(*b*) find a second equation in a and b,
(3)

(*c*) hence find the value of a and the value of b.
(3)

(*d*) Find, to 3 significant figures, the median of X.
(3)

(*e*) Comment on the skewness. Give a reason for your answer.
(2)

TOTAL FOR PAPER: 75 MARKS

END

PMT

Paper Reference(s)
6684/01R

Edexcel GCE

Statistics S2 (R)

Advanced/Advanced Subsidiary

Friday 24 May 2013 – Morning

Time: 1 hour 30 minutes

Materials required for examination	**Items included with question papers**
Mathematical Formulae (Pink)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation or symbolic differentiation/integration, or have retrievable mathematical formulae stored in them.

This paper is strictly for students outside the UK.

Instructions to Candidates

In the boxes above, write your centre number, candidate number, your surname, initials and signature.
Check that you have the correct question paper.
Answer ALL the questions.
You must write your answer for each question in the space following the question.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for the parts of questions are shown in round brackets, e.g. **(2)**.
There are 7 questions in this question paper. The total mark for this paper is 75.
There are 24 pages in this question paper. Any blank pages are indicated.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

P42832A

1. A bag contains a large number of counters. A third of the counters have a number 5 on them and the remainder have a number 1.

A random sample of 3 counters is selected.

(*a*) List all possible samples.
(2)

(*b*) Find the sampling distribution for the range.
(3)

---

2. The continuous random variable Y has cumulative distribution function

$$\mathrm{F}(y) = \begin{cases} 0 & y < 0 \\ \frac{1}{4}(y^3 - 4y^2 + ky) & 0 \le y \le 2 \\ 1 & y > 2 \end{cases}$$

where k is a constant.

(*a*) Find the value of k.
(2)

(*b*) Find the probability density function of Y, specifying it for all values of y.
(3)

(*c*) Find $\mathrm{P}(Y > 1)$.
(2)

PMT

3. The random variable X has a continuous uniform distribution on $[a, b]$ where a and b are positive numbers.

Given that $\mathrm{E}(X) = 23$ and $\mathrm{Var}(X) = 75$,

(a) find the value of a and the value of b.
(6)

Given that $\mathrm{P}(X > c) = 0.32$,

(b) find $\mathrm{P}(23 < X < c)$.
(2)

---

4. The random variable X has probability density function $\mathrm{f}(x)$ given by

$$\mathrm{f}(x) = \begin{cases} k(3+2x-x^2) & 0 \leq x \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant.

(a) Show that $k = \frac{1}{9}$.
(3)

(b) Find the mode of X.
(2)

(c) Use algebraic integration to find $\mathrm{E}(X)$.
(4)

By comparing your answers to parts (b) and (c),

(d) describe the skewness of X, giving a reason for your answer.
(2)

---

5. In a village shop the customers must join a queue to pay. The number of customers joining the queue in a 10 minute interval is modelled by a Poisson distribution with mean 3.

Find the probability that

(a) exactly 4 customers join the queue in the next 10 minutes,
(2)

(b) more than 10 customers join the queue in the next 20 minutes.
(3)

When a customer reaches the front of the queue the customer pays the assistant. The time each customer takes paying the assistant, T minutes, has a continuous uniform distribution over the interval $[0, 5]$. The random variable T is independent of the number of people joining the queue.

(c) Find $\mathrm{P}(T > 3.5)$.
(1)

In a random sample of 5 customers, the random variable C represents the number of customers who took more than 3.5 minutes paying the assistant.

(d) Find $\mathrm{P}(C \geq 3)$.
(3)

Bethan has just reached the front of the queue and starts paying the assistant.

(e) Find the probability that in the next 4 minutes Bethan finishes paying the assistant and no other customers join the queue.
(4)

PMT

6. Frugal bakery claims that their packs of 10 muffins contain on average 80 raisins per pack. A Poisson distribution is used to describe the number of raisins per muffin.

A muffin is selected at random to test whether or not the mean number of raisins per muffin has changed.

(*a*) Find the critical region for a two-tailed test using a 10% level of significance. The probability of rejection in each tail should be less than 0.05.

(4)

(*b*) Find the actual significance level of this test.

(2)

The bakery has a special promotion claiming that their muffins now contain even more raisins.

A random sample of 10 muffins is selected and is found to contain a total of 95 raisins.

(*c*) Use a suitable approximation to test the bakery's claim. You should state your hypotheses clearly and use a 5% level of significance.

(8)

7. As part of a selection procedure for a company, applicants have to answer all 20 questions of a multiple choice test. If an applicant chooses answers at random the probability of choosing a correct answer is 0.2 and the number of correct answers is represented by the random variable X.

(*a*) Suggest a suitable distribution for X.

(2)

Each applicant gains 4 points for each correct answer but loses 1 point for each incorrect answer. The random variable S represents the final score, in points, for an applicant who chooses answers to this test at random.

(*b*) Show that $S = 5X - 20$.

(2)

(*c*) Find $\mathrm{E}(S)$ and $\mathrm{Var}(S)$.

(4)

An applicant who achieves a score of at least 20 points is invited to take part in the final stage of the selection process.

(*d*) Find $\mathrm{P}(S \geq 20)$.

(4)

Cameron is taking the final stage of the selection process which is a multiple choice test consisting of 100 questions. He has been preparing for this test and believes that his chance of answering each question correctly is 0.4.

(*e*) Using a suitable approximation, estimate the probability that Cameron answers more than half of the questions correctly.

(5)

TOTAL FOR PAPER: 75 MARKS

END

PMT

Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced Level

Friday 24 May 2013 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Pink)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulas stored in them.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 7 questions.
The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

P42035A

1. A bag contains a large number of 1p, 2p and 5p coins.

50% are 1p coins
20% are 2p coins
30% are 5p coins

A random sample of 3 coins is chosen from the bag.

(*a*) List all the possible samples of size 3 with median 5p. **(2)**

(*b*) Find the probability that the median value of the sample is 5p. **(4)**

(*c*) Find the sampling distribution of the median of samples of size 3. **(5)**

---

2. The number of defects per metre in a roll of cloth has a Poisson distribution with mean 0.25.

Find the probability that

(*a*) a randomly chosen metre of cloth has 1 defect, **(2)**

(*b*) the total number of defects in a randomly chosen 6 metre length of cloth is more than 2. **(3)**

A tailor buys 300 metres of cloth.

(*c*) Using a suitable approximation find the probability that the tailor's cloth will contain less than 90 defects. **(5)**

PMT

3. An online shop sells a computer game at an average rate of 1 per day.

(*a*) Find the probability that the shop sells more than 10 games in a 7 day period.
(3)

Once every 7 days the shop has games delivered before it opens.

(*b*) Find the least number of games the shop should have in stock immediately after a delivery so that the probability of running out of the game before the next delivery is less than 0.05.
(3)

In an attempt to increase sales of the computer game, the price is reduced for six months. A random sample of 28 days is taken from these six months. In the sample of 28 days, 36 computer games are sold.

(*c*) Using a suitable approximation and a 5% level of significance, test whether or not the average rate of sales per day has increased during these six months. State your hypotheses clearly.
(7)

---

4. A continuous random variable X is uniformly distributed over the interval $[b, 4b]$ where b is a constant.

(*a*) Write down $\mathrm{E}(X)$.
(1)

(*b*) Use integration to show that $\mathrm{Var}(X) = \dfrac{3b^2}{4}$.
(3)

(*c*) Find $\mathrm{Var}(3 - 2X)$.
(2)

Given that $b = 1$, find

(*d*) the cumulative distribution function of X, $\mathrm{F}(x)$, for all values of x,
(2)

(*e*) the median of X.
(1)

---

5. The continuous random variable X has a cumulative distribution function

$$\mathrm{F}(x) = \begin{cases} 0, & x < 1, \\ \dfrac{x^3}{10} + \dfrac{3x^2}{10} + ax + b, & 1 \le x \le 2, \\ 1, & x > 2, \end{cases}$$

where a and b are constants.

(*a*) Find the value of a and the value of b.
(4)

(*b*) Show that $\mathrm{f}(x) = \dfrac{3}{10}(x^2 + 2x - 2), \quad 1 \le x \le 2$.
(1)

(*c*) Use integration to find $\mathrm{E}(X)$.
(4)

(*d*) Show that the lower quartile of X lies between 1.425 and 1.435.
(3)

---

6. In a manufacturing process 25% of articles are thought to be defective. Articles are produced in batches of 20.

(*a*) A batch is selected at random. Using a 5% significance level, find the critical region for a two tailed test that the probability of an article chosen at random being defective is 0.25.

You should state the probability in each tail, which should be as close as possible to 0.025.
(5)

The manufacturer changes the production process to try to reduce the number of defective articles. She then chooses a batch at random and discovers there are 3 defective articles.

(*b*) Test at the 5% level of significance whether or not there is evidence that the changes to the process have reduced the percentage of defective articles. State your hypotheses clearly.
(5)

PMT

7. A telesales operator is selling a magazine. Each day he chooses a number of people to telephone. The probability that each person he telephones buys the magazine is 0.1.

(*a*) Suggest a suitable distribution to model the number of people who buy the magazine from the telesales operator each day.

(1)

(*b*) On Monday, the telesales operator telephones 10 people. Find the probability that he sells at least 4 magazines.

(3)

(*c*) Calculate the least number of people he needs to telephone on Tuesday, so that the probability of selling at least 1 magazine, on that day, is greater than 0.95.

(3)

A call centre also sells the magazine. The probability that a telephone call made by the call centre sells a magazine is 0.05. The call centre telephones 100 people every hour.

(*d*) Using a suitable approximation, find the probability that more than 10 people telephoned by the call centre buy a magazine in a randomly chosen hour.

(3)

TOTAL FOR PAPER: 75 MARKS

END

Paper Reference(s)
WST02/01

Pearson Edexcel International Advanced Level

Statistics S2

Advanced/Advanced Subsidiary

Friday 17 January 2014 – Afternoon

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Blue)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.

Instructions

- Use **black** ink or ball-point pen.
- If pencil is used for diagrams/sketches/graphs it must be dark (HB or B). Coloured pencils and highlighter pens must not be used.
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- Answer **all** questions and ensure that your answers to parts of questions are clearly labelled.
- Answer the questions in the spaces provided – *there may be more space than you need.*
- You should show sufficient working to make your methods clear. Answers without working may not gain full credit.
- Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information

- The total mark for this paper is 75.
- The marks for **each** question are shown in brackets – *use this as a guide as to how much time to spend on each question.*

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.

P42962XA

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)

---

2. Bill owns a restaurant. Over the next four weeks Bill decides to carry out a sample survey to obtain the customers' opinions.

(*a*) Suggest a suitable sampling frame for the sample survey.

(1)

(*b*) Identify the sampling units.

(1)

(*c*) Give one advantage and one disadvantage of taking a census rather than a sample survey.

(2)

Bill believes that only 30% of customers would like a greater choice on the menu. He takes a random sample of 50 customers and finds that 20 of them would like a greater choice on the menu.

(*d*) Test, at the 5% significance level, whether or not the percentage of customers who would like a greater choice on the menu is more than Bill believes. State your hypotheses clearly.

(6)

---

3. The continuous random variable X has cumulative distribution function given by

$$F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{1}{6}x(x+1) & 0 \leq x \leq 2 \\ 1 & x > 2 \end{cases}$$

(*a*) Find the value of a such that $P(X > a) = 0.4$.

Give your answer to 3 significant figures.

(3)

(*b*) Use calculus to find (i) $E(X)$

(ii) $Var(X)$.

(8)

---

4. The number of telephone calls per hour received by a business is a random variable with distribution $Po(\lambda)$.

Charlotte records the number of calls, C, received in 4 hours.

A test of the null hypothesis H_0: $\lambda = 1.5$ is carried out.

H_0 is rejected if $C > 10$.

(*a*) Write down the alternative hypothesis.

(1)

(*b*) Find the significance level of the test.

(3)

Given that $P(C > 10) < 0.1$,

(*c*) find the largest possible value of λ that can be found by using the tables.

(3)

---

5. A school photocopier breaks down randomly at a rate of 15 times per year.

(*a*) Find the probability that there will be exactly 3 breakdowns in the next month.
(3)

(*b*) Show that the probability that there will be at least 2 breakdowns in the next month is 0.355 to 3 decimal places.
(2)

(*c*) Find the probability of at least 2 breakdowns in each of the next 4 months.
(2)

The teachers would like a new photocopier. The head teacher agrees to monitor the situation for the next 12 months. The head teacher decides he will buy a new photocopier if there is more than 1 month when the photocopier has at least 2 breakdowns.

(*d*) Find the probability that the head teacher will buy a new photocopier.
(5)

---

6. The continuous random variable X has probability density function given by

$$f(x) = \begin{cases} k(x+1)^2 & -1 \leq x \leq 1 \\ k(6-2x) & 1 < x \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

where k is a positive constant.

(*a*) Sketch the graph of f(x).
(2)

(*b*) Show that the value of k is $\frac{3}{20}$.
(5)

(*c*) Define fully the cumulative distribution function F(x).
(5)

(*d*) Find the median of X, giving your answer to 3 significant figures.
(3)

---

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

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)

TOTAL FOR PAPER: 75 MARKS

END

Paper Reference(s)

WST02/01

Pearson Edexcel International Advanced Level

Statistics S2

Advanced/Advanced Subsidiary

Tuesday 24 June 2014 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Blue)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.

Instructions

- Use **black** ink or ball-point pen.
- If pencil is used for diagrams/sketches/graphs it must be dark (HB or B). Coloured pencils and highlighter pens must not be used.
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- Answer **all** questions and ensure that your answers to parts of questions are clearly labelled.
- Answer the questions in the spaces provided – *there may be more space than you need.*
- You should show sufficient working to make your methods clear. Answers without working may not gain full credit.
- Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information

- The total mark for this paper is 75.
- The marks for **each** question are shown in brackets – *use this as a guide as to how much time to spend on each question.*

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.

P43173A

1. (*a*) State the conditions under which the Poisson distribution may be used as an approximation to the binomial distribution.
(1)

A farmer supplies a bakery with eggs. The manager of the bakery claims that the proportion of eggs having a double yolk is 0.009.
The farmer claims that the proportion of his eggs having a double yolk is more than 0.009.

(*b*) State suitable hypotheses for testing these claims.
(1)

In a batch of 500 eggs the baker records 9 eggs with a double yolk.

(*c*) Using a suitable approximation, test at the 5% level of significance whether or not this supports the farmer's claim.
(5)

---

2. The amount of flour used by a factory in a week is Y thousand kg where Y has probability density function

$$\mathrm{f}(y) = \begin{cases} k\left(4 - y^2\right) & 0 \leq y \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

(*a*) Show that the value of k is $\dfrac{3}{16}$.
(4)

Use algebraic integration to find

(*b*) the mean number of kilograms of flour used by the factory in a week,
(4)

(*c*) the standard deviation of the number of kilograms of flour used by the factory in a week,
(5)

(*d*) the probability that more than 1500 kg of flour will be used by the factory next week.
(3)

---

3. The continuous random variable T is uniformly distributed on the interval $[\alpha, \beta]$ where $\beta > \alpha$.

Given that $\mathrm{E}(T) = 2$ and $\mathrm{Var}(T) = 163$, find

(*a*) the value of α and the value of β,
(5)

(*b*) $\mathrm{P}(T < 3.4)$.
(2)

---

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)

---

5. A company claims that 35% of its peas germinate. In order to test this claim Ann decides to plant 15 of these peas and record the number which germinate.

(*a*) (i) State suitable hypotheses for a two-tailed test of this claim.

(ii) Using a 5% level of significance, find an appropriate critical region for this test. The probability in each of the tails should be as close to 2.5% as possible.
(4)

(*b*) Ann found that 8 of the 15 peas germinated. State whether or not the company's claim is supported. Give a reason for your answer.
(2)

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

---

6. A continuous random variable X has cumulative distribution function $\mathrm{F}(x)$ given by

$$\mathrm{F}(x) = \begin{cases} 0 & x < 2 \\ \dfrac{x^2}{20}(9-2x) & 0 \leq x \leq 2 \\ 1 & x > 2 \end{cases}$$

(*a*) Verify that the median of X lies between 1.23 and 1.24.
(3)

(*b*) Specify fully the probability density function $\mathrm{f}(x)$.
(3)

(*c*) Find the mode of X.
(2)

(*d*) Describe the skewness of this distribution. Justify your answer.
(2)

---

7. Flaws occur at random in a particular type of material at a mean rate of 2 per 50 m.

(*a*) Find the probability that in a randomly chosen 50 m length of this material there will be exactly 5 flaws.
(2)

This material is sold in rolls of length 200 m. Susie buys 4 rolls of this material.

(*b*) Find the probability that only one of these rolls will have fewer than 7 flaws.
(6)

A piece of this material of length x m is produced.

Using a normal approximation, the probability that this piece of material contains fewer than 26 flaws is 0.5398.

(*c*) Find the value of x.
(8)

TOTAL FOR PAPER: 75 MARKS

END

Paper Reference(s)

6684/01R

Edexcel GCE

Statistics S2 (R)

Advanced/Advanced Subsidiary

Tuesday 24 June 2014 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Pink)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation or symbolic differentiation/integration, or have retrievable mathematical formulae stored in them.

This paper is strictly for students outside the UK.

Instructions to Candidates

In the boxes above, write your centre number, candidate number, your surname, initials and signature.
Check that you have the correct question paper.
Answer ALL the questions.
You must write your answer for each question in the space following the question.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for the parts of questions are shown in round brackets, e.g. **(2)**.
There are 7 questions in this question paper. The total mark for this paper is 75.
There are 24 pages in this question paper. Any blank pages are indicated.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

P43141A

1. Before Roger will use a tennis ball he checks it using a "bounce" test. The probability that a ball from Roger's usual supplier fails the bounce test is 0.2. A new supplier claims that the probability of one of their balls failing the bounce test is less than 0.2. Roger checks a random sample of 40 balls from the new supplier and finds that 3 balls fail the bounce test.

Stating your hypotheses clearly, use a 5% level of significance to test the new supplier's claim.

(5)

---

2. A bag contains a large number of counters. Each counter has a single digit number on it and the mean of all the numbers in the bag is the unknown parameter μ. The number 2 is on 40% of the counters and the number 5 is on 25% of the counters. All the remaining counters have numbers greater than 5 on them.

A random sample of 10 counters is taken from the bag.

(*a*) State whether or not each of the following is a statistic

(i) S = the sum of the numbers on the counters in the sample,

(ii) D = the difference between the highest number in the sample and μ,

(iii) F = the number of counters in the sample with a number 5 on them.

(3)

The random variable T represents the number of counters in a random sample of 10 with the number 2 on them.

(*b*) Specify the sampling distribution of T.

(2)

The counters are selected one by one.

(*c*) Find the probability that the third counter selected is the first counter with the number 2 on it.

(2)

---

3. Accidents occur randomly at a road junction at a rate of 18 every year.
The random variable X represents the number of accidents at this road junction in the next 6 months.

(*a*) Write down the distribution of X.
(2)

(*b*) Find $P(X > 7)$.
(2)

(*c*) Show that the probability of at least one accident in a randomly selected month is 0.777 (correct to 3 decimal places).
(3)

(*d*) Find the probability that there is at least one accident in exactly 4 of the next 6 months.
(3)

---

4. The random variable X has probability density function $f(x)$ given by

$$f(x) = \begin{cases} 3k & 0 \le x < 1 \\ kx(4-x) & 1 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant.

(*a*) Sketch $f(x)$.
(3)

(*b*) Write down the mode of X.
(1)

Given that $E(X) = \frac{29}{16}$

(*c*) describe, giving a reason, the skewness of the distribution.
(2)

(*d*) Use integration to find the value of k.
(5)

(*e*) Write down the lower quartile of X.
(1)

Given also that $P(2 < X < 3) = \frac{11}{36}$

(*f*) find the exact value of $P(X > 3)$.
(2)

---

5. Sammy manufactures wallpaper. She knows that defects occur randomly in the manufacturing process at a rate of 1 every 8 metres. Once a week the machinery is cleaned and reset. Sammy then takes a random sample of 40 metres of wallpaper from the next batch produced to test if there has been any change in the rate of defects.

(*a*) Stating your hypotheses clearly and using a 10% level of significance, find the critical region for this test. You should choose your critical region so that the probability of rejection is less than 0.05 in each tail.
(4)

(*b*) State the actual significance level of this test.
(2)

Thomas claims that his new machine would reduce the rate of defects and invites Sammy to test it. Sammy takes a random sample of 200 metres of wallpaper produced on Thomas' machine and finds 19 defects.

(*c*) Using a suitable approximation, test Thomas' claim. You should use a 5% level of significance and state your hypotheses clearly.
(7)

---

6. In an experiment some children were asked to estimate the position of the centre of a circle. The random variable D represents the distance, in centimetres, between the child's estimate and the actual position of the centre of the circle. The cumulative distribution function of D is given by

$$F(d) = \begin{cases} 0 & d < 0 \\ \frac{d^2}{2} - \frac{d^4}{16} & 0 \le d \le 2 \\ 1 & d > 2 \end{cases}$$

(*a*) Find the median of D.
(4)

(*b*) Find the mode of D.

Justify your answer.
(5)

The experiment is conducted on 80 children.

(*c*) Find the expected number of children whose estimate is less than 1 cm from the actual centre of the circle.
(3)

PMT

7. A piece of string AB has length 9 cm. The string is cut at random at a point P and the random variable X represents the length of the piece of string AP.

(*a*) Write down the distribution of X.

(1)

(*b*) Find the probability that the length of the piece of string AP is more than 6 cm.

(1)

The two pieces of string AP and PB are used to form two sides of a rectangle.

The random variable R represents the area of the rectangle.

(*c*) Show that $R = aX^2 + bX$ and state the values of the constants a and b.

(2)

(*d*) Find E(R).

(6)

(*e*) Find the probability that R is more than twice the area of a square whose side has the length of the piece of string AP.

(4)

TOTAL FOR PAPER: 75 MARKS

END

Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced /Advanced Subsidiary

Tuesday 24 June 2014 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Pink)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulas stored in them.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 6 questions.
The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

P43172A

1. Patients arrive at a hospital accident and emergency department at random at a rate of 6 per hour.

(*a*) Find the probability that, during any 90 minute period, the number of patients arriving at the hospital accident and emergency department is

(i) exactly 7,

(ii) at least 10.

(5)

A patient arrives at 11.30 a.m.

(*b*) Find the probability that the next patient arrives before 11.45 a.m.

(3)

---

2. The length of time, in minutes, that a customer queues in a Post Office is a random variable, T, with probability density function

$$f(t) = \begin{cases} c(81 - t^2) & 0 \leq t \leq 9 \\ 0 & \text{otherwise} \end{cases}$$

where c is a constant.

(*a*) Show that the value of c is $\frac{1}{486}$.

(4)

(*b*) Show that the cumulative distribution function F(t) is given by

$$F(t) = \begin{cases} 0 & t < 0 \\ \frac{t}{6} - \frac{t^3}{1458} & 0 \leq t \leq 9 \\ 1 & t > 9 \end{cases}$$

(2)

(*c*) Find the probability that a customer will queue for longer than 3 minutes.

(2)

A customer has been queuing for 3 minutes.

(*d*) Find the probability that this customer will be queuing for at least 7 minutes.

(3)

Three customers are selected at random.

(*e*) Find the probability that exactly 2 of them had to queue for longer than 3 minutes.

(3)

---

3. A company claims that it receives emails at a mean rate of 2 every 5 minutes.

(*a*) Give two reasons why a Poisson distribution could be a suitable model for the number of emails received.

(2)

(*b*) Using a 5% level of significance, find the critical region for a two-tailed test of the hypothesis that the mean number of emails received in a 10 minute period is 4. The probability of rejection in each tail should be as close as possible to 0.025.

(2)

(*c*) Find the actual level of significance of this test.

(2)

To test this claim, the number of emails received in a random 10 minute period was recorded.

During this period 8 emails were received.

(*d*) Comment on the company's claim in the light of this value. Justify your answer.

(2)

During a randomly selected 15 minutes of play in the Wimbledon Men's Tennis Tournament final, 2 emails were received by the company.

(*e*) Test, at the 10% level of significance, whether or not the mean rate of emails received by the company during the Wimbledon Men's Tennis Tournament final is lower than the mean rate received at other times. State your hypotheses clearly.

(5)

---

4. A cadet fires shots at a target at distances ranging from 25 m to 90 m. The probability of hitting the target with a single shot is p. When firing from a distance d m, $p = \frac{3}{200}(90 - d)$.

Each shot is fired independently.

The cadet fires 10 shots from a distance of 40 m.

(*a*) (i) Find the probability that exactly 6 shots hit the target.

(ii) Find the probability that at least 8 shots hit the target.

(5)

The cadet fires 20 shots from a distance of x m.

(*b*) Find, to the nearest integer, the value of x if the cadet has an 80% chance of hitting the target at least once.

(4)

The cadet fires 100 shots from 25 m.

(*c*) Using a suitable approximation, estimate the probability that at least 95 of these shots hit the target.

(5)

---

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)

---

6. The continuous random variable X has probability density function $\mathrm{f}(x)$ given by

$$\mathrm{f}(x) = \begin{cases} \dfrac{2x}{9} & 0 \leq x \leq 1 \\ \dfrac{2}{9} & 1 < x < 4 \\ \dfrac{2}{3} - \dfrac{x}{9} & 4 \leq x \leq 6 \\ 0 & \text{otherwise} \end{cases}$$

(*a*) Find $\mathrm{E}(X)$.
(4)

(*b*) Find the cumulative distribution function $\mathrm{F}(x)$ for all values of x.
(6)

(*c*) Find the median of X.
(3)

(*d*) Describe the skewness. Give a reason for your answer.
(2)

TOTAL FOR PAPER: 75 MARKS

END