

## EXAMINATION PAPER

**Examination Session:**  
Sum2

**Year:**  
2021

**Exam Code**  
F\_FSTATS\_J\_EXAM\_B

**Module Title: Foundations of Statistics**

Time Allowed:	It is advised that you spend 1.5 hours to complete this paper. You do not need to complete the examination in one sitting and may take breaks if necessary.	
Additional Material provided:	n/a	
Materials Permitted:	n/a	
Calculators Permitted:	n/a	Models Permitted: n/a
Students permitted to use dictionaries: n/a		

<b>Instructions to Candidates:</b>	<b>SUBMISSION DEADLINE: Weds 4<sup>th</sup> August 1pm</b>	
	<p>You may use textbooks, lecture notes, DUO, and other web educational resources, but the examination attempt must be your own work. We advise you very strongly not to try to do any fresh research during the exam.</p> <p>You should not spend significantly longer than the scheduled duration of the exam in preparing and completing your answers. The work must be all your own and you must not have consulted any other person to assist you in any way when preparing and completing your answers.</p> <p>If you have any problems, please do not contact your subject tutors. You must contact <a href="mailto:DurhamISCExams@studygroup.com">DurhamISCExams@studygroup.com</a></p>	
		Revision:

### **Instructions**

There are 8 questions. Answer **all** questions.

The total marks available for this exam is 60. This accounts for 60% of your overall marks for this module.

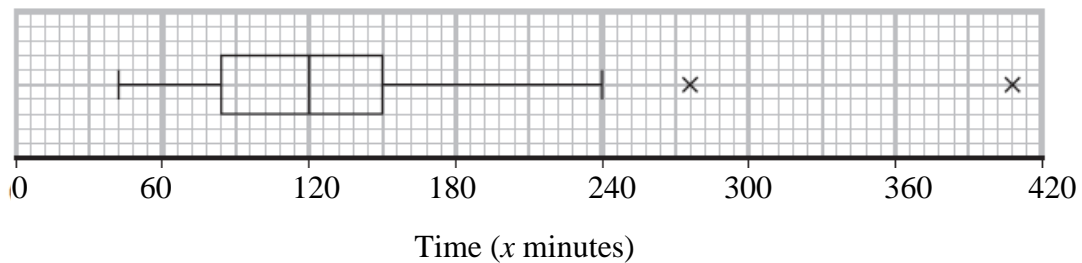
Ensure all working is shown. Answers with no working may gain no credit.

Any necessary formulae are given on the formula sheet provided on DUO.

1. Each member of a group of 35 people went walking on a weekend.

The time taken for the walk,  $x$  minutes, for each member of the group was recorded.

These times are summarised in the following box and whisker plot.



- (a) Find the range of the times. (1)

- (b) Find the interquartile range of the times. (1)

For these 35 people  $\sum x = 3850$  and  $\sum x^2 = 575,960$

- (c) Calculate the mean time taken on the walk, (1)

- (d) Calculate the standard deviation of the times taken on the walk. (2)

Louise defines an outlier as a value more than 3 standard deviations above or below the mean.

- (e) State how many outliers Louise would say there are in these data, giving a reason for your answer. (2)

2. The table gives information about the numbers of students from different types of schools who applied to Cambridge University in 2016:

Type of school	Gender		Total
	Male	Female	
Maintained	3674	2899	6573
Independent	1510	1268	2778
Other and overseas	300	312	612
Total	5484	4479	9963

Richard is going to take a sample of 200 of these students stratified by gender.

- (a) Work out how many female students there should be in this sample. (2)

- (b) Describe a situation when it would not be appropriate to take a sample stratified by gender. (1)

Richard could have used a different category for his stratified sample.

(c) What is this different category? (1)

A student is to be chosen at random from the 9963 students.

F is the event that the student chosen is female.

I is the event that the student chosen is from an independent school.

M is the event that the student chosen is from a maintained school.

(d) Explain why the event F and the event I are not mutually exclusive. (1)

(e) Find the probability of I or M. (2)

(f) Find the probability of F given M. (2)

3. The table shows information about the amount of time that each member of a group of 46 teenagers spent on social media during one day.

Time spent on social media (t minutes)	Frequency
$0 \leq t < 50$	1
$50 \leq t < 100$	4
$100 \leq t < 150$	8
$150 \leq t < 200$	17
$200 \leq t < 300$	16

- (a) (i) Use linear interpolation to find an estimate of the median time spent on social media by the 46 teenagers. (2)

The average person will spend approximately 116 minutes on social media each day. (Source: [www.socialmediatoday.com](http://www.socialmediatoday.com))

- (ii) Compare the amount of time spent on social media by the 46 teenagers with the amount of time spent on social media by the average person. (1)

The table below gives the mean, the standard deviation and the median for the times spent on social media during one day by a sample of sixty-year-olds.

Mean (minutes)	Standard deviation (minutes)	Median (minutes)
125	25	130

(b) Calculate the skew of the times spent on social media by the sample of sixty-year-olds. (2)

(c) Interpret your answer to part (b). (2)

4. A company director wants to introduce a performance-related pay structure for her managers. A random sample of 15 managers is taken and the annual salary,  $y$  in £1000, was recorded for each manager. The director then calculated a performance score,  $x$ , for each of these managers.

The results are summarised in the following statistics.

$$\sum x = 465 \quad \sum y = 562 \quad S_{xx} = 2492 \quad \sum y^2 = 23\,140 \quad \sum xy = 19428$$

(a) (i) Show that  $S_{xy} = 2006$  (1)

(ii) Find  $S_{yy}$

(2)

(b) Find the product moment correlation coefficient between performance score and annual salary. (2)

(c) Describe the correlation between performance score and annual salary. (2)

The director believes that there is a linear relationship between performance score and annual salary.

(d) Calculate the equation of the regression line of  $y$  on  $x$ , in the form  $y = a + bx$ . Give the value of  $a$  and the value of  $b$  to 3 significant figures. (4)

(e) Give an interpretation of the value of  $b$ . (1)



5. A geography student is investigating the relationship between the size of a shopping centre (measured by the number of shops it contains) and the mean distance travelled by shoppers to reach the shopping centre.

She obtains the following data:

Centre	A	B	C	D	E	F
No. of shops	6	20	15	30	30	60
Mean distance (km)	0.5	2.1	2.4	4.1	3.9	6.2

Calculate Spearman's rank correlation coefficient between mean distance travelled and number of shops

[7]

6. Fruitees sweets come in different flavours.

There are 8 sweets in a pack of mixed flavours and the flavours for each pack are chosen at random

The mean number of strawberry flavour Fruitees in a pack of 8 sweets is 2

Ed suggests that the number of strawberry flavour Fruitees in a pack of 8 sweets can be modelled by a binomial distribution.

- (a) By considering the conditions that make a binomial distribution a suitable model, explain why Ed's suggestion is appropriate. (2)

One sweet is selected at random from a pack of Fruitees.

- (b) Find the probability that the flavour of this sweet is strawberry. (1)

Ed buys a pack of Fruitees.

- (c) Find the probability that there will be exactly 3 strawberry flavour Fruitees in the pack. (2)

7. A machine puts crisps into packets. The weight of crisps in each packet,  $M$  grams, follows a normal distribution with mean 40 g and standard deviation 2.38g

(a) Find, to 2 decimal places, the value of  $k$  such that  $P(k < A < 40) = 0.40$  (4)

Fifteen packets of crisps are selected at random.

(b) Given that 20% of packets contain more than 42 g of crisps; find the probability that fewer than 3 of these packets contain more than 42 g of crisps. (2)

8. A health centre claims that the time a dentist spends with a patient can be modelled by a normal distribution with a mean of 25 minutes and a standard deviation of 12 minutes.

(a) Using this model, find the probability that the time spent with a randomly selected patient is more than 40 minutes.

(3)

Some patients complain that the mean time the dentist spends with a patient is more than 25 minutes.

The receptionist takes a random sample of 30 patients and finds that the mean time the dentist spends with a patient is 28 minutes.

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

(4)