A Level Statistics
AQA Past Exam Questions

TOPIC: Sampling Methods

Candidates may use any calculator allowed by Pearson regulations. Calculators must not 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 ***on paper***

• You should show sufficient working to make your methods clear. Answers without working may not gain full credit.

• Unless otherwise stated, statistical tests should be carried out at the 5% significance level.

• When a calculator is used, the answer should be given to three significant figures unless otherwise stated.

Information

• ***You may use the*** booklet 'Statistical Formulae and Tables'

• There are **8** questions in this question paper. The total mark for this paper is ***165***

• 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.

• If you change your mind about an answer, cross it out and put your new answer and any working underneath.

• Check your answers if you have time at the end.

AQA_JUNE_2015_6

The Goodwell Medical Practice has a total of 3200 registered patients.

The local health authority has asked this practice to complete questionnaires about a sample of 40 patients, but has not said how this sample should be chosen.

The practice has available a list of patients ordered alphabetically by their family name. The patients are numbered in the list from 1 to 3200. The details of 20 patients are printed on each page of the list.

(a) Dr Dobry suggests picking two pages at random from the list and using the patients on those pages as the sample.

(i) Name this method of sampling.

(ii) Give a reason why Dr Dobry's method would be unlikely to give a representative sample.

[2 marks]

(b) Dr Kalos suggests obtaining the sample of 40 from the list by a systematic method. Explain briefly how this could be done.

[2 marks]

(c) Dr Bueno says the sample should be chosen at random in such a way as to proportionately represent the age distribution of the patients.

(i) Name this method of sampling.

(ii) Given that there are 737 registered patients aged over 60 years, how many of these should be chosen in a sample of 40 chosen using Dr Bueno's method?

[3 marks]

(d) Dr Mabuti suggests obtaining a simple random sample by the following method.

* Obtain a four-digit random number from tables, rejecting any number above 9599 .
* Divide the number by 3200 and find the remainder.
* Add 1 to this remainder.
* Select the patient in the list corresponding to this number.
* Carry out this procedure 40 times.

(i) Using this method, which number of the patient in the alphabetical list would be generated by the random number 5817 ?

(ii) Explain why it is necessary to reject any random number above 9599 .

(iii) Explain why it is necessary to add 1 .

(iv) Dr Mabuti has omitted one instruction which is needed to make sure this method gives a simple random sample of patients. What instruction needs to be added?

[5 marks]

AQA_JUNE_2012_4a

An exam board was developing a new examination for Year 11 pupils. It was intended that the mean mark obtained in the examination should be 50. A specimen examination paper was produced. It was suggested that this examination paper should be taken by 80 pupils. Four Year 11 classes would be chosen, one class from each of four schools, and from each of these classes, 20 pupils would be chosen to take the examination paper.

(a) (i) Name this type of sampling.

(1 mark)

(ii) Explain why data obtained by this sampling method would not give a reliable result for a test of the hypothesis that the mean mark of all Year 11 pupils on this examination was 50.

(1 mark)

AQA_JAN_2012_6

Adam owns a farm. On the farm, there are 330 sheep, 160 cattle and 10 goats.

The animals are each tagged with a number which is recorded in a stock book. The sheep are numbered from 001 to 330, the cattle are numbered from 331 to 490 and the goats are numbered from 491 to 500. Adam wants to carry out a blood test to check the health of his animals but can only afford to carry out the test on 50 animals.

(a) (i) Describe how Adam could use random numbers to select a simple random sample of 50 animals for the test.

(4 marks)

(ii) Explain why the small number of goats might cause a problem when using a random sample.

(1 mark)

(b) Adam's farm manager, Ellie, suggests a different sampling method. She suggests randomly choosing a number from 1 to 10, testing the animal corresponding to that number in the stock book, and then testing every tenth animal after that. For example, if 7 were the randomly chosen number, then animals 007, 017, 027 and so on would be tested.

(i) Name the sampling method suggested by Ellie.

(1 mark)

(ii) State, with a reason, whether a sample chosen in this way is a random sample.

(2 marks)

(c) The animals are kept in many different fields. Adam decides to carry out a quota sample.

(i) Describe how Adam might carry out a quota sample of size 50.

(2 marks)

(ii) State one advantage to Adam of carrying out a quota sample.

(1 mark)

(iii) State one disadvantage of carrying out a quota sample.

(1 mark)

AQA_JAN_2013_6

The managers of a fast-food chain with several hundred outlets nationwide have hired a small research company based in Manchester to carry out some customer-satisfaction interviews. They categorise their customers into four distinct groups: pensioners, other adults, teenagers and children.

They believe that these four groups are in roughly equal proportions. They wish to know the opinions of a sample of 50 pensioners, 50 other adults, 50 teenagers and 50 children.

(a) Name this type of sample.

(1 mark)

(b) Give three reasons why simple random sampling from all the outlets' customers would be impractical.

(3 marks)

(c) The research company suggests that quota sampling of customers from outlets in Manchester should be used.

(i) Give an advantage to the research company of using this quota sampling.

(ii) Give a reason why the managers might think that this quota sampling is not suitable.

(2 marks)

(d) Describe how the research company might aim to obtain a representative sample by the use of cluster sampling.

(3 marks)

AQA_JUNE_2014_5

Bronwyn wants to conduct a survey of a sample of 50 pupils in her school in order to collect information about the ways they travel to school. She plans to use a stratified sample to reflect the proportions of boys and girls and also the number of pupils in each year.

The data for the whole school are given in Table 1.

Table 1

	Year 7	Year 8	Year 9	Year 10	Year 11
Boys	66	71	75	84	87
Girls	61	72	75	80	79

(a) Calculate the numbers of boys and girls from each year that Bronwyn should have in her sample and enter them in Table 2, below.

[4 marks]

(b) Bronwyn plans to use systematic sampling to select her sample of pupils from each year. She can produce, on the school computer system, a list of the pupils by year which also identifies them as boy or girl. For Year 9:

(i) state how the list should be arranged so that the systematic sample will be stratified for gender;

(ii) describe in detail how Bronwyn selects the pupils to be in her sample.

[4 marks]

(c) Bronwyn might have chosen to select her sample using quota sampling of pupils as they arrive at school. Give one advantage and one disadvantage of quota sampling in the context of this question.

[2 marks]

AQA_JUNE_2017_7ab

A gym has 800 members, of whom 460 are men and 340 are women. Half of the men are under 30 years old and 55 per cent of the women are under 30 years old. The manager, James, wants to survey a sample of 80 gym members to find out what developments they think the gym should undertake. James considers three options for how he might obtain his sample.

(a) In Option 1, James will calculate, for each combination of age and sex, the number for his sample that is in proportion to the corresponding age and sex for gym members. He will then select gym members at random, until he has the required numbers for his survey.

(i) Name this sampling method.

[1 mark]

(ii) For each combination of age and sex, calculate the number of gym members that James should have in his sample.

[3 marks]

(b) In Option 2, James will wait in the gym cafeteria and select groups of gym members there until 80 have been selected.

(i) Name this sampling method.

[1 mark]

(ii) State one advantage and one disadvantage of Option 2 compared with Option 1.

[2 marks]

AQA_JUNE_2018_7

Padraig is a student investigating the effects of running, on people who do not exercise regularly but are taking part in a fun run organised by 12 charities. He plans to select a sample of 40 runners out of 2000 taking part. The organisers have given Padraig a list of the runners, numbered from 0001 to 2000, and stating the charity supported by each runner.

Padraig plans to measure the pulse rate of each person in his sample at the end of the run. He wants to investigate how the pulse rate is associated with the sex and age of the runner.

(a) Explain why the list given to Padraig is insufficient to allow him to select a sample stratified to help his investigation.

[2 marks]

(b) (i) Explain how Padraig could use the numbers on the list to select a simple random sample.

(ii) State, in this context, one practical disadvantage of using this sampling method to collect his data.

[4 marks]

(c) Padraig considers selecting his sample of 40 systematically from the runners in the order that they finish the run.

(i) Explain in detail how Padraig would do this.

(ii) State, in this context, one advantage and one disadvantage of using this sampling method to collect his data, compared to using a random sample.

[4 marks]

(d) Runners from each charity taking part will wear T-shirts with the name of their charity clearly printed on the front and back. Padraig decides to obtain his sample by selecting 4 charities at random and then selecting 10 runners who finish the run from each of those charities.

(i) Name this method of sampling.

(ii) State, in this context, one advantage and one disadvantage of using this sampling method to collect his data, compared to using a sample selected systematically.

[3 marks]

AQA_JAN_2007_7

A bus company has 950 employees who are divided into four employment categories as follows:
620 drivers;
120 mechanics;
130 clerical staff;
and 80 managers.

For each category, a list of the names of the employees is available. The company hopes to improve morale among the employees by providing better and healthier meals in its canteen. As a first step, it plans to distribute a questionnaire to a sample of employees on the meals currently provided in the canteen.

(a) Describe how random numbers could be used to select a sample of 95 employees stratified by employment category.

(6 marks)

(b) Describe how a systematic sample of 95 employees could be selected.

(3 marks)

(c) Data from a similar questionnaire, used at another company, had been analysed. The results suggested that there was little difference in the opinions on canteen meals between the different employment categories but that there was a difference in the opinions on canteen meals between males and females. The bus company is to decide whether to use the stratified sample from part (a), a differently stratified sample or the systematic sample from part (b). Make three points to be considered before this decision is made.

(3 marks)

A university employs 820 staff at a city centre site. Natasha, an administrator in charge of car parking, wishes to survey members of staff as to their views on the present parking arrangements. A list of the 820 members of staff is available.

(a) Describe how Natasha could select a simple random sample of size 25 for the survey.

(4 marks)

(b) The university has a car park with 200 parking spaces. The spaces are numbered from 1 to 200. There are 300 members of staff with permits to park in this car park and each morning the spaces are allocated on a first come, first served basis. Permit holders arriving when the car park is full have to park elsewhere. There are 220 members of staff on a waiting list for a permit and the remaining 300 members of staff do not wish to have a permit.

(i) It is suggested that, instead of a simple random sample, Natasha should take a stratified sample. Suggest two relevant factors that Natasha could use in the stratification of the staff.

(2 marks)

(ii) A second suggestion is as follows: Select a systematic sample of 25 parking spaces. Natasha and her assistant will wait in the car park in the morning and when a car parks in one of the selected spaces the driver will be given a questionnaire and asked to complete it.

(A) Describe how the systematic sample of 25 parking spaces could be selected.

(B) State one advantage of this method of obtaining completed questionnaires.

(C) State two sources of bias in this method of data collection.

(5 marks)

AQA_JAN_2010_6

An organisation employs a large number of examiners to mark scripts. The organisation wishes to consult a sample of examiners on various matters, including whether they would prefer their meetings to be held in London or Manchester.

The following table summarises the locations of those examiners who live in England.

Region	**Number of examiners**
East Anglia	34
East Midlands	46
North	54
North West	21
South East	98
South West	73
West Midlands	62
Yorkshire and Humberside	12

(a) Describe how a simple random sample of size 40 could be selected from these examiners.

(5 marks)

(b) David proposes that four regions should be selected at random and then ten examiners selected at random from each of these regions. A visit would be arranged to each of the chosen examiners.

(i) Name this method of sampling.

(1 mark)

(ii) Give one advantage of this method of sampling.

(1 mark)

(iii) Give, in context, one disadvantage of this method of sampling.

(2 marks)

(iv) Would each examiner have an equal chance of being included in this sample? Justify your answer.

(2 marks)

(c) It is pointed out to David that examiners could be asked for their views by telephone or by e-mail. Explain whether or not your answer to: (i) part (b)(ii) remains valid; (ii) part (b)(iii) remains valid.

(2 marks)

AQA_JAN_2011_5

A hospital has 820 employees in 4 categories. The table shows the number of employees in each category.

Category	Number of employees
Medical	390
Ancillary	220
Administrative	140
Managerial	70

For each category, a list of employees is available. The hospital is seeking to improve its standard of hygiene and, as a first step, it plans to ask a sample of employees to answer a questionnaire.

(a) Describe how random numbers could be used to select a stratified sample of size 82 in such a way that all employees have an equal chance of being selected.

(6 marks)

(b) (i) Describe how a systematic sample of size 100 could be selected.

(ii) Explain whether or not all employees have an equal chance of being selected in the systematic sample that you described in part (b)(i).

(5 marks)

(c) Data from a similar questionnaire, used at a nearby hospital, found that there was little difference in the opinions of the different categories but there was a difference between the opinions of males and females. The hospital is to decide whether to select a stratified sample using the categories in the table, a differently stratified sample or a systematic sample. Make three points which may help the hospital in making this decision.

(3 marks)

AQA_JUNE_2009_6

John organises the house-to-house delivery of free newspapers in a small town. He divides the town into areas and employs students to deliver the newspapers to the houses in each area. After a delivery, John calls on eight houses in each area to check that the free newspaper has been delivered.

A particular area consists of:
North Street, which contains 63 houses;
East Street, which contains 77 houses;
South Street, which contains 46 houses;
West Street, which contains 94 houses.

(a) Describe how John could select a simple random sample of eight houses from this area.

(5 marks)

(b) Describe how John could select a systematic sample of eight houses from this area.

(3 marks)

(c) John decides to select two of the four streets in this area at random and to call at four houses selected randomly from each of these two streets. Name this type of sampling.

(1 mark)

(d) Socrates is a student who is sometimes employed by John to deliver the free newspapers in this area. If Socrates does not have time to deliver to all the houses, he misses out one whole street.

(i) Explain why John should not use the type of sampling described in part (c) when checking that Socrates has delivered to all houses in this area.

(ii) State, giving a reason, which, if either, of a simple random sample or a systematic sample would be preferable when checking that Socrates has delivered to all houses in this area.

(3 marks)

(e) Mary is another student who is sometimes employed by John to deliver the free newspapers in this area. If Mary does not have time to deliver to all the houses, she misses out houses haphazardly. State, giving a reason, which, if either, of a simple random sample or a systematic sample would be preferable when checking that Mary has delivered to all houses in this area.

(2 marks)

AQA_JUNE_2006_5

A company, with 9320 employees, provides refuse collection services for 47 councils in the United Kingdom. The company asks a market research firm to carry out an opinion poll of its employees concerning union membership.

(a) Describe how the market research firm could obtain a simple random sample of size 120 from the 9320 employees.

(4 marks)

(b) The market research firm selects 4 of the 47 councils at random.

(i) What further step(s) would be necessary to obtain a cluster sample of size 120 from the 9320 employees?

(2 marks)

(ii) Give a reason why the market research firm might prefer a cluster sample to a random sample.

(2 marks)

(c) It is proposed that a stratified sample be used.

(i) Suggest two factors which could be used to stratify the sample.

(2 marks)

(ii) Suggest a reason why a stratified sample might be preferred to a cluster sample.

(1 mark)

AQA_JUNE_2008_5

A train travelling between two major cities has ten carriages. Four of the carriages are for first-class passengers only and each contains 48 seats, numbered from 1 to 48. Six of the carriages are for standard-class passengers only and each contains 72 seats, numbered from 1 to 72.

The railway company wishes to survey passengers with a view to making the catering facilities more profitable. Interviews are to be carried out during the middle section of the journey when it can be assumed that all seats will be occupied. For practical reasons, only seated passengers will be interviewed.

It is planned that between 25 and 30 interviews in total will be carried out.

Owen suggests that 4 carriages should be selected at random and that 7 passengers selected at random from those seated in each of these carriages should be interviewed.

Xavier suggests that 18 passengers selected at random from the seated standard-class passengers and 8 passengers selected at random from the seated first-class passengers should be interviewed.

Jada suggests that three numbers between 1 and 48 should be selected at random and that the passengers in these seats in each carriage should be interviewed.

(a) (i) Name the type of sampling suggested by Owen.

(ii) Describe how random numbers could be used to select 7 passengers at random from the 72 passengers seated in a standard-class carriage.

(5 marks)

(b) (i) Name the type of sampling suggested by Xavier.

(ii) Suggest reasons why Xavier chose the numbers 18 and 8 for his sample sizes.

(4 marks)

(c) For Jada's method of sampling, state, giving a reason, whether or not:

(i) all seated passengers have an equal chance of being included in the sample;

(ii) all seated first-class passengers have an equal chance of being included in the sample.

(3 marks)

(d) The railway company requires an estimate of the mean income of all passengers on the train. State, giving a reason, whether the mean income of a random sample of size 26 from all seated passengers or the mean income of the sample suggested by Xavier would be a preferable estimate.

(2 marks)

AQA_JUNE_2010_7

A cooperative society owns 420 shops. The society decides to employ a team of energy efficiency experts to carry out audits of the electricity consumed by these shops and to recommend ways in which this could be reduced. The team will be able to audit about 40 shops during the next three months.

(a) Describe how random numbers could be used to select a simple random sample of 40 shops to be audited.

(4 marks)

(b) The energy efficiency experts arrange the shops in order of electricity consumption and number them from 000 to 419. The shop numbered 000 consumes the least electricity and the shop numbered 419 consumes the most.

The following three suggestions are made for selecting the shops for audit.

Suggestion P

The sample is made up of random samples of 10 shops from those numbered 000 to 099, 10 shops from those numbered 100 to 199, 10 shops from those numbered 200 to 299 and 12 shops from those numbered 300 to 419.

Suggestion Q

Select a single digit at random and choose this shop and every tenth shop thereafter. For example, if 002 is selected, shops numbered 002, 012, 022,… and 412 would constitute the sample.

Suggestion R

Shops numbered 378 to 419 are selected.

(i) For Suggestion P:

(A) name the method of sampling;

(1 mark)

(B) state whether or not each shop is equally likely to be selected;

(1 mark)

(C) give a reason why the sample is not a random sample.

(1 mark)

(ii) For Suggestion Q:

(A) name the method of sampling; (1 mark) (B) state whether or not each shop is equally likely to be selected;

(1 mark)

(C) give a reason why the sample is not a random sample.

(1 mark)

(iii) Give a reason why Suggestion R might be preferred to both Suggestion P and Suggestion Q.

(2 marks)