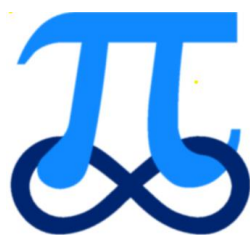


Edexcel

Statistics 3

Sampling Methods



Edited by: K V Kumaran

2 Sampling

Methods of collecting data

Taking a census

A **census** involves observing **every** member of a population and is used if
the size of the population is small
or if extreme accuracy is required.

Advantages

it should give a completely accurate result, a full picture.

Disadvantages

very time consuming and expensive
it cannot be used when testing process destroys article being tested
information is difficult to process because there is so much of it.

Sampling

Sampling involves observing or testing a part of the population. It is cheaper but does not give such a full picture. The **size** of the sample depends on the accuracy desired (for a varied population a large sample will be required to give a reasonable accuracy).

Simple random sampling

Every member of the population must have an **equal chance** of being selected.

Using random number tables

To take a simple random sample of size n from a population of N sampling units first make a list and give each member of the population a number. Then use random number tables to select the sample.

We ignore any numbers which do not refer to a member of the population – for example using three figure random numbers for a population numbered from 001 to 659 we would ignore numbers from 660 to 999.

Also (for sampling without replacement) we ignore the second occurrence of the same number.

Advantages

the numbers are truly random and free from bias
it is easy to use
each member has a known equal chance of selection

Disadvantages

it is **not** suitable when the sample size is large.

Lottery sampling

A sampling frame is needed – identifying each member of the population. The name or number of each member is written on a ticket (all the same size, colour and shape), and the tickets are all put in a container which is then shaken. Tickets are then drawn **without** replacement.

Advantages

the tickets are drawn at random.
it is easy to use.
each ticket has a known chance of selection (considered as constant as long as the sample size is much smaller than the total number of tickets).

Disadvantages

it is not suitable for a large sample
a sampling frame is needed.

Systematic sampling

First make an ordered list.

Second select every 50th (or ??) member from the list.

In order to make sure that the first on the list is not automatically selected random numbers could be used to select the first member then select every 50th (or ??) after that.

Used when the population is too large for simple random number sampling.

Advantages

- simple to use
- suitable for large samples

Disadvantages

- only random if the ordered list is truly random.
- it can introduce bias

Quota sampling

This is a non-random method.

First decide on groups into which the population is divided and a number from each group to be interviewed to form *quotas*.

Then go out and interview and enter each result into the relevant quota.

If someone refuses to answer or belongs to a quota which is already full then ignore that persons reply and continue interviewing until **all** quotas are full.

Used when it is not possible to use random methods - for example when the whole population is not known (homeless in a big city).

Advantages

- can be done quickly as a representative sample can be obtained with a small sample size
- costs are kept to a minimum
- administration is fairly easy.

Disadvantages

- it is not possible to estimate the sampling errors (as it is not a random process)
- interviewer may not put into correct quota
- non-responses are not recorded
- it can introduce interviewer bias

Sampling with and without replacement

Simple random sampling is sampling *without* replacement in which each member of population can be selected at most *once*.

In sampling *with* replacement each member of the population can be selected more than once: this is called **unrestricted random sampling**.

Stratified sampling

First divide the population into **exclusive (distinct)** groups or *strata* and then select a sample so that the proportion of each stratum in the sample equals the proportion of that stratum in the population.

Example: How would you take a stratified sample of 50 children from a school of 500 pupils divided as follows:

	Boys	Girls
Upper sixth	30	40
Lower sixth	30	30
Fifth form	70	60
Fourth form	60	70
Third form	50	60

Solution: As 50 is $\frac{1}{10}$ of the total population, $\frac{1}{10}$ of each stratum should be selected in the sample. Thus the sample would comprise

	Boys	Girls
Upper sixth	3	4
Lower sixth	3	3
Fifth form	7	6
Fourth form	6	7
Third form	5	6

and simple random number sampling would be used within each stratum.

Used when

the sample is large
the population divides naturally into mutually exclusive groups.

Advantages

it can give more accurate estimates (or a more representative picture) than simple random number sampling when there are *clear strata* present.
It reflects the population structure.

Disadvantages

within the strata the problems are the same as for any simple random sample
if the strata are not clearly defined they may overlap.

Primary data

Primary data is data collected by or on behalf of the person who is going to use the data.

Advantages

- collection method is known
- accuracy is known
- exact data needed are collected

Disadvantages

- costly in time and effort

Secondary data

Secondary data is data **not** collected by or on behalf of the person who is going to use it. The data are second-hand – e.g. government census statistics.

Advantages

- cheap to obtain
- large quantity available (e.g. internet)
- much has been collected year on year and can be used to plot trends

Disadvantages

- collection method may not be known
- accuracy may not be known
- it can be in a form which is difficult to handle
- bias is not always recognised.

1. A hotel has 160 rooms of which 20 are classified as De-luxe, 40 Premier and 100 as Standard. The manager wants to obtain information about room usage in the hotel by taking a 10% sample of the rooms.

(a) Suggest a suitable sampling method.

(1)

(b) Explain in detail how the manager should obtain the sample.

(4)

Q1, May 2002

2. Explain how to obtain a sample from a population using

(a) stratified sampling,

(2)

(b) quota sampling.

(2)

Give one advantage and one disadvantage of each sampling method.

(4)

Q1, June 2003

3. There are 64 girls and 56 boys in a school.

Explain briefly how you could take a random sample of 15 pupils using

(a) a simple random sample,

(3)

(b) a stratified sample.

(3)

Q1, June 2004

4. (a) State two reasons why stratified sampling might be chosen as a method of sampling when carrying out a statistical survey.

(2)

(b) State one advantage and one disadvantage of quota sampling.

(2)

Q1, June 2005

5. A school has 15 classes and a sixth form. In each class there are 30 students. In the sixth form there are 150 students. There are equal numbers of boys and girls in each class. There are equal numbers of boys and girls in the sixth form. The head teacher wishes to obtain the opinions of the students about school uniforms.

Explain how the head teacher would take a stratified sample of size 40.

(7)

Q1, Jan 2006

6. Describe one advantage and one disadvantage of

(a) quota sampling,

(2)

(b) simple random sampling.

(2)

Q1, June 2006

7. A researcher is hired by a cleaning company to survey the opinions of employees on a proposed pension scheme. The company employs 55 managers and 495 cleaners.

To collect data the researcher decides to give a questionnaire to the first 50 cleaners to leave at the end of the day.

(a) Give 2 reasons why this method is likely to produce biased results.

(2)

(b) Explain briefly how the researcher could select a sample of 50 employees using

(i) a systematic sample,

(ii) a stratified sample.

(6)

Using the random number tables in the formulae book, and starting with the top left hand corner (8) and working across, 50 random numbers between 1 and 550 inclusive were selected. The first two suitable numbers are 384 and 100.

(c) Find the next two suitable numbers.

(2)

Q5, June 2008

8. A telephone directory contains 50 000 names. A researcher wishes to select a systematic sample of 100 names from the directory.

(a) Explain in detail how the researcher should obtain such a sample.

(2)

(b) Give one advantage and one disadvantage of

(i) quota sampling,

(ii) systematic sampling.

(4)

Q1, June 2009

9. A lake contains 3 species of fish. There are estimated to be 1400 trout, 600 bass and 450 pike in the lake. A survey of the health of the fish in the lake is carried out and a sample of 30 fish is chosen.

(a) Give a reason why stratified random sampling cannot be used.

(1)

(b) State an appropriate sampling method for the survey.

(1)

(c) Give one advantage and one disadvantage of this sampling method.

(2)

(d) Explain how this sampling method could be used to select the sample of 30 fish. You must show your working.

(4)

Q2, May 2012

- 10.** A college manager wants to survey students' opinions of enrichment activities. She decides to survey the students on the courses summarised in the table below.

Course	Number of students enrolled
Leisure and Sport	420
Information Technology	337
Health and Social Care	200
Media Studies	43

Each student takes only one course.

The manager has access to the college's information system that holds full details of each of the enrolled students including name, address, telephone number and their course of study. She wants to compare the opinions of students on each course and has a generous budget to pay for the cost of the survey.

(a) Give one advantage and one disadvantage of carrying out this survey using

(i) quota sampling,

(ii) stratified sampling.

(2)

The manager decides to take a stratified sample of 100 students.

(b) Calculate the number of students to be sampled from each course.

(3)

(c) Describe how to choose students for the stratified sample.

(2)

Q3, May 2013

11. A gym club has 400 members of which 300 are males.

Explain clearly how a stratified sample of size 60 could be taken.

(3)

Q1, May 2013_R

12. (a) Explain what you understand by a random sample from a finite population.

(1)

(b) Give an example of a situation when it is not possible to take a random sample.

(1)

A college lecturer specialising in shoe design wants to change the way in which she organises practical work.

She decides to gather ideas from her 75 students.

She plans to give a questionnaire to a random sample of 8 of these students.

(c) (i) Describe the sampling frame that she should use.

(ii) Explain in detail how she should use a table of random numbers to obtain her sample.

(3)

Q1, May 2014

13. (a) State two reasons why stratified sampling might be a more suitable sampling method than simple random sampling.

(2)

(b) State two reasons why stratified sampling might be a more suitable sampling method than quota sampling.

(2)

Q1, May 2016

- 14.** A company director decides to survey staff about changes to the company calendar. The company has staff in 4 different job roles

72 managers, 108 drivers, 180 administrators and 360 warehouse staff.

The director decides to take a stratified sample.

- (a) Write down one advantage of using a stratified sample rather than a simple random sample for this survey. (1)
- (b) Find the number of staff in each job role that will be included in a stratified sample of 40 staff. (3)
- (c) Describe how to choose managers for the stratified sample. (2)

Q1, May 2017

- 15.** (a) Explain what you understand by a random sample from a finite population. (1)
- (b) Give an example of a situation when it is not possible to take a random sample. (1)

A college lecturer specialising in shoe design wants to change the way in which she organises practical work.

She decides to gather ideas from her 75 students.

She plans to give a questionnaire to a random sample of 8 of these students.

- (c) (i) Describe the sampling frame that she should use.
- (ii) Explain in detail how she should use a table of random numbers to obtain her sample. (3)

Q1, IAL May 2014

- 16.** The names of the 720 members of a swimming club are listed alphabetically in the club's membership book. The chairman of the swimming club wishes to select a systematic sample of 40 names. The names are numbered from 001 to 720 and a number between 001 and w is selected at random. The corresponding name and every x th name thereafter are included in the sample.

- (a) Find the value of w . (1)
- (b) Find the value of x . (1)
- (c) Write down the probability that the sample includes both the first name and the second name in the club's membership book. (1)
- (d) State one advantage and one disadvantage of systematic sampling in this case. (2)

Q1, IAL May 2015

17. A company wants to survey its employees' attitudes to work. The company's workforce is located at three offices. The number of employees at each location is summarised in the table below.

Office location	Number of employees
Bristol	856
Dudley	429
Glasgow	1215

Each employee is located at only one office.

A personnel assistant plans to survey the first 50 employees who arrive for work at the Bristol office on a Monday morning.

- (a) Give two reasons why this survey is likely to lead to a biased response. (2)

A personnel manager has access to the company's information system that holds details of each employee including their place of work.

The manager decides to take a stratified sample of 150 employees.

- (b) Describe how to choose employees for this stratified sample. (3)
- (c) Explain an advantage of using a stratified sample rather than a quota sample. (1)

Q3, IAL May 2016