Stats1 Chapter 1: Data Collection

1.3 Non-Random Sampling

Non-Random Sampling

Consider the following scenario: You wish to conduct a survey in the UK **on whether being left-handed affects IQ**. We need to choose people to assess.

Why would random sampling be problematic?
Because <u>we don't know the sampling frame</u>, i.e. don't have a list of all left-handed (and non-left-handed) people in the UK.

Famous Lefties




For this scenario we'd likely use **quota sampling**, i.e.

1. As with stratified sampling, divide population into groups according to characteristic of interest, then determine size of each group in sample to reflect proportions within the population.
2. But instead of random sampling within each group, we actively choose people within each group via suitable means (e.g. advertising), **until the 'quota' for each group is filled**.

A variant of this is **<u>opportunity sampling</u>**, where we find people **<u>at the same time the survey is being carried out</u>** (e.g. exit polls at polling stations). This is not a suitable method for the left-handed example, because giving the likely time-consuming nature of assessment coupled with resources required, we'd likely arrange with the people taking part before the actual assessment tasks took place.

Quota & Opportunity Sampling

Type	How to carry out	Advantages	Disadvantages
Quota Sampling	**What is it :** Population divided into groups according to characteristic. A quota of items/people in each group is set to try and reflect the group's proportion in the whole population. Interviewer selects the actual sampling units.	• Allows small sample to still be representative of population. • No sampling frame required. • Quick, easy, inexpensive. • Allows for easy comparison between different groups in population.	• Non-random sampling can introduce bias. • Population must be divided into groups, which can be costly or inaccurate. • Increasing scope of study increases number of groups, adding time/expense. • Non-responses are not recorded.
Opportunity/ Convenience Sampling	Sample taken from people who are available at time of study, who meet criteria.	• Easy to carry out. • Inexpensive.	• Unlikely to provide a representative sample. • Highly dependent on individual researcher.

Example Question

Edexcel S3 June 2010 Q2

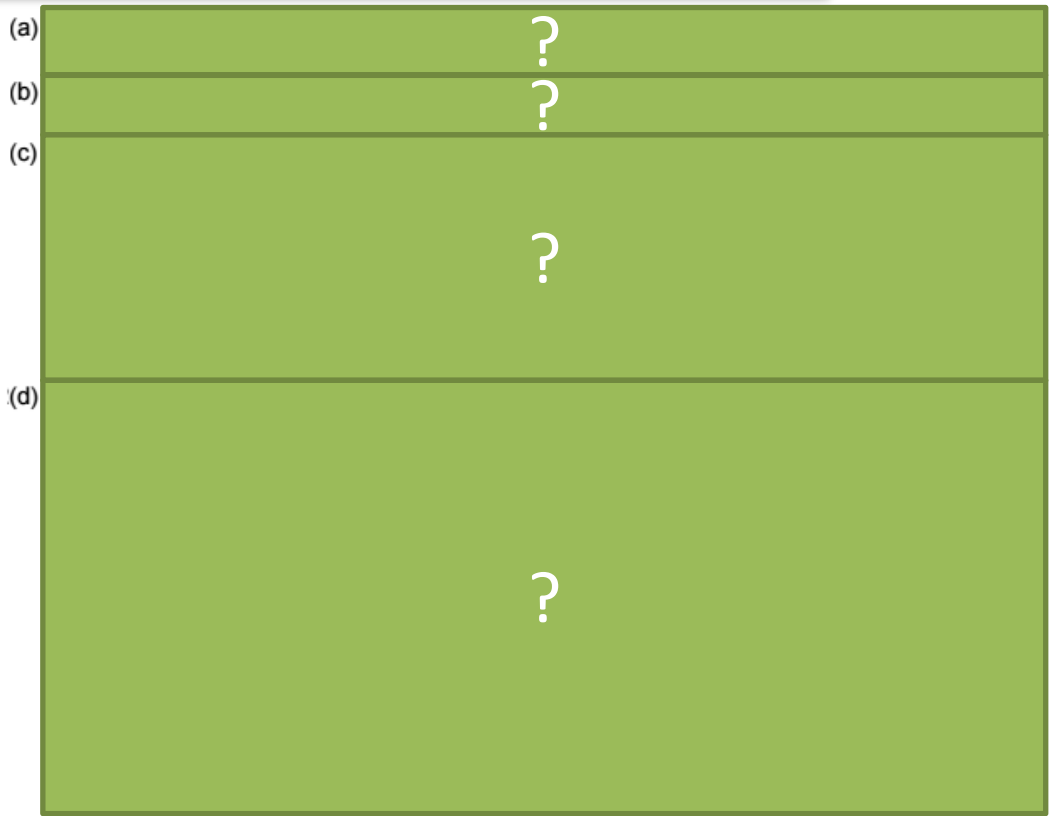
A lake contains 3 species of fish. There are estimated to be 1400 trout, 600 bass and 450 pike in the lake. A survey of the health of the fish in the lake is carried out and a sample of 30 fish is chosen.

(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)**



Example Question

Edexcel S3 June 2010 Q2

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)**

(a) Sampling frame within each species of fish in the lake impossible to obtain. B1 **(1)**

(b) Quota sampling B1 **(1)**

(c) Advantages:
Sample can be obtained quickly
Costs are kept to a minimum B1
Administration of survey is easy
Disadvantages:
Not possible to estimate sampling errors B1
Process not random
Surveyor may not be able to identify species of fish easily **(2)**

(d)

Species	Quota
Trout	$\frac{1400}{2450}\times 30 = 17.14$
Bass	$\frac{600}{2450}\times 30 = 7.35$
Pike	$\frac{450}{2450}\times 30 = 5.51$

Fish are caught from the lake until the quota of 17 trout, 7 bass and 6 pike are reached. B1B1B1

If a fish is caught and the species quota is full, then this is ignored. B1 **(4)**

Exercise 1.3

Pearson Statistics & Mechanics Year 1/AS
Pages 3-4

Homework Exercise

1 Interviewers in a shopping centre collect information on the spending habits from a total of 40 shoppers.

a Explain how they could collect the information using:
i quota sampling **ii** opportunity sampling

b Which method is likely to lead to a more representative sample?

2 Describe the similarities and differences between quota sampling and stratified random sampling.

3 An interviewer asks the first 50 people he sees outside a fish and chip shop on a Friday evening about their eating habits.

a What type of sampling method did he use?

b Explain why the sampling method may not be representative.

c Suggest two improvements he could make to his data collection technique.

4 A researcher is collecting data on the radio-listening habits of people in a local town. She asks the first 5 people she sees on Monday morning entering a supermarket. The number of hours per week each person listens is given below:

4 7 6 8 2

a Use the sample data to work out a prediction for the average number of hours listened per week for the town as a whole.

b Describe the sampling method used and comment on the reliability of the data.

c Suggest two improvements to the method used.

Homework Exercise

5 In a research study on the masses of wild deer in a particular habitat, scientists catch the first 5 male deer they find and the first 5 female deer they find.

a What type of sampling method are they using?

b Give one advantage of this method.

The masses of the sampled deer are listed below.

Male (kg)	75	80	90	85	82
Female (kg)	67	72	75	68	65

c Use the sample data to compare the masses of male and female wild deer.

d Suggest two improvements the scientists could make to the sampling method.

6 The heights, in metres, of 20 ostriches are listed below:

1.8, 1.9, 2.3, 1.7, 2.1, 2.0, 2.5, 2.7, 2.5, 2.6, 2.3, 2.2, 2.4, 2.3, 2.2, 2.5, 1.9, 2.0, 2.2, 2.5

a Take an opportunity sample of size five from the data.

b Starting from the second data value, take a systematic sample of size five from the data.

c Calculate the mean height for each sample.

d State, with reasons, which sampling method is likely to be more reliable.

Hint An example of an opportunity sample from this data would be to select the first five heights from the list.

Homework Answers

1 a i Divide the population into groups according to given characteristics. The size of each group determines the proportion of the sample that should have that characteristic. The interviewer assesses which group people fall into as part of the interview. Once a quota has been filled, no more people in that group are interviewed.

ii Opportunity sampling consists of taking the sample from the people who are available at the time the study is carried out, e.g. the first 40 shoppers who are available to be interviewed.

b Quota sampling.

2 Similarities: The population is divided according to the characteristics of the whole population (into strata for stratified sampling, and groups for quota sampling)
Differences: Stratified sampling uses random sampling whereas quota sampling does not.

3 a Opportunity sampling

b Sample is likely to be biased towards people who eat fish and chips on a Friday.

c Survey people at different times of day. Survey people in other parts of the town, not outside the fish and chip shop.

4 a 5.4 hours

b Opportunity sampling; unlikely to provide a representative sample of the town as a whole

c Increase the number of people asked. Ask people at different times/in different locations.

5 a Quota sampling.

b ANY ONE FROM: no sampling frame required, quick, easy, inexpensive, allows for comparison between male and female deer.

c Males are on average heavier and have a greater spread.

d Increase the sample size. Catch deer at random times during the day.

6 a Student's opportunity sample: For example, first five values

b 1.9, 2.0, 2.6, 2.3, 2.0

c 1.96 m, 2.16 m

d Systematic sample – is random and likely to be more representative. Opportunity sample might get all the small values, for example.