ISBN 0 902246 96 8

© C. Dixon and B. Leach 1977

SAMPLING METHODS FOR GEOGRAPHICAL RESEARCH

C.Dixon & B.Leach

CATMOG

(Concepts and Techniques in Modern Geography)

CATMOG has been created to fill a teaching need in the field of quantitative methods in undergraduate geography courses. These texts are admirable guides for the teachers, yet cheap enough for student purchase as the basis of class-work. Each book is written by an author currently working with the technique or concept he describes.

1. An introduction to Markov chain analysis - L. Collins
2. Distance decay in spatial interactions - P.J. Taylor
3. Understanding canonical correlation analysis - D. Clark
4. Some theoretical and applied aspects of spatial interaction shopping models - S. Openshaw
5. An introduction to trend surface analysis - D. Unwin
6. Classification in geography - R.J. Johnston
7. An introduction to factor analytical techniques - J.B. Goddard & A. Kirby
8. Principal components analysis - S. Daultrey
9. Causal inferences from dichotomous variables - N. Davidson
10. Introduction to the use of logit models in geography - N. Wrigley
11. Linear programming: elementary geographical applications of the transportation problem - A. Hay
12. An introduction to quadrat analysis - R.W. Thomas
13. An introduction to time-geography - N.J. Thrift
14. An introduction to graph theoretical methods in geography - K.J. Tinkler
15. Linear regression in geography - R. Ferguson
16. Probability surface mapping. An introduction with examples and Fortran programs - N. Wrigley
17. Sampling methods for geographical research - C. Dixon & B. Leach
18. Questionnaires and interviews in geographical research - C. Dixon & B. Leach

Other titles in preparation

This series, Concepts and Techniques in Modern Geography is produced by the Study Group in Quantitative Methods, of the Institute of British Geographers.
For details of membership of the Study Group, write to the Institute of British Geographers, 1 Kensington Gore, London, S.W.7.
The series is published by Geo Abstracts, University of East Anglia, Norwich, NR4 7TJ, to whom all other enquiries should be addressed.

CONCEPTS AND TECHNIQUES IN MODERN GEOGRAPHY No. 17

SAMPLING METHODS FOR GEOGRAPHICAL RESEARCH

by

C. Dixon and B. Leach

(City of London Polytechnic and Thames Polytechnic)

CONTENTS

		Page
I	BASIC CONSIDERATIONS	
(i)	Sampling in research design	3
(ii)	Sampling and probability	3
(iii)	What is a good sample?	4
(iv)	Sample size	7
(v)	Target population and sampling frame	11
II	PROBABILITY SAMPLES	
(i)	Simple random samples	13
(ii)	Systematic samples	14
(iii)	'Systematic random' samples	15
(iv)	Stratification	16
(v)	The procedure for stratifying	17
(vi)	The choice of sampling fraction	18
(vii)	Weighting	19
(viii)	Stratification after selection	20
III	CLUSTERING	
(i)	Reducing costs over a large area	20
(ii)	Creating clusters	23
(iii)	Selecting clusters	23
(iv)	Sampling with probability proportional to size	24
IV	SAMPLING FRAMES	
(i)	Convenient sampling frames	25
(ii)	Inconvenient sampling frames	26
(iii)	Sampling from lists: the example of the British Electoral Register	27

(iv) Selecting within small clusters: the example of households 28
(v) Constructing frames from maps 30
(vi) Constructing frames by mapping 30
(vii) Longitudinal studies 31
(viii) Sampling continuous flows 32

V AREAL SAMPLING METHODS
(i) Frames and methods 33
(ii) Problems of spatial frames 38
(iii) Sampling in space and time 39

VI NON-PROBABILITY SAMPLES
(i) Purposive samples 40
(ii) Quota samples 40
(iii) Random routes 41
(iv) Combining quota sampling with a random route 42
(v) Snowball samples 42
(vi) Justifying non-probability samples 43

VII RESEARCH DESIGN AND THE CHOICE OF SAMPLING METHOD 43

BIBLIOGRAPHY 46

I BASIC CONSIDERATIONS

(i) Sampling in research design

We might want to design a research project to investigate the livestock of farms in an area to be visited on a field course, or to find out where the households on a particular estate do their shopping, or to measure an island's soil properties. In each of these cases, we want to make generalisations about a large number of units (elements) in a population consisting of all the elements (farms, households, possible measuring points) whose characteristics we want to describe.

In some cases, we can undertake a complete study of these elements, as we might with shops in a small centre, and on other occasions we might deliberately pick out for study a few items, for example, several towns of the British Isles. But generally our resources do not stretch to collecting data about all the elements in the population, and we have to make inferences about the whole population by studying only some of them.

Even if our resources were adequate, a more economical way of carrying out our study would be to draw a sample from the population. Clearly a sample will only approximately represent the characteristics of the parent population, since it only contains part of it, but sampling theory enables us to estimate the likelihood of our sample being a good representation of the population, provided we have followed certain rules in the choice of elements for the sample.

(ii) Sampling and probability

This book is mainly concerned with probability sampling, which requires that each element in the population has a nown c ance of se ection for the sample. Its simplest form is the 'simple random sample' in which all elements have an equal chance of inclusion. This means that any particular sample, any combination of elements, drawn from a population would have the same chance of occurring as any other sample of the same size.

If a researcher only looks at the fields he can see from the road, or interviews the people who happen to be in the street, his sample is clearly biased. It does not surprise us that students sent to interview shoppers in a market each return with a sample consisting of a different sort of people, elderly ladies, young men, or whoever the student found it easiest to approach. However, it is harder to realise that we cannot choose elements 'at random', without any selection bias. Researchers sampling from a list of names may tend to select more from one part than another; a biogeographer throwing a stick to find a random point may subconsciously direct it towards some interesting vegetation; if we place a small quadrat frame over a part of the crop in a field, can we really do so without considering if the location chosen is above, below or close to the average for the field? Only a method

in which all the subjective choice is removed ensures that each sample, each possible combination of elements, has exactly the same chance of selection. Only if we have strict rules which prevent the researcher, or his agents such as field-workers, using 'judgement' at any stage can a sample be subjected to statistical testing on the basis that it is truly a probability sample. One of the theoretical texts such as those in section A of the bibliography should be consulted by those wanting a detailed discussion of sampling theory. In the sections which follow, we summarise the main conclusions which have the greatest relevance for the practical design of samples.

(iii) What is a 'good sample'?

The major paradox of sampling is that, since we are trying to find out about a population whose characteristics we do not know, we can rarely measure the success with which our sample represents that population. Sampling theory enables us to estimate the range around each sample statistic within which we expect the true population value or parameter to lie, with a calculable margin of error.

If a number of samples were drawn from a population they would provide a number of estimates of such population parameters as the mean. While each sample might give us a different estimate of the mean, the spread of all possible samples of that size that we could conceivably select forms a distribution scattered around the true mean, the population parameter. We can calculate statistics to describe the characteristics of any distribution, the most important for our purpose being the variance, a measure of the spread around the mean, or its square root, called the standard deviation.

The standard deviation is found by calculating the deviation of each value in the sample, x, from the mean, $\bar{x}$, squaring the deviations, adding item together, dividing by the size of the sample, and finally taking the square root, so that it will be in the same units as the mean. Where the sample is small, dividing by (n-1) rather than n, the sample size, will give a significantly larger answer. This compensates for the fact that the smaller the sample, the less likely it is to have as large a variability as the population itself.

The usual formula for the standard deviation of a sample is:

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}} \qquad (1)$$

where s is the standard deviation of a sample, x is the value of each observation, $\bar{x}$ is the mean, n is the sample size, and Σ is the summation sign which means add together.

In sampling theory, the standard deviation of the distribution of all the possible sample means around the true, population mean is called the standard error of the estimate, because each sample mean will be a more or less accurate estimator of the true value. The more variable the population itself, the more likely we are to have a sample which is very different from the population, and the greater the range of possible estimates. The standard error is therefore larger when we are measuring something which is very variable in the elements of the population.

We need to know this standard error if we are to have an indication of the likelihood of having drawn a 'good' sample, one which closely corresponds to the parent population.

Although we cannot actually calculate the standard error without knowing the true population mean - the very thing we want our sample to estimate - the distribution of values in our sample can give us some indication of the probable standard error. The more variable the parent population, the more variable the samples from it will tend to be. If we draw a very small sample, it might not reflect all the variability in the population, because sometimes, by chance, all the elements in the sample will be rather similar. However, the larger our sample, the more likely it is to be as variable as the population itself.

We therefore calculate the standard deviation of the values in our sample, using formula 1. Then we estimate the standard error from it:

$$SE = \left\{ \frac{s}{\sqrt{n}} \right\} \sqrt{1 - n/N} \qquad (2a)$$

where SE is the standard error, s is the standard deviation of the sample (formula 1), n is the sample size, and N is the size of the population.

The expression (1-n/N) is designed to take into account the sampling fraction, the proportion of the population elements included in the sample, (n/N), because we should expect the sample to resemble the population more closely if it is a large proportion of it. Unless n is a very large proportion of N, it will make little difference, and for practical purposes this factor is frequently ignored. Thus the standard error of the mean is usually estimated:

$$SE = \frac{s}{\sqrt{n}} \qquad (2b)$$

Although realistically we would never be able to draw all the possible samples from the population in which we were interested, we are able to use their theoretical distribution as the basis for estimating the standard error of any individual estimate, the value found in one sample. There are further assumptions that we can make about the theoretical sample distribution.

If we drew all the possible samples from the population, we would expect that as many of the samples would have means above the true mean as below. The mean of the sample means would be the population mean.

In addition, the distribution of sample means would always form a normal curve, a distribution whose main mathematical property is that, regardless of any particular mean and standard deviation, there will always be a constant proportion of observations (in this case, sample means) within each standard deviation unit (in the case of a sampling distribution, standard error unit) of the mean. However variable the population, and thus however large the variation in the possible samples, we would always expect 95 per cent of the sample means to be closer to the true population mean than 1.96 times the standard error; within 2.58 SE there will always be 99 per cent of sample means. Statistical tables give the percentage at other standard error distances, but these are the most generally used, because we are usually interested in ensuring that we will be able to predict that our sample value is

within a stated range of the true value, with a 95 per cent or 99 per cent chance of being right.

This chance is called the confidence level, and is usually expressed as a 95 or 99 per cent chance of being right, or conversely as a probability of 5 per cent or 1 per cent (.05 or .01) of making a statement which is wrong about the relationship between the estimate and the parameter.

If 95 per cent of the sample means lie within 1.96 SE of the population mean, 95 per cent of the time we expect the population mean to lie within the same distance of our sample mean. Therefore, in order to be 95 per cent confident of our prediction of the population value, we can add or subtract 1.96 times the standard error to our estimate, and assert that the true mean will lie within this range.

The distance from the sample mean (in either direction) within which we expect to find the population value is the confidence limit, calculated by:

$$c = \pm z.SE \qquad (3)$$

where c is the confidence limit and z is the standard-error-unit measure for the desired confidence level. The sign ± indicates that the figure given is the range on either side of the mean, since we do not know whether our sample mean is above or below the true value.

For the 95 per cent confidence level, the confidence limit is:

$$c = \pm 1.96\ SE$$

If the mean land holding in an area were found from a sample of 100 farms to be 53 hectares, and the standard deviation to be 26, the confidence limits could be calculated for the 95 per cent probability level:

$$\begin{aligned} c &= + 1.96\ SE \\ &= \pm 1.96\ s / \sqrt{n} \\ &= \pm \frac{1.96 \times 26}{\sqrt{100}} \\ &= \pm 5.1 \end{aligned}$$

We could therefore say that we were 95 per cent confident that the true mean lies in the range 53 ± 5.1, that is, from 47.9 to 58.1 hectares. If we wanted to be 99 per cent confident in our prediction of the population value, we would calculate c using 2.58 instead of 1.96 in formula (3), and this would give a confidence limit of 6.7; our population mean could then be said to lie within the range 46.3 to 59.7 hectares with 99 per cent confidence.

When making assertions about population values on the basis of an estimate from a sample, we can choose a high confidence level, a low chance of being wrong, but then we have to make a less definite statement, that is, a statement with wider confidence limits, a bigger range within which we expect to find the population value.

The standard error, and thus the size of the confidence limits, depends primarily on the variability of each characteristic studied. Samples which are quite small may give very good estimates of phenomena which are fairly uniform in the population as a whole, but we might find that when we calculate the confidence limits for variables with a more erratic distribution, our assertions about the value for the population as a whole must be very qualified indeed. In the case of the farms mentioned above, we might find that the number of cows on each farm varied very little, and that we could say that the average herd was 28 with a confidence limit of 2.2 and 99 per cent confidence. In other words, we would be 99 per cent sure that the true value was between 25.8 and 30.2 cows - a very close estimate indeed.

It would be rare for research reports to qualify all their assertions about the population by the addition of confidence limits and a probability of accuracy to each estimate, because their fluency would be greatly impeded. However, the calculation should be performed for several variables reported, and the results presented in a footnote or appendix, to indicate the confidence with which the reader should approach the statements made in the text. It is also important to give the standard deviation of variables in the study. This will serve as an indicator of the standard error; a reader who sees a large standard deviation and knows that the sample was small will approach statements made about the population with caution, knowing that the confidence limits will be large.

(iv) Sample size

In general, a sample of 30 is the smallest that can be expected to conform to the normal distribution on which sampling theory, as outlined above, is based. However, the larger the sample, the more accurate it will be, that is, the nearer values calculated from it will be to the true population values.

Intuitively, we can see that if we want to estimate what proportion of all students study geography, a sample of ten could not be expected to reflect the whole student body as closely as a sample of a hundred would.

It is primarily sample size, and not the percentage of the population included in the sample, which determines the accuracy of a sample. If we deal out four hands of playing cards, each will tend to contain a mixture of cards, of all suits, and of both high and low values. If we increase the number of packs of cards to twenty, and the size of the hand to 52, the spread of suits and values will be more even, although the sampling fraction will have decreased from 1 in 4 to 1 in 20.

Thus the larger our sample, the more confident we can be in our prediction, or the narrower the range in which the predicted value lies. That is why in the formula (3) for the confidence limit, the sample size, n, was included. However, we should expect diminishing returns. Increasing the number of cards in each hand from 13 to 18 would have a greater effect than adding 5 cards to a hand of 52. For this reason sampling theory indicates that it is the square root of n, rather than n, that should be used in these calculations.

The second factor which, as has already been mentioned, will affect the accuracy of a sample is the variability of the population. The more different the elements are from each other, the larger the sample needed to represent them accurately.

In a pack of playing cards, there are two colours, four suits, and 13 different values. We should need a comparatively small sample from the 20 packs to have a balance of the red and black cards, but a larger one would be

needed before the suits would be found in equal numbers. A still larger hand would have to be dealt before the hand contained all 13 values in their correct proportions.

The researcher will not usually know much about the things he sets out to measure, but he will need some indication of the variability before he can determine the sample size he needs for accuracy.

Equation (3) related the confidence limit to the standard-error-unit measure for the desired confidence level and the standard error:

$$c = \pm (z.SE) \tag{3a}$$

Because $SE = s / \sqrt{n}$ from equation 2(b), equation 3 can be written:

$$c = \pm (z.s / \sqrt{n}) \tag{3b}$$

It therefore follows that we can determine the sample size needed for a given confidence level and confidence limit by turning this information round:

$$\sqrt{n} = \frac{z.s}{c} \tag{4a}$$

or, more conveniently:

$$n = \left\{\frac{z.s}{c}\right\}^2 \tag{4b}$$

In the playing card example, we were concerned with attributes like colour and suit which a certain proportion of the cards have. In the case of a proportion rather than a mean, we cannot calculate a standard deviation. The measure of variability which we use instead, for example in calculating the standard error of a proportion, is v, measured by:

$$v = \sqrt{p\,(100 - p)} \tag{5}$$

where p is the percentage with the characteristic.

Therefore, instead of formula (4b), we estimate the necessary sample size using v instead of s:

$$n = \left\{\frac{z\ .\ v}{c}\right\}^2 \tag{6}$$

Using these equations, if we need a sample to estimate the proportion of households with cars to within 2 per cent (the confidence limit) with 95 per cent confidence (the confidence level), in an area where we expect that only half will have cars, we substitute in equation (6):

$$v = \sqrt{50\,(100 - 50)}$$

$$= 50$$

$$\therefore \quad n = \frac{1.96 \times 50}{2}^{\ 2}$$

$$= 49^2$$

$$= 2401$$

This means that we would need a sample of 2401 households to ensure the precision specified.

Some small scale studies will draw samples from small populations, where the sampling fraction, the percentage included in the sample, will be so large as to increase the accuracy of the sample significantly. The answer obtained by formula (4) or (6) can be corrected to take account of the sampling fraction:

$$n' = \frac{n}{1 + (n/N)} \tag{7}$$

where n' is the corrected sample size, n is the sample size calculated with formula (4b) or (6), and N is the population size.

If there were 10,000 households in the area, we could estimate car ownership to within 2 per cent with a sample of 1936 instead of the previously estimated sample of 2401. However, if there were 100,000 households, the corrected sample size from formula (7) would be 2345, not a very great reduction from the original estimate.

Normally only professional studies with considerable resources can conduct pilot exploratory work to find out the likely variability in the population, but it is possible to draw on the experience of other research. For human populations, sources such as population censuses can be used to make an informed guess of the variability in the population to be sampled.

A more serious problem is that we are rarely concerned with only one feature of the population, and the sample size needed for each will depend on its variability. It is therefore the case that only an approximation to the desired sample size can be made, and for this purpose Table 1 may be adequate.

Table 1 gives the sample size needed to estimate population values to within a chosen percentage (the confidence limit) with a desired probability of being right (the confidence level) if the variability of the population is 50 per cent. This was the case in the car ownership example, where half the households had the characteristic. For a proportion, this is the maximum possible variability, because v is the square root of 50 times 50, 2500, (49 times 51 is 2499), and the table therefore gives a conservative estimate of sample size.

For a continuous variable, 50 per cent would not be the maximum possible value. Not infrequently, the standard deviation is more than half the mean, giving a coefficient of variability (100 s / R) over 50 per cent. The table may therefore give an underestimate of the sample size needed for continuous variables.

It can be seen that in order to obtain answers to within a narrow range it is necessary to have a larger sample size if the population is variable. In choosing a research project for a small scale study, it might be possible to scale down the original ambitions to reduce the sample size needed.

Reducing the variability of the population could reduce the sample size needed; to generalise about shopping patterns, a study of all households would almost certainly need a larger sample than one concentrating on one particular type of household, such as families with small children.

The size of the target population can also be reduced, so that instead of drawing the sample from a large area, a smaller one with fewer elements is chosen. The sampling fraction may then have some effect on the sample size needed, which can be adjusted with equation (7). This could not be done in cases where the population was of unlimited size, for example consisted of all possible sampling points in a river basin; in such a case, unless reducing the area reduced the variability, there would be no increase in accuracy.

TABLE 1

Sample sizes needed to estimate population values with given levels of confidence, assuming a variability of 50%[1] and a very large population.'

(from equations 4b or 6)

Confidence limit[3] (±%)[4]	Confidence level 99%	95%
1	16587	9604
2	4147	2401
3	1843	1067
4	1037	600
5	663	384
6	461	267
7	339	196
8	259	150
9	205	119
10	166	96
15	74	43
20	41	24

[1] The table assumes a variability of 50%. For an attribute, this means $v = \sqrt{p(100-p)} = 50$, which is the case when p, the proportion with the quality, is .5. In the case of a continuous variable, this means the coefficient of variability ($100\ s/\bar{x}$) is 50%, in other words, the standard deviation is 50% of the mean.

[2] Where the sample size given by the table is a substantial proportion of the population size, the figure on the table should be corrected by formula (7) in the text.

[3] For a complex sample design which is expected to increase the standard error by 1.5, the confidence limit should be set to 2/3 of the desired limit, for example, for a desired confidence limit of ± 6%, read the sample size for ± 4% (Section III, i).

[4] The confidence limit is measured as a percentage. For an attribute, it is expressed as the percentage having the characteristic. For a continuous variable, it is given as a percentage of the mean.

Redefinition of the target population is not something which a client or sponsor of a commercial study would usually accept, but in academic research the choice may be between an inadequate sample of a highly diverse population, or a less wide-ranging one about which it is possible to infer characteristics with greater precision.

In considering sample size, it is also necessary to decide the minimum numbers needed in particular sub-groups of the population which are to be compared. Statistical tests will require a minimum in each category for any meaningful conclusions to be reached.

A final consideration is the size of the data body that can be handled at the processing stage. If the measurement to be undertaken on each element is very lengthy, then the sample size may be limited accordingly.

Efficiency and resources of time and money have generally to be weighed against each other. To reduce the sample error by 30 per cent, the sample size will have to be doubled. This can be seen in Table 1, where the reduction of the confidence limit from ± 6% to ± 4% necessitates an increase in sample size from 267 to 600 at the 95% confidence level.

In the case of sampling from maps or lists, this increase in size may not lead to a doubling of time and effort, but in the case of a questionnaire survey, one of the leading research agencies estimates that doubling the sample size would increase costs by 80%. Other studies involving field work would expect similarly large increases.

In practice, in student work, sample size is usually determined by time and resources. It may be that, having calculated the confidence limits for the sample size that is feasible, it will be clear that statements could not be made about the population with any degree of reliability. If we are unable to conduct a study with a sample size sufficient to improve our population estimates above the level of guesswork, then we might as well not conduct it at all. Small scale work, even if it is only a very exploratory study, intended to generate ideas rather than to draw conclusions about a population, must recognise its limitations. Before embarking on any sampling procedure the researcher should calculate the likely confidence limits for his estimates. It will be in many cases a discouraging exercise, but is better done beforehand than after collecting some meaningless data.

(v) Target population and sampling frame

The results of a sample study are only applicable to the population from which the sample was drawn, although they may stimulate other research, refute or support generalisations made elsewhere, or contribute to the formulation of ideas about a wider universe.

The researcher will normally define a population of a size and variability that he can study with the resources he has available. However, he may also have to re-define this target population to enable it to be drawn from a convenient source, such as a list or a map, known as the sampling frame.

Establishing a sampling frame, which contains all the elements in the target population, is the first stage in sampling. If the frame is imperfect, the researcher has to decide in the circumstances of his particular study whether to accept its deficiencies by re-definition of the study population, or to amend the frame. It is only possible to compensate for obvious

inaccuracies; often one can only evaluate the frame after completing the research. If an old frame is being used, it may need up-dating.

There are two possibilities if the frame excludes part of the original target population. For example, if the land use in a country is to be studied from a map, but this does not cover the north east corner, we could obtain another map, or prepare an additional map in some way, or alternatively we could re-define the population in terms of the existing map. This would mean stating that the study only concerned the part of the country shown. Provided that this was made clear in any discussion of findings, the results for the rest of the area would be by no means invalidated.

Similarly, in dealing with human populations, it is quite usual to exclude those living in institutions such as army barracks, halls of residence or hospitals, and the report will point this out.

Where an ineligible element is drawn in a sample, it cannot be replaced by the next one on the list, as this would give the one following a second chance of selection. If a list contains a large number of ineligible entries, or substantial parts of a map are to be excluded, enough elements must be chosen to enable the final total of eligible elements to correspond to the desired sample size.

Sometimes an element will occur twice in a frame. In special cases, duplication will be so rare that it can be conveniently ignored. For example, so few individuals would qualify as members of more than one household that, although they would actually have had two chances of selection, they would hardly ever receive special treatment.

More usually, though, if an element occurs more than once, one occurrence, normally the first, is taken as the one qualifying for inclusion in the sample. If any of the others are selected, they are rejected as ineligible.

It is not always possible to find a comprehensive source to serve as a sampling frame, but there may be two or more sources, which, with some overlap, include all the elements in the population. If all the sources were used, elements which occurred in more than one would have an increased chance of selection. The various sources are therefore placed in order, starting with the most comprehensive or best organised. The first entry of each element is then taken to be the decisive one. The initial selection of elements is made from all the sources, using the same sampling fraction. All those in the first list or on the first map are automatically included, but those chosen in the others are checked to see if they had an entry in an earlier list, or were on an earlier map; they would then be rejected as ineligible.

The form of the sampling frame will to a large extent determine the possible sampling methods, but before looking in detail at the different types of frame (chapters IV and V), it is helpful to consider the alternative methods of selecting a sample.

II PROBABILITY SAMPLES

(i) Simple random samples

The most straightforward sort of sample is the simple random sample already mentioned, where elements from a list or other sampling frame are selected randomly, usually by using random numbers.

Each element in the sampling frame is allocated a unique identifying number. Numbers are then read from a random number table, and elements in the frame with those numbers are included in the sample. Random numbers, published in every set of statistical tables, are simply lists of numbers in which each digit has an equal chance of occurring in each position. As many can be used at once as are necessary to correspond with the size of the identifying numbers on the frame. For example, if the list were numbered 1 to 1000, three digits could be read at a time, with the number 000 corresponding to 1000. If the frame included 1001, four digits would be needed. If numbers are read which do not correspond to numbers on the sampling frame, then these ineligibles are ignored, and the reading of numbers continues until enough valid numbers have been read. It is usual to start from a randomly selected point on the table.

Sampling theory assumes that if the same element is picked twice, it will be included twice in the sample. In practice, this would not generally be done, since it would tend to reduce the variability in the sample, because two elements included would be identical. The technical term for not allowing the same item to occur twice in the sample is sampling without replacement; once an element has been selected, it ceases to have another chance of being drawn, and those remaining have a slightly increased chance - although this is only significant if the sample is a very large proportion of the population.

The procedure of the simple random sample will produce a sample which can be of exactly the desired size, from any part of the sampling frame. It might happen, since each digit has an equal chance of occurrence and therefore all combinations of elements in the frame could theoretically be encountered, that the chosen sample is concentrated in one part of the population. If the list contained males and females, the chosen sample might be of one sex only, and in theory it would be sometimes. If the sampling frame covered different rock types, only one might be found in a sample drawn at random.

Although simple random sampling is in many cases reasonably quick to use, provided that the sampling frame can be readily numbered (as lists and maps with grids usually are), better coverage could be ensured by other methods.

Deviations from the simple random sample will be used to increase precision, that is, to reduce the spread of sample estimates around the population value, and eliminate the sort of 'freak' samples just mentioned. They will also be used for practical reasons, or to reduce the cost of drawing the sample. The alternative methods are discussed in the succeeding sections.

(ii) Systematic samples

One simple and convenient method of ensuring even coverage throughout a sampling frame is the systematic sample. Instead of using random numbers which could indicate elements in any part of the population, a regular spacing is used, taking every kth individual, or the intersections of a regular grid laid over a map.

The sampling interval, k, is the reciprocal of the desired sampling fraction, that is, k is N/n. If the sample is to contain 50 elements and the sampling frame lists 5000, one element in 100 would be chosen by setting k at 100.

The sample is selected systematically from a random starting point. This is usually a number between 1 and k, in order to ensure that all the items in the list before the first one chosen have a chance of selection. This is the only random number which needs to be drawn. Thereafter the selection of the components of the sample proceeds simply, with every kth one chosen.

It is subsequently very easy to check that the sampling method has been correctly applied, because the sample will be regularly spaced; the sample will come from every part of the original frame in the correct proportions. However, there are two major reasons why the systematic sample, for all its attractions and its frequent use, is not strictly a random sample.

Firstly, the systematic method does not enable all possible different samples of the given size to be selected. The number of different samples is limited by the sampling interval k. The selection of members of the sample is not independent, because once the first individual has been chosen, all the others are fixed. This violates one of the assumptions on which sampling theory is based.

Secondly, if the sample produced by this method is to be analysed as if it were a simple random sample, the frame needs to be in a random order. Many lists can be assumed to be randomly ordered, but many lists and certainly most spatial arrangements, whether plants, landscape features or the built environment, contain order. In many cases there is a periodicity which may not be immediately apparent. If a list of households is arranged in order corresponding to their position along a street, taking a sampling interval with an even number might produce a sample consisting entirely of the left-hand pair of semi-detached houses, which might be different in size or other important features. Terrain or crops, too, might have periodic features which could be picked up by a sampling method which used a regular interval.

An additional problem can occur when a list is arranged in an order, for example, in ascending order of magnitude of some variable, (either one being studied or one which correlates with variables being studied). In such cases, the initial random start can affect the estimate of the population mean, for a high starting point will bias the sample upwards, a low start downwards. The effect may be reduced by treating the list as circular. After drawing every kth element, those left at the end are added to those at the beginning, and if there are more than k, another one is drawn; this is more likely to give equal representation to high and low values but will not completely solve the problem.

A systematic sample is not equivalent to a random sample, but it may be treated as one if the pitfalls discussed above are considered, and the researcher feels confident that regularities in the sampling frame will not lead to bias in the sample.

(iii) 'Systematic random' samples

Some of the problems of systematic samples can be minimised, while retaining the advantages of even coverage, by combining systematic sampling with a greater random component. There are a number of ways of doing this.

The simplest method where periodicity is suspected is to use a series of random starts scattered throughout the frame, and to select with the intercal k between them, as if each part constituted a separate list. For example, in order to select 100 from a population of 1000, k would be 10. We might choose 5 random numbers, starting with a random number between 1 and 10. If these happened to be 4, 26, 364, 505 and 787, we would start by sampling 4, 14, 24, and then 26, 36, 46 .... 364, 374, 384 ....

Alternatively, a series of low starts could be made and a systematic sample drawn with the same interval from each. If we wanted to take 100 from 1000 with 5 starts, we would use k = 50, and make 5 starts between 1 and 50. Using an interval of 50 from each of these starts would have the same effect as using an interval of 10 throughout the whole sequence.

A spaced random sample has the advantage of even coverage, but lacks the advantage of quick selection. Random numbers are read for each member of the sample, but the choice of an individual element excludes all its neighbours up to a chosen distance away. This distance must be less than k/2, and the closer to k/2 it is, the more evenly spaced this sample will be. Sampling from maps might lend itself to this method, where the selection of a point could exclude all the points in a circle around.

A randomly varying interval can be used instead of a fixed k. The interval is varied each time by a small random number, r, added or subtracted, so that the elements are chosen at intervals of k ± r, from a start between 1 and k. Whether r is to be added or subtracted each time should be randomly determined. One way of doing this would be to read an extra random digit, an odd number indicating plus, an even number (or 0) minus. Small random numbers can be drawn quickly using playing cards; if r was to have a maximum value of ± 7, the higher cards (and the picture cards) would be removed. Red cards could indicate positive values, black cards negative values of r.

The desired k and the limits of r would have to be decided first, and the random start made between 1 and k. With k = 50 and r between 1 and 7, a starting point of 48 might be made. The next element would be chosen by drawing a card. Red four would mean that the interval was 50 + 4, and the element chosen would be 102. The next card might be black 1, and the interval therefore 49, giving element number 151.

The final sample size using the k ± r method would be very close to the sample size using a fixed k.

Many sampling frames will lend themselves to division into systematic blocks, where the number of blocks is the same as the desired sample size. One element is then chosen from each at random. The advantage of this method is that only small random numbers will have to be drawn.

If the sample size is not an exact division of the population, the final block will be smaller than the others. 'Blanks' are therefore added to it to make it the same size, and if one of the blanks is chosen, no element from the last section will be included in the sample. Thus each element in the smaller section will have the same chance of inclusion as those in the other blocks.

(iv) Stratification

Systematic samples effectively choose elements from even-sized blocks of the population. Strata, sub-divisions of the frame within which elements are chosen, can be simply even divisions of the population (or approximately even, as in the case of the randomly varying interval systematic method), but they can be sub-groups of the population of any size that are expected to differ from each other in ways important to our study.

If we want a sample to contain both men and women, it would be sensible to sub-divide the frame on the basis of sex before sampling (if we had the information to do this), rather than run the risk of obtaining a sample unrepresentative of the population. If we expect different areas to have different characteristics, it will be as well to consider them as separate strata, and choose within each of them.

There are two benefits from stratifying in this way. Firstly, it will ensure that each part of the population which we might want to compare will be represented with a sample of sufficient size for the comparison to be made. If it is known in advance what groups are to be tested, and they can be identified in the frame, even approximately, it will be beneficial to stratify.

The second major advantage of stratifying is that it will tend to improve population estimates, provided that the variable used for the stratification is related to the subject of the study. For example, we might expect soil types to be important determinants of aspects of plant communities. The more different the plants on each soil type are, and the more internally homogeneous the communities are, the more the estimates of the population will gain from the use of the strata.

If the frame were randomly ordered in terms of the variables being studied, then there would be no gain from stratification (or from systematic sampling) and a simple random sample would be as good. However, this completely random arrangement will be unusual, and although we will not normally be able to calculate how much we have gained from stratifying, we can expect that in almost all circumstances our population estimates will be more precise as a result. Greater precision means a reduction in the spread of sample estimates around the population parameter. Each individual sample may not be better than each simple random sample, but freak samples will be less likely to occur. This means that, at a given confidence level, the confidence limits will be narrower for a given sample size - but the same effect might be achieved more cheaply by drawing a larger simple random sample.

Even if the simple random sample will not produce great savings, it will facilitate calculations, for even when the same sampling fraction has been used in each stratum, the standard error is estimated by summing the variance in each stratum:

$$SE_{strat} = \sqrt{\frac{\Sigma (n_i \, s_i^2) \, (1-n/N)}{n^2}} \qquad (8a)$$

where SE is the standard error of a stratified sample, n_i is the number chosen in stratum i, n is the total sample size, and s_i is the standard deviation in stratum i; Σ means sum for all the strata. For the standard error of a proportion, we would use v_i, the variability of a proportion in each stratum (equation 5), instead of s_i.

Unless the sampling fraction were very large, we could ignore the expression (1-n/N), and calculate the standard error:

$$SE_{strat} = \sqrt{\frac{\Sigma n_i \, s_i^2}{n^2}} \qquad (8b)$$

Many researchers make the assumption that the variability in all the strata will be very similar and continue to use the formulas given in section I, which are applicable to simple random samples. This has the advantage of greatly easing the calculations, but each variable studied may have a different variability in the strata, and a different degree of correlation with the variable used to create the strata. For most variables the standard error calculated by the simple random sample formula would be an over-estimate of the sampling error.

(v) The procedure for stratifying

In some cases, the sampling frame will be already divided into suitable strata, for example, a trade directory may be arranged in different categories of businesses, but in others the researcher will have to decide how to divide up the frame.

If the sampling frame is conveniently divided, it is possible to draw the part of the sample in each stratum separately, either randomly or systematically. Alternatively, a random sample with stratum quotas can be drawn by dividing the sampling frame into 'strata', which are either blocks of equal size, or which correspond to differentiated parts of the frame, for example soil types or areas of a town. Each part is allocated a proportion of the sample, and random numbers are read until the quota in each stratum is full; numbers are discarded if they fall in a part whose quota is completed.

There is no need for the strata to be of identical size, and there is no rule about the appropriate number of strata, although Kish (1965, p.102) suggests that between three and ten will normally be best.

If the sampling frame provides little information with which to perform the stratification, supplementary information from another source may be used, provided that it correlates with the desired stratifying variable.

Thus, if we wanted to include even coverage of areas with small and large percentages of pensioners or particular house types, we could use a frame such as the Electoral Register (section IV, iii) which did not contain any information on these variables, but distinguish our strata by Census,

airectory or map information for the same areas. This would probably be valid even if the source was out of date, but the researcher would have to be sensitive to any changes (new buildings, demolition and so on) which might radically have altered the areas.

This suggests another important point: the researcher should always know his study area as well as possible before commencing the formal data collection stage. He may in fact know the characteristics of the frame sufficiently to make an approximate division on a subjective basis. The categorisation of towns, house types, central and peripheral parts of a settlement or remote and accessible parts of a nature reserve might be used to create strata which would improve the population estimates. Provided that, having created the strata, sampling proceeds rigorously, there is no obligation to use only information internal to the sampling frame.

(vi) The choice of sampling fraction

The simplest form of stratification involves applying the same sampling fraction, that is, taking the same proportion of elements, in each stratum. There are occasions, though, when the use of a different sampling fraction in particular strata can be beneficial.

The first case is when the variability of the population is very different from one stratum to another, and is known or can be guessed in advance. Thesma leh the variability, the smaller the sample size needed to provide an estimate of its characteristics (section I, iv), and the same is true of the part of the population in a stratum. A smaller number can be taken from a more homogeneous stratum, saving resources which can profitably be applied to the more variable strata.

If our sampling frame consists of mothers of children attending a particular nursery, for example, we might expect the mothers with jobs to differ from those without. If the topic under investigation is perception of the immediate vicinity, we might expect those not working to be more alike than those who travel away from the area to work. If, on the other hand, we were studying the behaviour patterns during the time between leaving and collecting the children, the working mothers would almost certainly be less variable.

If we feel sufficiently confident in our decision about the differential variability of the two strata, we might produce a more efficient sample (one which gives a better population estimate than a simple random sample of the same size) by using a larger sampling fraction in the more variable stratum. However, it might happen that the working mothers are more uniform in their perceptions, since they all work on a single industrial estate, and we would then discover that we had made the wrong decision if we had over-sampled them. Our allocation of the samples to the strata would then produce a less efficient sample than a simple, unstratified, random sample. Only if we have strong evidence for differential variability should we use different sampling fractions, both because of the possibility of producing a worse sample than a simple random sample, and because unequal allocation will necessitate weighting, which, as the next section shows, complicates calculations.

The second case in which we might want to use different sampling fractions is when one stratum is very small, and we need to have a sufficient sample in that part to enable it to be compared with the rest. If the proportion of working mothers is very small, it might be necessary to sample them at a higher rate even if we expect that they will prove more homogeneous. The effect of this would be to increase the overall sampling error, making the sample a less efficient way of determining the characteristics of the whole population of mothers, but this disadvantage might be over-ridden by the desire to be able to make comparisons between the two groups.

(vii) Weighting

If the strata are sampled with different fractions, in order to generalise about the population as a whole, the sample values for the individual elements must be weighted by the reciprocal of the sampling fraction to represent their true proportion in the population.

Table 2 illustrates a stratified sample for a study of three villages; they have 200, 300 and 500 households respectively, and these are therefore the stratum population sizes, N_i. The sample of 100 has been divided unequally between the villages because it is expected that the third village will be less variable than the first. The Table shows the calculations for the percentage of households with cars. The number with a car in each village (u_i) is multiplied by the weight, the reciprocal of the sampling fraction, in other words, by N_i/n_i, in order to estimate the number of units in the population possessing cars. The sum of u_iw_i is equal to u, the estimated number of car-owning households in the total population of the three villages. We can estimate the proportion of car-owning households by dividing u by N.

TABLE 2

Estimating population values with a weighted sample

Strata	Number of households in stratum	Sample size in stratum	Number of car-owning households in sample from stratum i	Weight for stratum i	
	N_i	n_i	u_i	$w_i = N_i/n_i$	u_iw_i
Village one	200	35	18	200/35 = 5.7	102.9
Village two	300	30	21	300/30 = 10.0	210.0
Village three	500	35	14	500/35 = 14.3	200.0
Total	N = 1000	n = 100			Σu_iw_i = 512.9 = u

u, the estimated number with the characteristic, car-ownership, is 512.9
p, the estimated proportion with cars, is calculated:

$$p = u / N$$
$$= 512.9 / 1000$$
$$= .5129 \text{ or } 51.29\%$$

Kish (1965, p.427) makes the point that there is no need for the weights to be precise to several places of decimals; in most situations where weights vary from 10 to 99, rounding the weights to 2 digits will be sufficiently precise. Using 6, 10 and 14 as weights in the example above gives u = 514, not very different from the precise answer. Weights do not have to equal the reciprocal of the sampling fraction; more workable numbers may be obtained by using numbers which are in proportion. Instead of 6, 10 and 14, we could use 3, 5 and 7, but then instead of N in the calculations we would use $\Sigma n_i w_i$, that is, 500 instead of 1000.

The formula for estimating the standard error of a weighted sample is:

$$SE_{strat} = \sqrt{\frac{\sum\{(n_i w_i)^2 \ (s_i^2/n_i) \ (1 - n_i/N_i)\}}{(\Sigma n_i w_i)^2}} \qquad (9a)$$

where SE_{strat} is the standard error of a weighted sample, n_i is the sample size in stratum i, N_i is the population in stratum i, s_i is the standard deviation in stratum i, and w_i is the weight to be applied to the sample in stratum i; Σ means sum for all the strata.

If the weightings are the reciprocal of the sampling fraction, as in the example in Table 2, and n_i times w_i is equal to N_i (for village one 5.7 times 35 is 200) and therefore the sum of $n_i w_i$ for all the strata is equal to N, the total population (this is true in the example, where the sum of $n_i w_i$ is 1000), the formula can be simplified to:

$$SE_{strat} = \sqrt{\frac{\sum\{N_i^2 \ s_i^2 \ / \ n_i \ (1 - n_i \ / \ N_i)\}}{N^2}} \qquad (9b)$$

Unless the sampling fractions in the strata were very large, we could ignore the expression $(1 - n_i \ N_i)$, and calculate the standard error:

$$SE_{strat} = \sqrt{\frac{\sum\{N_i^2 \ s_i^2 \ / \ n_i\}}{N^2}} \qquad (9c)$$

Weighting will counteract to some extent the beneficial effects of stratifying because if some elements in the sample have very large weights, it will reduce the variability of the sample. As a general rule, though, for most variables we would expect our stratification to show appreciable gains over a simple random sample, even with some weighting.

The most efficient sample will be obtained when the sampling fraction in each stratum is proportional to the variability in that stratum, measured by the standard deviation. In other words, if we consider only one variable, and the standard deviations in three strata are 5, 6 and 9, the best sample is obtained when the sampling fractions are in these proportions.

If the costs of including each element in the sample vary between the strata, this factor can also be taken into account in determining the allocation of effort to the strata. The best sample is obtained when the sampling fraction in each stratum is proportional to the standard deviation divided by the square root of the cost per element, in other words $s_i \ / \ \sqrt{m_i}$, where s_i is the standard deviation in stratum i, and m_i is the cost per element there (Stuart, 1968, p.52).

It would be rare indeed to have sufficient information to obtain a sample which conformed to this optimum, but given the constraint about the complexity of processing samples with very large weights, some such allocation can be attempted. In addition, only specialised sample designs of very unusual populations would have strata where variability and costs differed substantially.

(viii) Stratification after selection

Occasionally, we may want to stratify on some variable which is not clear on the frame, but which is measured in the course of the study. In these situations, the original sample is sometimes weighted to correspond with the known distribution of the population. This method should be used with care, for if the final sample does not correspond closely to the characteristics of the population in one or two important respects, it may be that the elements that are included are not truly representative.

For example, we may want to weight a sample of people to correspond with the age groups in the population as at the most recent Census. However, there may be reasons why our sample does not have the correct proportions; our sampling frame may have excluded mobile sections of the population, for example, young people, or we may, by excluding those in institutions, have reduced the number of elderly people. Simply to weight those young people or those in the oldest age group whom we have included in the sample may not improve our sample's estimates for the whole population, and may well make it worse (Moser and Kalton, 1971, pp 99-100).

III CLUSTERING

(i) Reducing costs over a large area

Sampling over a large area, or from a very dispersed population, can produce a sample which is very precise, but enormously costly in time and effort of data collection. Examples occur when field-work is necessary, and a widely scattered sample, or one which is evenly spread as a result of using systematic sampling, is difficult to survey. Similar problems occur when data are from maps or other sources, and not all can be bought or sorted through for the sample. In physical geographical studies, problems of access over large areas might make a scattered sample very hard to use.

A common solution is to use a sample which is deliberately grouped in a convenient number of small areas. In other words, we divide the frame into parts, and not all of them are sampled. Small scale projects rarely need to use cluster samples because they do not cover large areas.

Yates (1960, p.19) writes that 'it will be more accurate to take 10 per cent of all farms in each parish than to take all the farms in 10 per cent of the parishes', in other words, a clustered sample will be less good at estimating the population values than a sample of similar size drawn from throughout the sampling frame.

Exactly how detrimental will be the effect of clustering on the precision of the sample is hard to determine. It will not be the same for all variables, having the most adverse effects on the variables for which the clusters are relatively homogeneous.

The ratio of the standard error of any sample to the standard error of a simple random sample of the same size is the square root of what is usually called the design effect (the ratio of the variance of a sample to that of a simple random same of the same size). Since cluster sampling will almost always increase the standard error, the ratio will almost always be more than one while for stratified samples it will usually be less than one. Experience shows that with clustered samples it usually varies from 1 to 2, depending on the variable, and it is often around 1.5.

The calculation of the standard error of a clustered sample is complex, particularly when, as is usually the case, clustering has been combined with stratification. Rather than simply using the formulas for a simple random sample, it will be better to correct them by the 'rule of thumb' ratio of 1.5 to take account of the design effect. Instead of formula (2), we might therefore estimate the standard error by:

$$SE = \left\{ \frac{1.5\ s}{\sqrt{n}} \right\} \sqrt{(1-n/N)} \qquad (10)$$

where SE is the standard error of a sample where we expect the standard error to be 1.5 times that of a simple random sample of the same size, s is the standard deviation, and n is the sample size. The expression (1-n/N) will be ignored unless the sampling fraction is very large.

The formula for calculating the sample size needed would be altered to:

$$n = \left\{ \frac{1.5\ s\ .\ z}{c} \right\}^2 \qquad (11)$$

where n is the sample size needed, s is the standard deviation, z is the standard-error-unit measure for the desired confidence level, and c is the confidence limit.

Table 1 can still be used to give an approximation to the sample size needed if the result for the confidence limit 2/3 of the one desired is used. It can be seen that if we make the assumption that the ratio will be about 1.5, our sample size needs to be twice as large as a simple random sample to achieve the same precision. For example, at the 95% confidence level, if we want to be within 9 per cent of the true value, we would have to have a sample of 119 for a simple random sample, but 267 for a clustered sample.

Although the actual number used is arbitrary, multiplication by 1.5 will generally be better than ignoring the adverse effects of clustering on the ability of the sample to estimate the population values. The researcher should, however, endeavour to minimise the design effect.

The clusters should therefore be internally as heterogeneous as possible, /so that the omission of whole clusters from the sample will not remove unusual parts of the population and thus bias the final sample. In stratified sampling, the object is to create strata which are as different from each other as possible, since all will be included. The opposite principle is applied in creating clusters; each should be a microcosm of the population.

In many cases, clusters will be geographical sub-areas since that is why clustering is more economical and has to be used. If it is thought that these will be internally homogeneous, and different from each other, cluster sampling is not appropriate. Maximising heterogeneity and concentrating data collection in small areas are mutually contradictory for many of the features of interest to geographers.

(ii) Creating clusters

Clusters tend to be more homogeneous than the population as a whole, in other words, elements within them tend to exhibit high intra-cluster correlation. The procedures for creating clusters are designed to minimise the effects of this.

Larger clusters, from each of which a small number of elements are taken, will be best, because they are more likely to be heterogeneous.

The number of sampled elements in the cluster will be determined by the savings that can be made, since clustering is undertaken for economic reasons. If sample size can be increased by reducing travel time between elements, the clusters should contain a reasonable assignment for one field-worker or one day's work; a larger sample in one cluster would be unnecessary.

Heterogeneity can also be increased by skilful division of the sampling frame. In urban areas, for example, gains may be made by producing elongated clusters rather than compact ones, perhaps by aggregating wards; interviewing would still be facilitated, because the areas would be contiguous and delimited, but the possibility of a cluster containing a single house type, household type or 'social area' would be reduced.

If a list which is not spatially arranged is to be used as a sampling frame, and it will be prohibitively expensive to spread data collection throughout the area covered by the list, it can be arranged into clusters. The quickest way of doing this is to draw a sample several times the size of the required sample. These elements are then grouped into geographical areas each of which contains the same number of elements. These clusters can then be selected randomly, and the result will be a sample arranged in a more economical way, although clusters will clearly not be as compact as might be desired.

(iii) Selecting clusters

The simplest form of cluster sampling is when the population is divided into equal sized groups, and the resultant clusters randomly selected. All elements within these, or a fixed proportion within the clusters drawn randomly or systematically, are then included in the sample.

If clusters of different size are randomly selected, and a fixed sampling fraction is used, individuals in smaller clusters will have had a greater chance of selection than those in larger ones. Elements will have to be weighted by the size of the cluster from which they were drawn.

Where cluster sizes and sampling fractions both vary, two weights will be needed for each element. Firstly the variable sampling fraction is corrected by multiplying each element by the reciprocal of its cluster's sampling fraction, as in Table 2, and then a weight proportional to the size of the cluster is needed.

(iv) Sampling with probability proportional to size

Clusters are usually created to economise at the data collection stage, and the work-load in each is the smallest possible number of elements that will make these savings. It is also preferable for each field-worker to be allocated the same size of sample (except in circumstances where the work is expected to be substantially more difficult in some areas, because of terrain or greater difficulty in contacting respondents, perhaps). If the clusters vary greatly in size, the equalisation of the sample size in each can be achieved without weighting by selecting the clusters with a probability which depends on their size.

The need for weighting results from the fact that selecting uneven-sized clusters equally gives those in the smallest clusters a higher chance of selection. If the initial selection of clusters is dependent on the cluster size, this effect is compensated for.

The simplest method in practice is to allocate sequential numbers to the elements in each cluster. In other words, the clusters are listed with a cumulative total of elements within them.

Random or systematic sampling can be used to select from the list of cumulative numbers; the clusters within which numbers fall are used in the second stage of the sampling procedure. This is demonstrated in section IV, vi (Figure 1).

In a multi-stage sampling design which is designed to avoid weighting, every stage of selection except the last is made with probability proportional to size. For example, if constituencies are used as the first stage, and wards as the second, both should be chosen with probability proportional to their size (number of electors). An equal number of elements should be taken from each of the chosen wards.

The method requires that the population of each unit is known, or that any errors will be constant for all the clusters (for example, if the information on size is out of date, the error might be consistent if all the units have increased their population in the same proportion).

Sampling with .probability proportional to size is the most convenient form of cluster sampling.

IV SAMPLING FRAMES

(i) Convenient sampling frames

Where resources are limited or time is short, the rapid application of a sampling procedure to a readily obtainable sampling frame will greatly assist the project.

At the expense of reducing generality, it might be reasonable to limit a study to a sampling frame which can be easily obtained, and to devote proportionally more resources to other aspects of the data collection.

For example, pilot or exploratory work on time geography (Thrift, CATMOG 13 1977) might reasonably be limited to an institutional population, such as nurses in a home, if the co-operation of the appropriate authorities could be obtained. This might have the added advantage of reducing the expected variability in the sample (there are only a certain number of shifts) and thus reducing the necessary sample size.

Similarly, it might be possible to discover something useful about a small area, or about the users of a particular facility. Even an inexperienced researcher who had designed a sensible project might manage to gain access to membership records of a small organisation which felt its work might benefit from his study. Clearly, for example, a study of young people based on those who used youth clubs would be in no way representative of young people in general, but case studies are certainly not invalid because of their lack of generality.

(ii) Inconvenient sampling frames

If a researcher is interested in a population for which no listing or source is available, it will be necessary to investigate the possibility of constructing a sampling frame.

The compilation of large-scale sampling frames is complex and time-consuming. Where the population to be examined consists of a rare sub-group of a population some form of screening might be used. For example, to find a sample of freezer-owners to compare with other shoppers, either a short postal questionnaire or a brief interview might be necessary.

Screening is more economical when resources are concentrated in clusters, rather than spread out to create a frame over the whole area. An area of sufficient size should be screened or investigated to yield a sample of the desired size without the rejection of any, so that all the resources employed to identify the sampling frame would be used to the full. When the percentage with the rare characteristic is not known, clearly it must be estimated conservatively, with the aim of finding at least as many as are needed for the sample.

If at the preliminary stage there is doubt about whether an element is eligible or not, it should be included, because false inclusions can be removed at the data collection stage, but false exclusions will not be subsequently checked.

Screening involves a great deal of time and cost to find each element in the sample. Where each member of a research team wants to investigate different

sub-groups or features of a population, and could pool resources to identify the sampling frame, perhaps screening could yield a more justifiable return.

A similar process might be used where it is necessary to obtain a large amount of information from a small number of individuals, and a small amount from a large number. This might take the form of conducting short interviews and asking a sub-sample to complete diaries or undergo longer interviews. In other studies, it would be economical to utilise the differential variability of the information being collected by, in effect, having samples of different sizes, with the more uniform variables collected from only a small part of the larger sample from which the others would be collected. Farm size might be highly variable, crop yields much less so; this factor could be used in the sample design to speed up data collection by not determining yields on every farm. However, if one aim of the study were to discover if more extensive farms had higher or lower yields than smaller ones, the method would be inappropriate.

The arrangement of the list will determine the ease with which the 'first entry' method can be used. In rural areas, names may be listed in alphabetical order making it hard to work out the households; in urban areas, the listing is alphabetical within addresses, street by street. Problems arise when there are several surnames at the same address, who could be boarders, relatives, or separate single person households in bed-sitters. If a name is chosen and there is any doubt about whether others listed are members of his household, all the names at the address would have to be listed and checked by a field-worker, so that if the voter listed is not the first named in a household, he can be crossed off.

The definition of what constitutes a household will have to be rigorous; it is usual to include people who regularly live together and are catered for by the same person for at least one meal a day, but specialised surveys may need their own definition. The researcher will have to decide how to define any institutional households he wishes to exclude.

A supplementary frame can be used where the main frame is incomplete; an area of new building might be sampled from an address list. However, the sampling fraction would have to take into account the fact that the Register gives individuals, the list dwellings.

Another way of dealing with omissions is to use the half-open interval method. With each element drawn for the sample a note is made of the next element in the list. Any subsequently found to lie between them and not on the original frame are included. This works well for small amounts of new building, or where several households are found at one address. The method assumes that each element omitted is linked unambiguously to one included, and thus compensates for exclusions from the original frame. Problems arise when the spatial arrangement makes it difficult to ascertain to which included element an omitted ohe should be attached, and when the intermediate elements are very numerous, when an unacceptable level of clustering would arise. It also depends on field-work to identify all the exclusions.

(iv) Selecting within small clusters: the example of households

Sometimes our sampling frame constitutes, in effect, a list of small relatively homogeneous clusters. This would be the case with households

(possibly selected from the Electoral Register, as explained above) where we were really interested in individuals.

Often we would find that attitudes and behaviour were highly correlated for members of the same household, and it would be detrimental to the sample to include all the members of the household in the sample (section III, i). There is also an additional problem with interviewing more than one person at the same address where the first interview may influence the others, either by suggesting responses, or by taking up an unreasonable amount of time.

Therefore one member of the household is usually chosen at random. The problem is that we have to establish all the members of the household before making the selection; as well as being tiresome for the field-worker this procedure can be discouraging to the respondent and may prevent him from co-operating altogether.

If a rare part of a population, for example a particular crop, rock or group of people, is unevenly distributed, the reduction of generality to areas where the concentration is greatest might be essential in order not to dissipate energy searching out needles in the rest of the haystack - but the isolated part of the population might well have distinctive characteristics, and its exclusion would have to be explicit.

(iii) Sampling from lists: the example of the British Electoral Register

The ideal sampling frame would be a list which contained all the target population with no elements excluded and none included more than once. This is rarely encountered, but in many cases an imperfect frame can be adapted to serve. The British Electoral Register, the most commonly used sampling frame for studies of people in Britain, is by no means a perfect source, and the ways of using it illustrate the techniques adopted to deal with problems encountered more widely.

The Electoral Register lists all the adults aged eighteen and over eligible to vote who have complied with the legal obligation to complete registration forms on the basis of residence in October each year. Registers are available in public libraries or can be purchased from local authorities.

Although compiled in October, the Register is not published until the following spring, so that it will always be at least three months, and possibly nearly fifteen months, out of date. Towards the end of the life of the list (that is, in the early part of the year) alternative frames might be more attractive. It is possible to substitute a new occupant of the same address for someone who has died or moved, although it would have to be someone not already included on the Register, and randomly selected within the new household (see section IV, iv). If it is only the selected individual who is no longer there, no substitution can be made, because the others in the house would already have had a chance of selection.

Lists frequently contain a number of omissions: in the case of the Electoral Register, foreigners, immigrants or others not sure of their entitlement to vote, as well as people who have recently moved. The simplest solution is to ignore these deficiencies, and note in the research report any qualifications which result.

However, an alternative is available which may be preferred, particularly in areas of high mobility or when the list is out of date. It is more likely that one elector will have been left off than that all those in a household will have been. It is therefore a better source for households than for individuals.

We do not know how many voters are listed at each household. We therefore decide that the first entry on the list of any household will have to be chosen if that household is to be included. This gives each household only one chance of selection irrespective of the number of voters in it.

In order to know what sampling fraction to use when many of the names selected will be treated as ineligible because they are not the first listed, we need to know that on average there are 2.2 in each household, and we therefore start out by selecting 2.2 times the number of households we actually want in the sample.

It will be necessary to weight the answers for each individual by the reciprocal of the sampling fraction, within his cluster, that is, by the number of eligibles in the household divided by the number chosen. Someone who is the only eligible member of a household will therefore be weighted by 1/1, where there are two people eligible, the one chosen will be weighted by 2/1, where there are 3, 3/1, and so on. To avoid weighting, instead of using the 'first entry' selection method from the Electoral Register we include a household when any of its members is selected. This gives, in effect, selection with probability proportional to size (number of voters). Taking one from each household prevents the need for weighting (III, iv).

TABLE 3

Kish's tables for selecting one individual from a household

	Rows are allocated in the proportions:	Number of eligible individuals found: 1	2	3	4	5	6 or more
		select individual number:					
Row I	1/6	1	1	1	1	1	1
Row II	1/12	1	1	1	1	2	2
Row III	1/12	1	1	1	2	2	2
Row IV	1/6	1	1	2	2	3	3
Row V	1/6	1	2	2	3	4	4
Row VI	1/12	1	2	3	3	3	5
Row VII	1/12	1	2	3	4	5	5
Row VIII	1/6	1	2	3	4	5	6

An impartial selection within the household can be made by using Table 3, devised by Kish (1965, pp. 398-400). The interviewer sets out with

a list of households, to each of which is attached one row of the Table. The rows are allocated in the proportions set out in the first column, so that if an interviewer went to twelve addresses, two would have the first row, one the second, one the third, and so on. When the field-worker has listed all the eligible individuals in the house in a pre-determined order (usually by age), he selects for interview the one with the number indicated for that household size by the row of the Table given to that address. For example, if there were four people eligible in a household, and the address had been allocated row V, individual number three would be interviewed; using the same row, if there were five in the house, number four would be interviewed, if three, number two.

(v) Constructing frames from maps

Published maps or aerial photographs can be used to create lists of a variety of phenomena for which no suitable sampling frame already exists, such as vegetation, water courses, shops, factories and dwellings. Clearly there will be errors due to inaccuracy of the source map, changes such as demolition and construction or disturbance of vegetation, and misinterpretation. Comparison of the map with the area might indicate the extent of these problems.

Some phenomena can be identified readily from maps, but care would have to be taken over multiple-occupied premises, moveable stalls, mobile shops, and so on.

Specialised maps may name all the elements to be included. An example of this type of map is the series of Goad shopping centre maps, on which each shop is named and its main functions indicated (Rowley and Shepherd, 1976). Such maps might be preferable to directories if their coverage were more complete, and if a sample were desired with a spatial spread. One way would be to number all the units shown by a 'postman's walk' route, and then to use a form of systematic sampling.

Where such a map is out of date, it might be possible to up-date it without re-surveying the area. If, for example, a shop had been sub-divided into smaller units, all should be included in the sample if the original was included. The half-open interval method can be used for new units inserted, provided that they can be unambiguously linked to their neighbour on one side. The removal of a unit, such as a closure, would mean that element would be considered as an ineligible blank. A sufficient size of sample would have to be drawn to allow for some deletions in this way, because it would be quite incorrect to substitute a neighbouring unit which would then have an extra chance of selection.

(vi) Constructing frames by mapping

Where there are no obtainable maps which mark the population to be sampled and are up to date, a researcher may have to construct a sampling frame on the ground, or at least amend the existing maps in order to construct such a frame.

Buildings such as industrial or commercial premises in a small area, dwellings in rural areas, new buildings, or in countries where there is no suitable sampling frame for households, a listing may be prepared by mapping.

Clear procedures and forms for recording information will be needed,

and Kish (1965, pp.322-351) discusses the instructions needed to list buildings for such a frame. The area is usually divided into segments delimited by natural features or streets, clarified by sketch maps where necessary.

Complete enumeration of every element in the population may be possible, but will often be too costly and time-consuming. The segments may be used as clusters, and only those included in the sample need be mapped fully.

Clusters are most conveniently sampled with probability proportional to size, to avoid weighting (section III, iv). If population sizes for the segments are known, this method can be used, and then the same number of elements would be sampled from each one.

Where segments contain fewer than 6 or 8 elements, it may be more economical not to use probability proportional to size sampling, but to include all the elements in the selected clusters. The value for each element would have to be weighted by the size of the cluster from which it was drawn, but this may be acceptable if the segments vary little in size. The advantage is that there is no need to list all the elements first and then select within them.

If the selected clusters were subsequently found to contain only slightly more or slightly fewer elements than the enumerator first thought, there will be little difficulty. If more were to be found than expected, all should be included in the sample.

Figure 1 illustrates the kind of sheet that could be used to record the segments, in a convenient form for the selection of clusters with probability proportional to size. Random numbers would be drawn once the listing was complete, and whichever clusters they were in would be the ones included in the sample. In the example shown, if the numbers 9 and 21 were drawn, the clusters selected would be segments 2 and 5. If the same number of elements was chosen from each, the sample would be self-weighting (section III, iv). In this example, two elements would probably be included from each. The disadvantage of the method is that the chosen segments have to be visited again, and the elements within them listed.

It would be possible for the identifying numbers for the clusters to be drawn in advance, so that the field-worker could proceed directly to data collection after the listing of segments. In this case, to avoid conscious or unconscious 'cheating', he should not have the random numbers to hand when listing the segments, but should carry them with him in an envelope to be opened only after all the first stage had been completed. In some situations, particularly where the area to be covered is compact and the data to be collected readily available (as might be the case with observations rather than interviews), field-work could be conducted rapidly by this method.

(vii) Longitudinal studies

Where studies are designed to measure changes over time, it may be necessary to re-visit measuring points or a panel of respondents. This assumes that the process of measurement does not destroy or affect the variables being studied, as trampling vegetation or disturbing the soil does. A different sort of change takes place in people who have been interviewed once, for they

Segment Number		Approx. No. Of Dwellings	Cumulative Number
1	Church Row	5	1-5
2	Behind church (sketch 1)	8	6-13
3	Behind church (sketch 2)	4	14-17
4	Green north	3	18-20
5	Green West (including vicarage)	5	21-25
. . . .			

Fig. 1 Listing Segments: An Example

may become more conscious of their actions, or be conditioned to respond in a particular way. They may therefore be very different from the population they are supposed to represent at the subsequent interview.

With human populations, unless the time period between the interviews is very short, there will be substantial 'attrition' of the sample as people move out of contact or die. Substituting its replacement for a household which has moved is sometimes used, although it may not be the case that the new inhabitants resemble the old, and the method is anyway inapplicable to individual respondents. The response rate on the second interview is also likely to be lower than on the first, unless the interview is very enjoyable or the sample 'captive', as in the case of a school class.

To avoid repeating the measurement on the same individual or at exactly the same point, some studies select two samples with the second sample adjacent to the first. In the case of physical features, this method seems legitimate, but it may not be safe to assume that people will be like their next-door neighbours.

Another method of studying change, therefore, is to sample afresh. The problem is that each sample may have such large sampling error that it is impossible to distinguish differences due to that from real trends in the data. It is more reliable, although more costly, with large-scale longitudinal studies, to draw an initial master sample several times larger than the sample needed each time. After the initial observation, sub-samples are drawn from it for subsequent phases. Each individual is only examined twice, which reduces the effects of 'contamination' due to observation. Moser and Kalton (1971, pp 137-145) discuss panel studies and longitudinal surveys for human populations; their suggestions have applications for other studies.

(viii) Sampling continuous flows

When records are constantly being added to a card index, or people are passing a point, the researcher cannot start off with a complete sampling frame and then make his selection. The only feasible sample will be systematic, taking every kth element, or every (k ± r)th, where r is a small random number. The systematic sample incorporates the advantages of coverage through time, representing early and late arrivals equally.

It is not possible, however, to pre-determine the sample size, unless the rate of flow can be precisely predicted. The interval k will have to be guessed on the best available evidence.

The processing of individuals in the sample may hinder the counting of others. Counting and interviewing every kth shopper to pass through a supermarket checkout is a difficult task for this reason, and there will certainly have to be someone counting as well as sufficient people interviewing. This will inevitably mean that field-workers will be standing around for a large proportion of the time, and what seems superficially a straightforward and economical way of sampling is in fact less practical. It may be necessary to interview half as many in the busiest time periods, for example, and then weight them accordingly.

Another problem with sampling sequential flows is that what is being sampled is passage past a point or entry onto a list. If a researcher was really interested in users of a particular facility, rather than usage, the visits might have to be weighted to represent the users. The chance of selection of an individual would depend on the frequency of his use of the facility, so the results should be weighted by the reciprocal of the frequency of use to produce an unbiased estimate of the users. This means that the results for someone who uses the facility fortnightly should be given double the weight of those of a weekly user.

There is also the general problem that the survey may cover an unrepresentative time and may need to be repeated (section V, iii); different time periods might be treated as strata (section II, iv).

V AREAL SAMPLING METHODS

(i) Frames and methods

It might be expected that geographers would make considerable use of areal sampling techniques, but especially in human geography samples are generally drawn from frames which seldom have more than an incidental spatial aspect. Land use, vegetation, soils and other continuous phenomena are, however, sometimes sampled from maps or aerial photographs, and similar sampling methods may be used on a small scale in the field.

Many of the spatial sampling methods are directly analogous to the a-spatial methods already discussed.

Normally we use a grid system with co-ordinates, preferably one already present on the map as in the case of the National Grid on British Ordnance Survey maps. If the area is of irregular shape, the grid must overlap the study area, as shown in Figures 2 to 7. The grid reference for each point serves the same function as the identifying number in a list.

Figure 2 shows the spatial version of the simple random sample, using points as elements. Random number tables are used to provide co-ordinates on both axes. Points selected which fall outside the study area, in this case in the sea, are rejected, just as any blanks in a sampling frame would be.

This form of sampling is used to estimate land use on the assumption that the points will fall on each land use type in its correct proportions; the percentage of points on waste ground, for example, will give the percentage of the whole area covered by waste land.

It is likely that waste ground will not be distributed across the whole area; a small random sample might miss it altogether. Most work on an areal basis requires even coverage, and a systematic point sample (Figure 3) is therefore often used. This method is easier to use in the field than random sampling, because the field-worker can move easily from one point where he has to examine vegetation or soil to the next, perhaps using chains to measure the distance. Just as in sampling from lists, though, there is a danger that the interval between the points will correspond to some periodicity in the data, as it might with land forms, or man-made landscapes, or, within a field, the spacing of crops (Zarkovich, 1966).

Many of the methods of combining random and systematic sampling can be adapted for the spatial situation. If the area is divided into parts, a quota can be set for each which, when exceeded, leads to rejection and the selection of other points. Alternatively, a point can be selected at random in each square of a regular grid (Figure 4).

Berry and Baker (1968) advocate a stratified systematic unaligned sampling method (Figure 5), which utilises grid squares each with its own internal grid co-ordinates. We start by drawing two random co-ordinates, a and b, within the first square. The marginal grid squares along the top row are then allocated points by drawing a random y co-ordinate within each square and using the x co-ordinate, a, from the first square. The y co-ordinate, b, is used with a random x co-ordinate in each grid square of the marginal column. Each subsequent square's sample point is determined by the y co-ordinate of its column and the x co-ordinate of its row.

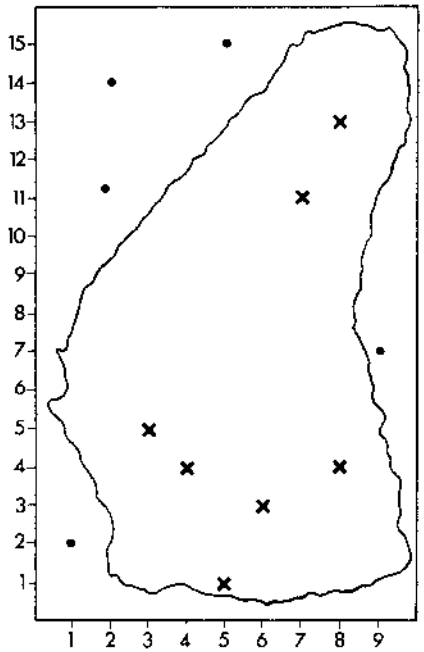

Fig. 2 Simple Random Sample

x Included
. Excluded

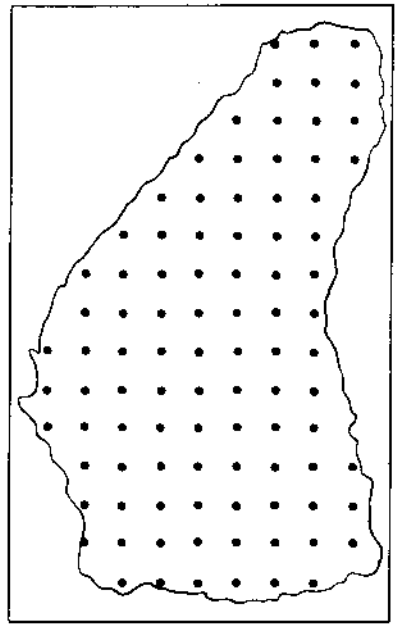

Fig. 3 Systematic Point Sample

While this method is not strictly one in which each point is selected independently, for all practical purposes it can be assumed to be a stratified sample with independent selection within the strata (the grid squares); Berry and Baker have demonstrated the successful use of this form of sample for land use proportions.

If part of the area, for example a soil or land use type, is very small or variable, and is to be compared with a larger and more uniform area, stratification with different sampling rates may be used. A larger quota of random numbers would be allocated to the smaller or more variable stratum.

The quickest way to draw the sample may be to take random numbers within grid squares. Approximate demarcation between the strata may be adequate; squares could be allocated to the more variable stratum if they contain more than 50 per cent of the more variable feature. This misallocation of the fringes may not matter if the grid squares are small relative to the size of the strata (Figure 6).

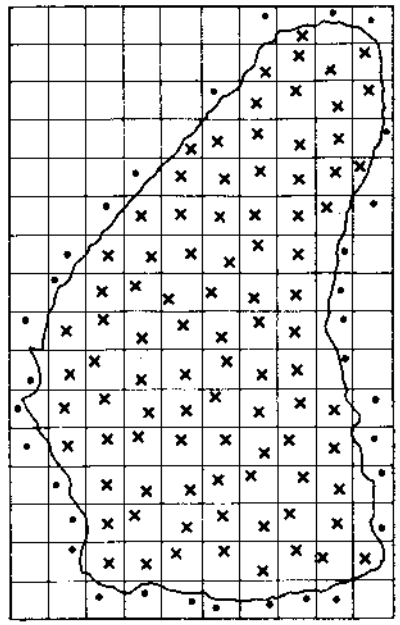

Fig. 4 Systematic Random Sample

x Included
. Excluded

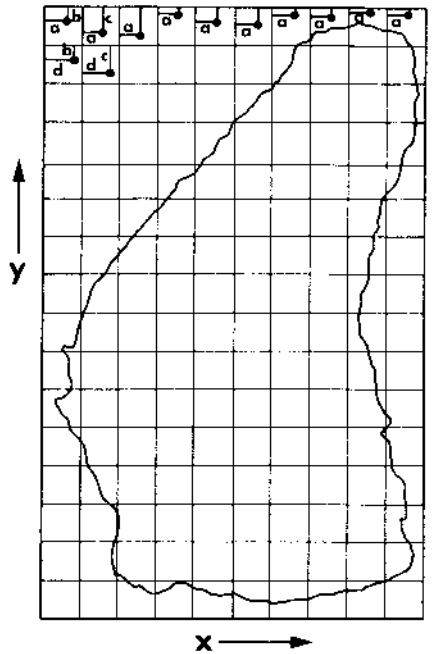

Fig. 5 Stratified Systematic Unaligned Sample

(For explanation see text)

Figure 7 shows the simplest form of random transect sample, in which lines are drawn across the area between pairs of random coordinates. It is assumed that the area of phenomena will be in proportion to the length of line crossing them. The percentage of each land use can therefore be simply assessed by measuring distances along the transects.

For a systematic transect sample, the transects are arranged across the study area, usually as a set of parallel lines. Purposively located transects are also commonly used; rather than using probability sampling the transects are sited by the researcher across the contours, or to correspond with an 'environmental gradient'. If we were to position our transects at right angles to the sea, at regular intervals, we would not have a probability sample, and would have to make this clear in discussion of the techniques of our project. However, along these transects we might use a random or systematic sampling procedure, treating the initial transects as sampling frames. Transects have the advantage that they can be illustrated readily by a cross-sectional diagram or a series of diagrams showing such features as topography, land use or vegetation along their length.

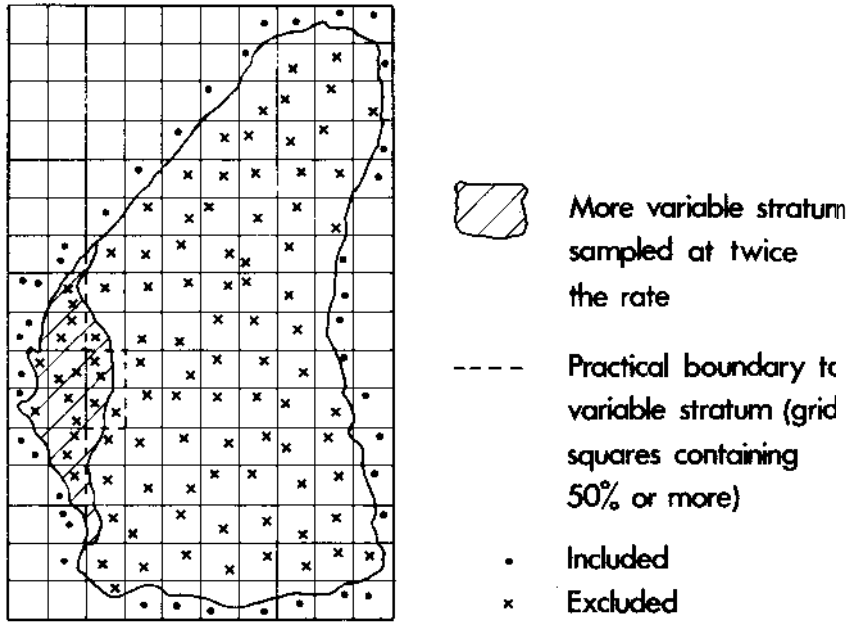

Fig. 6 Stratified Systematic Random Sample

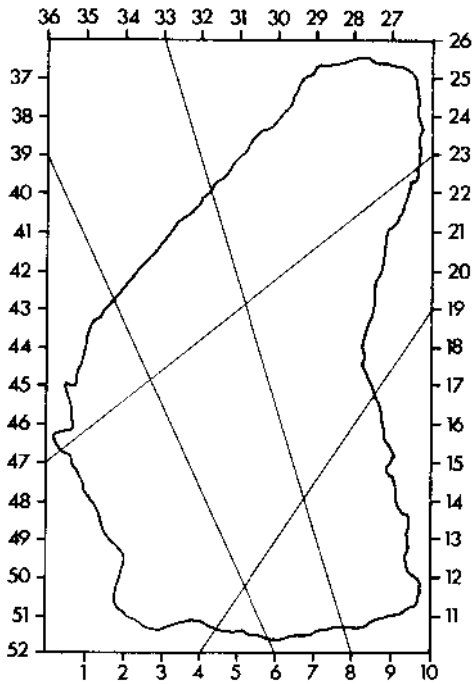

Fig. 7 Random Transect Sample

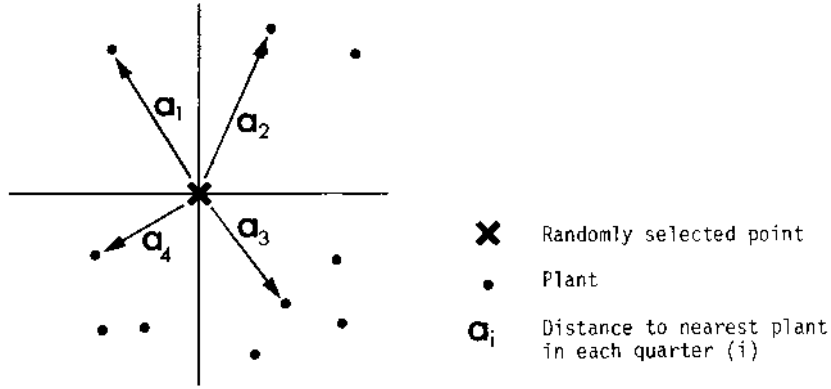

Fig. 8 Estimating Species Density By The Method of Cottam et al (1953)

All these spatial samples are easily applied to maps, but are more difficult in the field. Trying to identify points on the ground, or to follow straight lines, may be rendered impossible even without such obstacles as buildings, termite mounds or impassable boundaries. The samples normally used in the field are systematic rather than random, because it is convenient to relate each observation point to the last.

Biogeographers in particular make considerable use of quadrat sampling, which is not strictly a sampling method, but a data collection method at points chosen by one of the areal methods discussed. A quadrat is a rigid frame of fixed area and shape placed on the ground, for example with one corner touching the chosen point; the vegetation, soil or crop within the frame is studied.

The shape and size of the quadrat will to some extent determine the results. There seems no generally agreed improvement on the square quadrat, but the size has to be chosen for the particular features being studied. Examples and advice on choosing suitable sizes are given in a number of references (Peltier, 1962; Haggett, 1965, pp.198-199; Shimwell, 1971; Kershaw, 1973). A large number of small quadrats seems better than a small number of large ones, because they can be more spread out, but they have longer edges and therefore create more problems of determining inclusion or exclusions. Shimwell (1971, p.17) discusses the question of the 'minimal area needed for vegetation studies, the smallest area in which all species in the area will be found, and which is thus 'large enough to represent the characteristic structure and floristics of a plant community'.

Extra problems arise in the use of quadrats for crops, because if a field is planted in rows, it might be possible to place the quadrat in such a way as to contain either x rows, or x + 1. The solution clearly is to take several randomly aligned quadrats from a field. A decision has to be made about the border of the field, for the crop may be sparser at the margin. If quadrats

which overlap the boundary are excluded, there will be a biased over-estimate of yield (Zarkovich, 1966, pp.331-365).

Quadrat measurements can be taken at randomly or systematically located points, or along transects. They cannot legitimately be located by throwing objects at random', for this method will never approximate to a true random sample. Even if a field-worker is genuinely concerned not to direct his throw, the vegetation height may affect the point where the object falls, and points around the border of the study area might also be under-represented (particularly as a long throw may make the object hard to find!).

Not all vegetation studies use quadrats; there are alternative methods which have particular applications. For example, the contact between species can be determined by sampling points by any of the usual methods, recording the plant at each point, and any plants touching it; this produces an 'association matrix' between species (Yarranton, 1966) although it ignores association in which plants do not 'touch'. Recording the distance from the sample point to the nearest plant of a particular species in each of four 'quarters' (Figure 8) gives an indication of the density of the species and may be quicker than estimating coverage by a large random sample (Cottam et al., 1953).

In the field, particularly when a quick sampling method is required in open country, use has been made of the random walk to find sampling points, often points at which to take quadrat measurements, or as an alternative to transects. From a random start, the observer moves a set distance (or number of paces) in a direction randomly determined for a predetermined number of 'legs'. Such a method might give very poor coverage. The use of random walks in urban areas is discussed in section VI, iii.

(ii) Problems of spatial frames

The greatest problems with spatial sampling are concerned with access. There is a temptation to ignore areas which have the most difficult terrain, or to place transects across accessible points. Coastal vegetation studies have sometimes been limited to points at which landings could feasibly be made. These are not then probability samples, and the researcher would have to use his judgement to guess how different the inaccessible areas might be.

Similarly, a rigorous sample might be drawn from a very incomplete frame. The target population might include all the areas with a particular soil type, but the sample might be restricted to areas not built upon. The research report would have to indicate such deficiencies, and likely inadequacies would need to be discussed. Rigorous sampling within the accessible areas could not compensate for the initial lack of a probability sample, but as a case study the work would not be invalidated.

Interaction between the researcher and the study area can be detrimental, for example trampling across an area undoubtedly effects its vegetation. The problem might not be too severe the first time an area is studied, but if the field-work is to be repeated later, there could be problems analogous to the 'polluting', 'conditioning' or mortality of a human population. Placing markers to indicate points to be returned to is particularly likely to affect the vegetation, especially if the markers are posts on which birds may perch. In such cases, sample points can be a set distance from their markers.

One mistake is to use a spatial sample for a phenomenon which is not continuous. If random points were located, and the sample were then to consist of the nearest oak tree, or the closest farm building, then the method would be biased in favour of the more scattered, isolated elements.

If we were to include elements if a point fell on them, buildings or farms with large ground areas might be more likely to be included. The best method for phenomena such as these is to list (or mark) all the elements. For economy, we could do this in parts of the study area chosen by a probability method, and sample within these.

Occasionally, where phenomena are relatively evenly spread, as in the case of shops in a shopping centre or factories on an industrial estate, elements are allocated to a single dimensionless point. This might be their southwest corner, or their central point. The nearest one to each selected sampling point in a given direction (and perhaps within a specified distance) might then be chosen. The larger units would have a higher chance of selection by this method, but if their size does not vary greatly, and a speedy selection is required, the method might sometimes be justifiable.

(iii) Sampling in space and time

Sampling methods are sometimes used to locate points in time and space for traffic counts or stream flows, or measures of a variety of phenomena which are to be repeated.

Traffic counts are usually performed at purposively selected sites, for example on all major roads leading into a settlement, but occasionally they may be evenly spaced every k kilometres. Yates (1960, pp.363-364) discusses this use of sampling.

Occasionally it is necessary to divide the measurement period in order to cover a larger number of sites than there are field-workers. This might occur when studying atmospheric pollution at a variety of locations over a week (Haggett, 1975, pp.541-543).or when interviewing people using a number of public libraries during different opening shifts (SCPR, 1973, pp.78-80). In both cases a grid would be drawn up which ensured that each point was sampled the same number of times, at different time periods, for example, in the library case, in both morning, afternoon and evening equally.

Distinct time periods may be regarded as strata and sampled accordingly. However, when a researcher divides his sample between different times and places, he is in effect creating a clustered sample design. As was explained in chapter III, such a sample probably only includes part of the variability of the population as a whole. In order to counteract adverse effect, the largest possible number of small clusters (sampling points and times) should be used as will be economic.

Geographers might perhaps give more attention to choice of sampling time, because both time of day and season will affect certain kinds of data.

VI NON-PROBABILITY SAMPLES

(i) Purposive samples

When a geographer selects a 'typical' example, or set of examples, from the total population he purports to study, this sample cannot be analysed as if it were a probability sample. No matter how carefully chosen, it clearly cannot conform to the requirement of probability sampling that each element has a known probability of selection, and therefore estimates of standard error and confidence limits cannot be produced.

In fact, the purposive sample drawn in this way may well contain the extreme cases, or those which best illustrate the point the researcher wishes to make.

Within purposively chosen samples it is possible to use probability methods. Case studies of particular places are of this type. Transects of vegetation or topography can be purposively located, but probability samples drawn within them; data collected in origin and destination surveys, taken at deliberately selected points where the work can be undertaken, may be analysed quantitatively. In these cases, the researcher would make it clear exactly what the limitations to the study were.

(ii) Quota samples

If no convenient sampling frame exists, and costs or time preclude the creation of one, it may be permissible to include elements so long as they have certain specified characteristics of the target population. If the proportions in which some features occur in the target population are known, these features are used as quota controls, and elements are selected by the most convenient method until the number in each quota is complete.

The method is usually used in samples of people, and particularly in market research, but could be used for other phenomena. Quota samples depend on accurate knowledge of the target population. It is known, for example, from the Census how the population of Britain (or of particular areas of the country) divides into age and sex groups. If the variables in which the research project is interested are highly correlated with these variables whose distribution is known, a sample in which they are represented in the correct proportions can be expected to resemble the population. However, we are usually interested in variables that are not highly correlated with the quota control variables, and we can rarely assume this correlation; therefore the method can lead to a very misleading sample.

The bias stemming from interviewer or respondent selection could be considerable. If a researcher was interested in users of a particular central facility who came from a particular part of the town, he could approach people using the facility and ask them where they came from. If the resultant sample consisted wholly of people under 30 he would have no way of knowing if this was typical of users of the place, or if he had chosen an unrepresentative time of day to carry out the survey, or if he had failed to approach all sorts of people.

Quotas are established in one of two forms, depending on whether the variables used as controls are given in an independent or interrelated form.

In the interrelated form, the interviewer is given specific categories of people to find, and told how many of each to interview. For example, a quota sample of young people might control for occupation and sex. The quota would then consist of a specified number of young unemployed males, young unemployed females, young workers of each sex, and young people continuing their education of each sex, giving six separate categories with a target number for each. This would prevent the interviewer confining his attention to the most accessible young people, for example by interviewing a large number of unemployed or school pupils.

When the controls are independent, the interviewer would be told simply how many of each sex, and how many in each occupation category to interview, but not how many in each of the six categories. It would therefore be possible for one category, for example, working girls, to be omitted altogether while fulfilling the marginal totals. The interrelated form is therefore often preferable, but it depends on the researcher knowing how many of each category there are in the population, and the particular variables he is using may not be cross-tabulated anywhere.

It may be possible to determine quotas appropriate to a particular study, and indeed a researcher may feel confident that in doing so nothing is lost by this method. It is also possible to set quotas which over-sample particularly variable parts of the population, in the way that particular strata can be over-sampled in probability methods.

When an interviewer has to approach individuals in the street, there is a great temptation to omit certain people altogether. It is difficult to approach someone and ask them the questions necessary to establish if they are in the quota; asking age and occupation does not usually start off an interview well. Similarly, it requires a certain amount of tact to refuse to interview someone who does not happen to fit into a required category, and there may even be a temptation to squeeze a co-operative but ineligible person into an adjacent classification. Most interviewers are reluctant to approach someone and then reject them, and consequently will approach people who look like the stereotypes of the set categories. For example, if age groups are given, the interviewers are likely to produce a sample which clusters in the middle of the age ranges, although this is less likely when the quota controls are not linked. If interviewing is conducted door-to-door , the problem of excluding people not on the street is overcome, but the sample may instead be biased towards those at home, such as housewives or the housebound. Bias can never be removed from quota samples, but their greatest drawback is that we can never know how severe it is in any given sample.

(iii) Random routes

When there is no sampling frame, and the researcher wants to sample households or buildings of some kind, it may be quick and cheap to sample by means of a set of rules governing a walk around the study area.

A random start is chosen, and then instructions are followed, for example, first right, first left, first right, calling at addresses at a fixed interval along the way.

Instructions should be sufficiently unambiguous for two people following the same route to make the same decisions about where to turn and which addresses to count, but in practice seldom are. Decisions will be needed about maisonettes, tower blocks, multiple-occupied dwellings and cul-de-sacs. The technique is attractive in that it dispenses with a sampling frame, but it is open to misapplication, and the creation of a sampling frame would be the only way to ensure that buildings obscured from the road, or houses containing more than one address, were included.

The selected addresses should be treated as any other sample, and if there is no reply, they will have to be visited again. Random walks depend very much on the integrity of the practitioner, and are at best quasi-probability samples.

(iv) Combining quota sampling with a random route

To ensure that the initial selection of respondents is relatively unbiased, and that if a sample is to represent residents of a place, they are not all contacted in one part of it, interviewers can be sent to randomly located starting points and told to follow random route instructions, approaching households at a fixed interval, and interviewing until the quotas are full.

The interviewer's honesty is greatly strained, particularly if told to call at every kth house, because if the house next door seems more likely to fulfil the quota, it will be very tempting to include it. Permitting the interviewer to call at every house along the random route is preferable, and if the quota is for a comparatively rare trait, then the level of clustering may not be too great; after a set number of eligibles have been found, the interviewer can be told to move to a new point.

The method suffers from the disadvantage of quota sampling in general, that those who are not in the right place at the right time never get included. It might have applications for immobile populations such as morphological features or buildings.

(v) Snowball samples

One way of creating an approximation to a sampling frame for a rare trait in human populations is to ask a number of initial contacts for the names of any other eligible individuals, who are then approached and asked the same question. This has been successfully used for deaf people, or players of particular sports. Ideally the initial contacts should be randomly located, and the trait should be one which leads to a high degree of contact. The 'snowball' should be allowed to grow until each respondent is producing a list which consists largely of those already identified, but even then there is no guarantee that there is not a substantial number of isolated individuals, or a self-contained group not in touch with those found by the snowball.

Snowball sampling has limited applications, but where a researcher felt confident in the enclosed nature of the rare attribute qualifying for inclusion in the sample, and no sampling frame was available, it might be the only feasible method.

(vi) Justifying non-probability samples

If a probability sample can be used, in other words, if a reasonable sampling frame exists or can be created, it is hard to justify using a non-probability method. It is impossible to establish the absence or extent of bias in these methods, and the statistical tests applied to probability samples are invalid.

Their speed, cheapness and the ability to dispense with a sampling frame may occasionally make their use essential, and they may be justified in pilot or exploratory work. Paradoxically, though, they rely on experience and a certain amount of knowledge of the target population - which the researcher probably lacks when he is embarking on a pilot study. Whenever possible, probability methods should be used.

VII RESEARCH DESIGN AND THE CHOICE OF SAMPLING METHOD

The cost of using a sampling method is a combination of the cost of collecting the measurements, and of the frame; the time taken is similarly a combination of the time spent in the field and the time spent beforehand obtaining the frame and drawing the sample. It is these costs and time limits, as well as constraints at the analysis stage, which will largely determine the sample size and the methods which are feasible.

After fixing the relevant population and properties of it which are to be studied, and the method of measuring these characteristics (interviews, observations, measurements), we have to decide the approximate size of sample, the number of units, which can be studied. This size can be increased by reducing costs or time, usually by clustering, but this will increase the sampling error.

Conversely, stratification may be more costly, and lead to a reduced sample size, but to an increase in precision which is effectively equivalent to an increase in sample size. It is possible to cost ways of reducing the sampling error to see which are feasible.

Therefore, the initial decision about sample size may be modified as the research design proceeds, but it will be a major consideration and a rough approximation will be needed from the outset.

Figure 9 shows a flow diagram for choosing between probability and non-probability methods. It is only after discovering the absence of a sampling frame, and deciding that it is not feasible to amend any existing frame or create a new one, that non-probability methods should be resorted to. Having established a sampling frame, decisions about the form of probability sample are set out in Figure 10. There will be applications for all five methods at the bottom of the chart, and their use will be largely determined by the frame. A degree of clustering may be needed to reduce costs, and can be combined with other methods. Complex sample designs are harder to use, and calculations of standard errors are made more difficult. In small scale studies elaborate sample designs will rarely be needed.

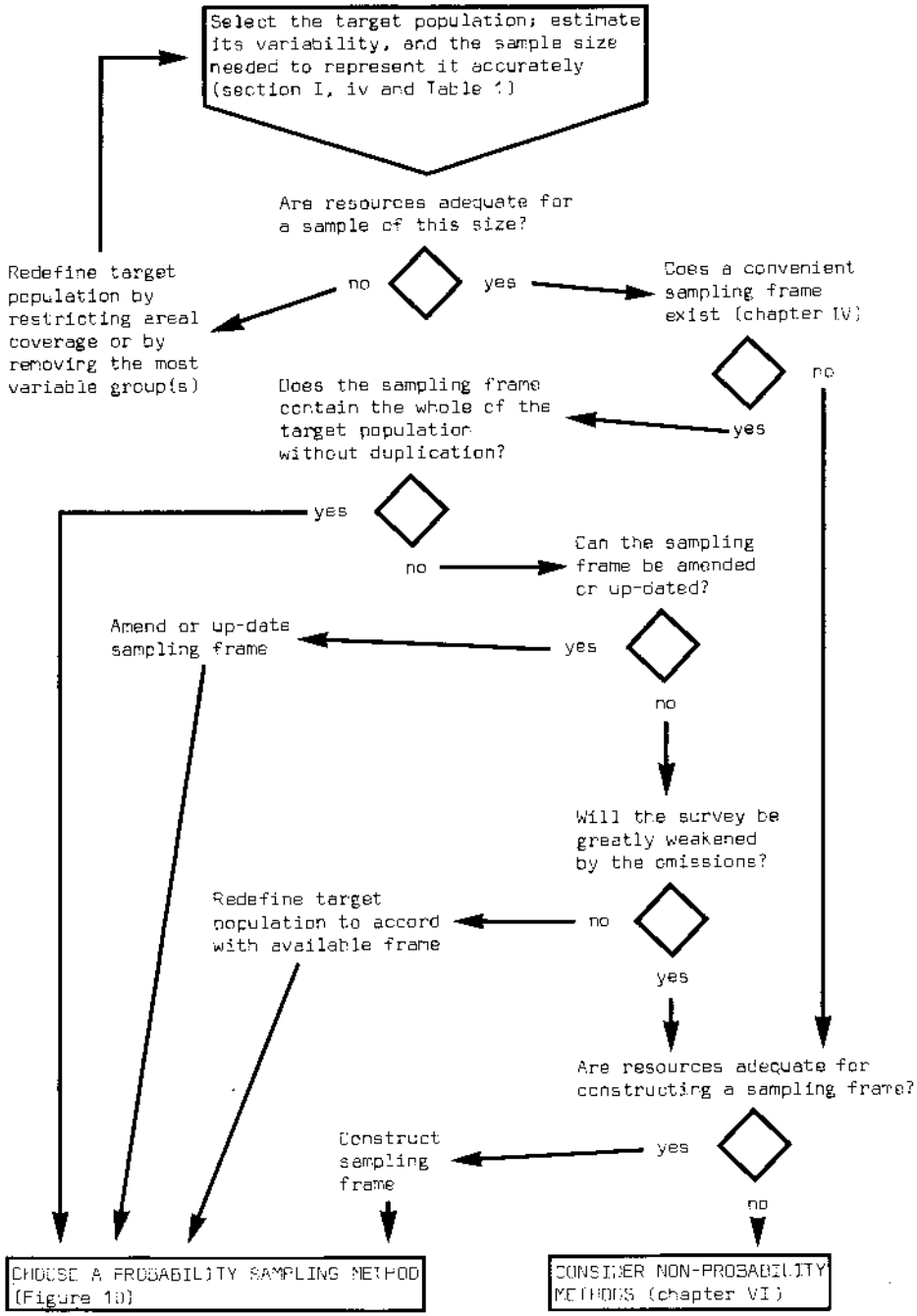

Fig. 9 The Process of Sample Design: probability and non-probability methods

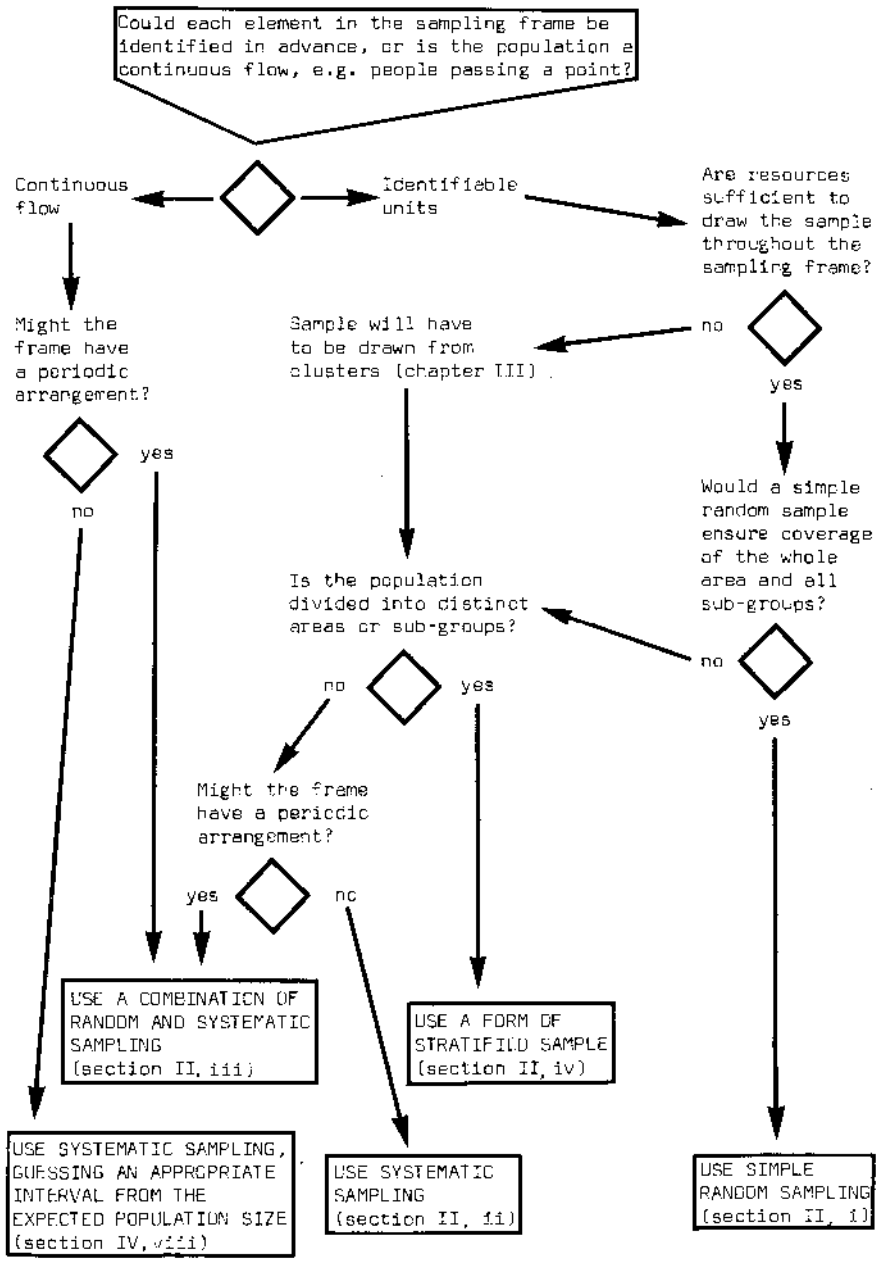

Fig. 10 The Process of Sample Design: Choosing between probability methods

The greater the knowledge the researcher has of the population he is to study, and of the frame from which he is to draw the sample, the better able he is to take decisions about the research design. But having made the various choices, it is necessary to set out rigorous rules of procedure and follow them, and finally, in any research report, to state exactly what they were, and if there were any problems with them. The simplest and smallest piece of research needs to have its methods clearly stated to enable anyone reading it to understand its strengths and weaknesses.

BIBLIOGRAPHY

A. Basic texts on sampling

Moser, C. and Kalton, G. (1971) *Survey methods in social investigation.* (London: Heinemann), 61-210.

Social and Community Planning Research, SCPR (1973) *Sample design and selection.* (Technical Manual No.2).

Stuart, A. (1962) *Basic ideas of scientific sampling.* (London: Griffin).

B. More Advanced texts

Kish, L. (1965) *survey sampling.* (New York: Wiley).

Yates, F. (1960) *Sampling methods for censuses and surveys.* (London: Griffin).

Cochran, W. (1963) *Sampling techniques.* (New York: Wiley).

C. Sampling frames for human geography in Britain

Gray, S. (1971) *The Electoral Register.* Government Social Survey M151, (London: HMSO).

Pickett, K. (1974) *sources of official data.* (London: Longmans), 115-131.

Rowley, G. and Shepherd, P. (1976) A source of elementary spatial data for town centre research in Britain. *Area,* 8(3), 201-208.

Davies, W., Giggs, J. and Herbert, D. (1968) Directories, rate books and the commercial structure of towns. *Geography,* 53(1), 41-54.

Oliver, J. (1964) Directories and their use in geographical enquiry. *Geography,* 49(4).

D. Introductions to areal sampling

Chorley, R. (1966) The application of statistical methods. in: *Essays in Geomorphology,* G. Dury, (London: Heinemann), 275-387, especially 292-307.

Haggett, P. (1965) *Locational analysis in human geography.* (London: Edward Arnold), chapter 7.

Kershaw, K. (1973) *Quantitative and dynamic plant ecology.* (London: Edward Arnold), 28-37 - particularly deals with quadrats and transects.

Shimwell, D. (1971) *The description and classification of vegetation.* (London: Sidgwick and Jackson), 4-23 - particularly quadrats, transects and 'plotless sampling'.

E. Examples of areal sampling

Berry, B. and Baker, A. (1968) Geographic sampling. in: *Spatial analysis,* B. Berry and D. Marble, (New Jersey: Prentice-Hall).

Cottam, G., Curtis, J. and Hale, B. (1953) Some sampling characteristics of a population of randomly dispersed individuals. *Ecology,* 34, 741-757.

Peltier, L. (1962) Area sampling for terrain analysis. *Professional Geographer,* 14(2), 24-28.

Yarranton, G. (1966) A plotless method of sampling vegetation. *Journal of Ecology,* 54, 229-237.

Zarkovich, S. (1966) *Quality of statistical data.* (Rome: Food and Agriculture Organisation of the United Nations).

F. Sampling in space and time

Haggett, P. (1975) *Geography: a modern synthesis.* (New York: Harper and Row) 2nd edn. 541-543.

G. Quota sampling

Moser, C. and Kalton, G. (1971) *Survey methods in social investigation.* (London: Heinemann), 127-137.

Stephan, F. and McCarthy, P. (1963) *Sampling opinions.* (New York: Wiley).