SOME USE OF STATISTICAL SAMPLING FOR BUSINESS DECISIONS

Gernold Frank
University of Frankfurt
Frankfurt, West Germany

At the University of Frankfurt, West Germany, there exists an own idea of teaching statistics, which was founded at the beginning of this century. It is connected to university teachers as e.g. F. Zizek, P. Flaskämper, H. Hartwig, A. Blind or G. Menges – only mentioned those who are not teaching anymore nowadays. But, what are the specialities or the ideas? Ahead of all this traditional teaching there is the claim: before new methods etc. are developed the request for their economic and social understanding has to be answered. This so-caled "Frankfurtarian School of Social-Statistics" therefore

- asks, first, for the reason than for the methods and

- asks for the contents of the data.

Or, to sketch it in a comprimated goal, every time there is an asking for the "adequation" of the methods, which should be used for an application as well as for the data.

Aside the university a teaching programme was installed some years ago for adults having already a university degree (mainly in economics) how and why to use sampling techniques. This programme was worked out for business decisions as well as there is a co-operating with people being engaged with private companies.[1] Within this programme we have to take care of the kind of teaching mentioned above and we have to consider that the participants of this programme (I) have enthusiasm and (II) are interested in the topic. But (III) do not want to know very much about sampling-theory and (IV) have to be convinced that using sampling techniques is better than their current way. On the following pages, some of the main topics of this programme are pointed out.

The necessity for decisions in business is well known; in the same way it is intended to have the right decisions. The problem of inferential statistics however is that we will never know reality before we make our decisions. Instead, there is a need for a variety of information; even in order to have better ideas for anticipation of what will happen.

The main problem, however, is not the variety but differing qualities of information; this can be structured into the following individual problems:

1. Problem of exact definition of what our goal shall be;

2. Problem of deficiencies in "adequation", i.e. a divergency between theoretical and empirical definitions;

3. Problem of time structure within information, i.e. a question of the kind of information – e.g. if it describes the reaction in the past or if it anticipates the behaviour in the future;

4. Problem of information date, which means that there is a mix of older and newer information every time.

Quality of information has to take care of every single topic; let us demonstrate this by problem four, examined from the field of auditing and accounting. If we consider the doubtfulness of accounts we have to think about risks concerning individual risks of the person and pretty often ore general risks (e.g. labour market, international behaviour etc.). But if this considering takes a long time then we often have the problem of time-gaps within the same kind of information because in between there could have been a change in the economic frame.

Our problem is described precisely by Arking (pp. 3-4):

> *The auditor is frequently confronted by tremendous masses of documentation or, in many instances, smaller numbers of entries but still too many for him to perform a 100 per cent detailed examination in the available time or at reasonable cost. He must then examine a proportion of the documentation in such a situation and base his decision as to whether there is evidence of effective operation of the system on conclusions drawn from this portion of the data.*

There are two ways how this could be done:

a) the auditor judges the whole by examining those records he had chosen through consideration what is most important for this subject – we call such a study a selective test;

b) the records which should be examined are determined by random numbers (respectively generated by systematic sampling with a random starting point or found in a way which is described e.g. by "proportional to size") – we call this random sampling.

Both methods are widely used and every method leads to a decision. These decisions have a variety of dimensions.[2] Our main interest could be an information

- whether the whole is reasonably accurate or not;

- if the internal control system does provide the desired protection;

- from which size the error-rate could be;

- about the different between book amount and audit amount.

The information we receive are depending on whether we are interested in special attributes – like true or false – or in the observed value; the latter we announce with variable sampling because the measured characteristic is expressed by variables (e.g. the amount of a record in US $). But when we

remember our problem of examining only a proportion of the whole, we'll never have the probability of 100 per cent for any decision.

The difference between the two ways mentioned is that the selective test will not give us any information for the whole – only for the audited records. In opposite are those methods based on any kind of random sampling which yield us to an information about the whole based on known probabilities.[3]

So, if we can agree to the statement, that we on one side need decisions in order to know for example about the difference at all between book and audit amount but on the other side need quick decisions (and furthermore have to take care of the costs) then we need statistical auditing based on sampling methods.

Therefore let us consider (A) about the advantages of statistical sampling in auditing, (B) about requests to sampling techniques in order to gain the advantages without leaving auditing standards and (C) about a proposal for one general application of statistical sampling.

To A: The idea of sampling is to draw a small proportion – called sample – out of the population which should be examined.[4] Only the elements in the sample have to be checked and their results to be fixed. With statistical software it is possible to state, with any (expect 100 percent) desired probability, the mean or total and a surrounding confidence interval for the unknown results of the whole population. The width of this interval depends on the observed population, the sample size, the sampling technique and the desired probability.[5] And we know out of sampling theory that – concerning our chosen probability – the result of the sample is no further away from the result we would obtain from a complete examination of all items than this interval. This provides a lot of advantages which should be taken into consideration:

1. Statistical sampling saves time

2. Statistical sampling saves money

3. The result of a statistical sample is objective and defensible

4. Multiple use of sample

5. "The statistical sampling approach may provide a more accurate method of drawing conclusions about a large mass of data than the examination of all the data" (Arking, p. 11).

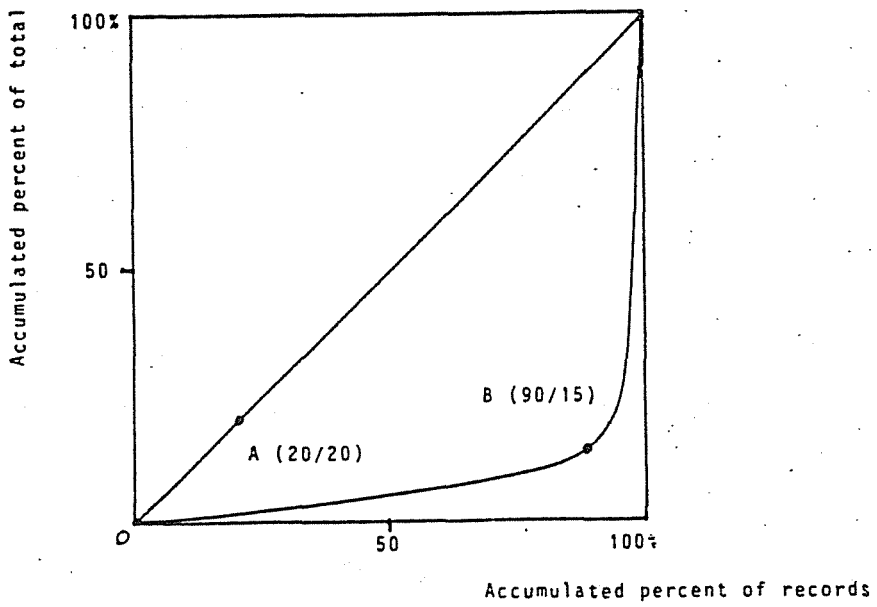
Though criterias (1) to (4) are pretty reasonable, No. 5 is something strange because we have heard of confidence intervals and of probabilities smaller than 100 percent – isn't it? Let us therefore consider in which way a complete inventory is done: either it is on a weekend or there is no selling etc. because we do not need any fluctuation. The examination therefore has to be done in time as short as possible and because the mass of goods for example we need a great lot of auditors. But normally we don't have enough auditors and we therefore take helping hands for counting

out all storage places within the firm. These people, however, will not know about specialities like size or normal volume etc. That is the reason why there will be errors in counting in spite of the examination of all data. And that is the reason for this advantage: because we do not need other people except those auditors who are well known with their job we will not have errors like those mentioned above. Therefore the estimation out of a statistical sampling may be more accurate.

To B: Let us now have a look at the requests to sampling techniques before we will evaluate a proposal for statistical sampling in cases of having large masses of data.[6] Because these requests sometimes are pretty difficult to be answered by those who would like to apply statistical sampling, there are institutions defining the frame for applications.[7]

To C: In the last step we will show a general but very effective use of statistical sampling in various fields of accounting. Before installing any kind of sampling methods we have to analyse our data. Very often we are confronted with a phenomena which has its origin in the fact that there is a great lot of records with (very) small amounts and only a few with (very) high amounts. If we connect information by accumulating the amount as well as the count then we are able to construct a concentration-curve.

Figure 1: Concentration-Curve



The information we receive depends on the extent of concentration. If there would be no concentration at all then we would receive that e.g. those 20 percent of records with the lowest amount would have a share of exactly 20 percent from the total (see Point A in Figure 1). If there is some concentration then the share from the total will decrease. For many

problems in our field we will find a concentration-curve such as the one shown with Point B (see Figure 1). B tells us that those 90 percent of the records with lower amount only have a share of around 15 percent of the total.

Our proposal takes care of this concentration-effect: if we, for example, will find a population of 5000 records with the obtained concentration-curve then we know that 4500 records only have a share of 15 percent. We use this information and draw a sample of perhaps 200 records out of this subpopulation and parallel a complete examination in the remaining part. Altogether we have to examine 700 records out of 5000, which is 14 percent.

In this case we use all the advantages of sampling to 90 percent of our population and connect the appraised result (total value together with the confidence interval) with the results for the complete examined remaining 500 records which have a share around 85 percent of the total. The effect for our decision is obvious: (I) We have an information about the total from what we know that it is complete examined in a very substantial part and an estimation for the rest based on permissible methods, (II) we have the information much quicker as proceeding a complete examination and (III), just in comparison to a selective test, we have the information for the whole population.[8]

Knowledge of the outlined possibilities to receive high quality information at reasonable time and costs should let forget the saying that any second-best but quick decision is better than the best one if it takes too long time. Using statistical sampling meets somewhere the glory idea of a best and quick decision. And that is the way we try to convince the participants.

Footnotes

1. Project "Statistical Sampling for Decisions in Business and Auditing" by G. Frank, V. Reinhardt and L. Scheweis in cooperation with the Frankfurt Chamber of Commerce/German Institute of Business Economics.

2. See e.g. Arking or Schlaifer.

3. See e.g. Neter/Loebbecke, 1975.

4. See e.g. Arking or Roberts.

5. See e.g. Netr/Loebbecke, 1977.

6. A good overview is given by Roberts in Chapter 7.

7. For the FRG it is the IdW and for the USA is it the AICPA.

8. A complete matrix of all alternatives concerning confidence intervals and their consequences for firm's policy with respect to the balance sheet is given by Frank/Schneweis.

Literature

Arking, H. (1983). Handbook of Sampling for Auditing and Accounting. New York.

Frank, G., & Schnewwis, L. (1984). Stichproben im bereich von revision und wirtschaftsprünfung. Der Betrieb, 37, 51-52.

Neter, J., & Loebbecke, J.K. (1975). Behavior of major statistical estimators in sampling accounting populations. New York.

Neter, J., & Loebbecke, J.K. (1977). On the behaviour of statistical estimators when sampling accounting populations. Journal of the American Statistical Association, 72, nr. 359 App. Section, 501-507.

Roberts, D. (1981). Statistical auditing. New York: AICPA.

Schlaifer, R. (1959). Probability and statistics for business decisions. New York.