Sondajul statistic în turism

Tomescu Dumitrescu, C.

2004

The Statistical Survey in Tourism

The statistical survey has a broad area of application in the sphere of tourism. It represents the statistical method of establishment, based on the sampler, on the opinions of different socio-professional categories, with the purpose of anticipating the behavior of the collectivity's members. The essential idea of the survey is that, at the basis of every person's opinions and behavior, stands a general system of references and meanings. Therefore, the individual, subjective opinions present an objective significance that can be generalized and measured.

The survey can be applied in the study of the touristical demand, in the researches that are made in order to get to know the tourists' requests and motivations.

The survey leans on a fundamental theory, because each of the operations, on which its results are based, has a conceptual foundation. Thus, the determination of the number of persons interrogated is based upon the calculus of the probabilities and the law of the big numbers, the establishment of the interrogated persons that support the sampler theory, the significance of the responses has as a support the attitudes' theory, the validity and the fidelity of the responses is founded on the theory of talk, the rapidity of the execution is assured by using computers.

This procedure was also called "selective research", supposing the establishment of a sample from a general collectivity, where to study the respective phenomenon's characteristic, that is, the different aspects of the demand of touristical services. Then, the results obtained through research are extended upon the whole collectivity.

The elements supposed to observation are the tourists. In statistics these ones are called "sample" or "selection collectivity".

Depending on the nature of the observation, the unity of observation can be the individual or the family. The study of the demands of the tourists, constituted in a sample is unfolded through the agency of different methods of inquiry (the questionnaire, the interview, the inquiry, the survey etc.).

The organization of an inquiry referring to the population's demand represents a very complex process. In the specialty literature there are numerous opinions regarding the stages of an investigation of this kind. Roger Mucchielli judiciously establishes the following stages of an inquiry: the determination of the objective, the pre-inquiry, the establishment of the objectives, the determination of the inquiry's universe, the sampler, the choice of the research and carrying out techniques, the processing of the obtained information, the analysis of the results and the redaction of the research rapport.

Every investigation will begin with the delimitation of its object, thing that will imply an abstraction from the multitudes of facts, phenomena and processes. Because, "if everything is essential, then nothing is essential".

The second stage, the pre-inquiry, consists in a logical analysis of the possible hypotheses, selecting the verifiable hypotheses in order to fix the objectives. It is about a critical reflection upon the research's hypotheses.

It has been remarked, without doubt, that "a good inquiry can be selected based on good hypotheses". The fact that we can elaborate an area of problems of a good level without the serious study of the corresponding literature, without the collegial research of the problems and without having a considerable experience in the domain is entirely implausible.

In this phase, the cost of the investigation, the calendar term are estimated and the difficulties from the terrain related to the development of the investigation are forecasted. The study of the problem's bibliography, that is the theoretical works which appeared and the applicative researches effectuated as well as the instruments used, is very important in this stage.

The determination of the objectives of the inquiry and the explicit formulation of the research's hypotheses represent the third stage. It is well known that the hypothesis constitutes a plausible explication that is about to be verified by the factual material, being able to be partially or totally confirmed or infirmed. Or, according to other authors' idea, the hypothesis is the utterance of a causal relation in a form that permits the empirical verification.

There's no doubt that the formulation of some judicious hypotheses permits their testing through an empirical research and constitutes a precious impulse in the research activity.

In the fourth stage, the universe of the inquiry or the population that will be investigated is determined. In this way, for instance, in order to develop the tourism in Malta, the Maltese Tourism Ministry effectuated a survey inquiry in 2004. The inquiry included 6000 tourists, that visited Malta by plane or ship, and that had higher incomes. During the survey they collected information that will respond to the following problems: the dates of the journey, the duration of the sojourning, the localities visited, the preferred type of housing, the main characteristics of the tourists: age, sex, profession, income, residence; the purpose of the journey (tourism, business, family); the travel type: in group or individually; the use of the capacities of the means of transport and the opinions about Malta. Then, in another stage, they proceed to the constitution of the sample, that is, to the determination of the sample's dimension and structure. The sample's representativity, reported to the general collectivity, represents the basic condition of the choice. Because only the sample's representativity permits the extrapolation, that is, the extension upon the ensemble of the results observed upon one part.

In this order of ideas, and without doubt, the marketing specialists appreciate that it will be a mistake to believe that it is necessary to investigate too many samples. In reality, the precision doesn't depend on the fraction, or the

rate of survey, that is, the number of samples reported to the total number of population, but on the sample's effective.

The constitution of the sample depends on the homogeneity of the population from which is extracted. Therefore, when the population is homogenous, a relatively reduced segment of touristical market, it is representative enough in order to characterize the entire content of the collectivity. The size of the sample is determined, depending on this measured eclecticism, through the mean squared error of the unity's characteristics (persons or families).

At the same time, it is also necessary to calculate the survey mean error, given the fact that a single sample is studied and not the entire population. The survey mean error is considered a measure of the representativity mean error and represents the mean squared error of all the samples' means from the statistical population's mean.

Actually, if the dispersion is null, all the unities would have the same values, equal to the mean and the extraction of a single unity, would constitute a representative enough sample. In the case in which the dispersion is weak, that is, the values of the studied characteristic are very much grouped around the mean, the ensemble is homogenous and a limited sample will give a sufficient precision. On the contrary, if the dispersion is high, the population being very heterogeneous, then it is imposed to have a more important sample, in order to obtain the necessary precision.

The probabilistic sampler offers the necessary guarantees of precision, because all the individuals are effectively tossed up, and it is possible to calculate the survey mean. In order to do this, it is necessary to dispose of a complete list of the individuals that figure on the surveyed universe, from which the sample can be extracted. At the same time, the characteristics of the total population, which constitutes the survey's basis, must be recognized in order to be able to find them in the sample. The probabilistic survey's technique sets

aside the personal experience of the researchers or of the organizers of the inquiry in the constitution of the sample, because they proceed according to the ballot principle. This principle is defined through the fact that it gives every unity of population a known probability, different from zero, in order to enter the survey. Therefore, the population's characteristics have thus equal chances to figure with their relative value in the sample. For this reason, it was written that "the samples must be rigorously assayed, aimlessly, in the same way in which at a lottery all the cards have equal chances of designating the winner".

The stratified survey, as opposed to the probabilistic one, is less expensive, because necessitates a minimum of observations and permits obtaining estimations with a smaller error. The population that is to be submitted to the survey is divided in strata, from which the sample is deduced.

The necessary condition for a judicious stratification is that the dispersion of the characteristic studied around the mean to be lower inside the group, than when it comes to the total population. Therefore, the strata should be more homogenous than the ensemble. But, for every stratum, the sample must be aimlessly extracted. The samples will have different sizes according to the degree of eclecticism existing in every stratum. The fraction from the stratum that is explored varies depending on the size of the dispersion.

Both in the probabilistic survey and in the stratified one, the survey basis isn't though perfect as long as the lists with the individuals that figure in the surveyed universe can contain errors and omissions.

And, because of this, the researchers are confronted not only with sampling problems (omissions, disproportions etc.), but also with practical problems (finding the persons they are looking for) or with problems regarding obtaining some comprising and exact responses. In order to remediate the absence or the insufficiency of a correct survey basis, several other methods were put in their place, as for example, the area sampling, the random route and the on quota or proportional sampling.

The area sampling represents a variant of the probabilistic survey. It consists in tossing some geographical areas and then systematically exploring the survey unities that figure in these areas. Through empirical methods and with lower expenses, they try to obtain the expected results from the probabilistic methods.

The random route aimlessly fixes starting points on a plan or on a map and beginning with them, it imposes rigid routes that permit the designation of some dwellings where the investigations are about to start (proceed to investigations).

The on quota or proportional sampling consists in the constitution of a miniature-sample structured depending on some significant criteria, exactly in the same way as the explored population.

These socio-economic criteria can be: the income, the age, the profession of the head of the family, the geographic region. The choice of the persons submitted to the observation is left at the inquisitors' initiative, by receiving indications in what concerns the quotas that they have to observe (of profession, of age, of income etc.). The geographic dispersion of the inquisitors, their diversity and the low number of observations lead to the interrogation of other persons, in the situation of a ballot. And in this situation, a sample, whose main characters are conformingly to the ones of the surveyed universe, is obtained. In order to apply the quotas' method, a necessary condition is disposing of precisely enough data upon the studied population. When the sample is small, the survey through quotas' method is preferable to the random quotas.

The essential problem, when it comes to constituting the sample, either if it's about probabilistic surveys or proportional surveys, is to proceed such as the sample's deformations are avoided.

The causes of these deformations are multiple; the universe of the inquiry is incorrectly defined, the choice of some persons that are easier to find than others, the impossibility of getting to the persons designed through ballot.

The value of a survey can be influenced by a series of errors. However, the most frequent are the errors of observations. They are followed by the sampling errors. Therefore, any estimation that results from a survey is effectuated by, at least, a double source of errors, coming from sampling, on the one hand, and observation, on the other hand.

Another important element of the survey is the questionnaire. In the specialty literature, the questionnaire was defined as a logical succession of written questions or graphical images, depending on the research's hypotheses, that, through the inquiry operators' administration, determine, from the inquired one, a behavior (verbal or non-verbal) that will be registered.

At the basis of the questionnaire, three categories can be retained: the content, the form, the application mode.

Reported to the first criterion, we can have factual questionnaires (they express collective data) and questionnaires of opinion (they contain subjective data), impossible to be directly observed. The second criterion permits the classification into: closed questions questionnaires, that only permit the choice of the responses previously fixed in questionnaires; open questions questionnaires, which leave the subject with the liberty of personally expressing the responses; mixed questions questionnaires. At last, the third criterion leads to the classification into: auto-administrated questionnaires (the registration of the responses by the subjects submitted to the investigation themselves) and the ones administrated by the inquiry operators.

The structure is extremely important in a questionnaire. Through the questionnaire's structure we understand the part formed by the different types of questions and the responses between them.

Reported to their function, in the structure of the questionnaires, the following types of questions can be emphasized:

a) the introductive, of contact or of "breaking the ice" questions have the role of "warming up" the atmosphere;

b) the transition questions indicate the apparition of a new group of questions referring to another problem;
c) the filter questions stop the access of some categories of subjects to some successive questions from the questionnaire, representing at the same time a control of the responses' quality.

Concurrently, the length of the questionnaire represents an essential problem, of research technique, because it expresses the capacity of choosing from the possible indicators' universe the essential ones.

They didn't make a mistake when they said that "a questionnaire interesting for the subject is shorter than the shortest uninteresting questionnaire". The practice proved that 25-50 questions don't cheese off neither the inquisitor, nor the inquired.

The formulation of the questions represents, as well, a problem of technique of the questionnaire. These ones are established depending on the object of the investigation and on the set of indicators that clarifies the theme of the inquiry. In the formulation of the questions, a series of conditions must be kept into account, and we mention some of them here: the content of the questions must be clear, simple, without equivocation, respecting the topic of the phrase or of the proposition; the negations and the very abstract terms will be avoided; the attractive formulation of the questions will be one of the major preoccupations; the neologisms and the archaisms, as well as the technical terms will be avoided, choosing words of broader circulation, and so on.

The operation of gathering the responses can take place in the resorts, at their residence, at expositions, on the street, and, depending on the way in which the information is gathered, we distinguish oral inquiry (the direct or on the phone interviewing) and the written inquiry (by completing a questionnaire). Without doubt, this operation is effectuated depending on the nature of the studied problems and on the expenses allocated to the carrying out of an inquiry.

The carrying out of the research instruments represents another stage of the investigation which is, in fact, a critical reflection upon the work tools (questionnaire, interview, observation guide), that is realized during the pilot inquiry.

Thus, for instance, the carrying out of a questionnaire can submit to reflection the terminology used, its form, and so on.

Before redacting the research report, a last stage would be constituted by the analysis of the results.

As we well know, the scientific theories and the laws represent, in fact, hypotheses confirmed by a series of experiences or explications upon some regularities observed in the nature. However, new observed facts can infirm certain hypotheses or theories, they start new researches in order to discover better hypotheses.

The most difficult problem during the verification of the hypotheses is "when it can be considered that we dispose of sufficient data in order to accept or reject a hypothesis?". We can't objectively respond to this question, but on the basis of gathering a fond of data and on their evaluation. Therefore, in the situations specific to the economical activities also, the researchers and the specialists must be informed both in what concerns "the evidence rules" on whose basis the hypothesis is formulated, and in what concerns its acceptation or rejection. And these rules are offered by the statistical instruments. Moreover, the non-recognition or the application of an inadequate method, when it comes to appreciating the results and testing the hypotheses, can have negative results upon the development of the entire research.

In the first place, in a research, the statistical interpretation of the interdependence relations and the association of the variables of touristical marketing are necessary.

Thus, for instance, let's examine the statistical significance of a survey regarding the opinions of some foreign tourists (100 persons) upon the

presentation of a folkloric program in the restaurants from the Romanian littoral, for three hours every day of the week. It is obvious that the partial observation is the source of a flux of information, but only the interpretation of the statistical significance can give us indications of either this one can replace the total observation.

Among the foreign tourists interrogated during the survey, 76% were in favor of the introduction of the special folkloric program, 20% were against, and 4% were undecided. The question "to what extent can the results of this survey, that comprised 100 tourists, be considered representative for the zone 'Littoral'" is inherently born? It is sure that, if the survey were to be effectuated upon other different 100 persons, the proportions of the responses obtained would be different. The differences of proportions, between the responses obtained in different samples of the same size, come from what we call the variation of survey or sampling. Because none of the samples is a perfect miniature of the entire population, the information is inevitably incomplete. And this interpretation can be realized by calculating the probabilities. The survey error is estimated as a probability. At the same time, we say that an event, that has equal chances of being produced or of not being produced, has a probability $p = 0,5$.

In this example, we must ask ourselves, which is the probability of being wrong, if we state that from the total number of the foreign tourists from the littoral, the ones that prefer the introduction of the special folkloric programs in the restaurants from the Littoral are more numerous than the ones that don't adhere to this idea. The problem that comes upon is the following: the difference between the two categories of tourists comes from sampling variations or from the fact that, in reality, the number of tourists from the first category is larger than the number of tourists from the second category. In different surveys, of course, certain differences will appear, however, bigger these differences are, bigger the probability that they don't come from sampling variations, but from other causes (errors of observation). The significance of a difference can be

evaluated by calculating the distance between the mean of the sample and the mean of the population in the terms of the survey mean errors. In other words, in order to measure the error interval that can be attributed to the respective estimation, the concept of mean squared error was created. According to this one, it is admitted, that the dispersal of the sample's unities, as compared to this mean of the sample, approximates the existent dispersal of the population and that the survey mean squared error of the mean from the sample approximates the error that is made by estimating the mean of the population on the basis of the mean from the sample. We mention the fact that the survey error is different from the mean squared error, because this one represents a measure of the effective distribution, while the survey error measures the dispersal around the mean of a hypothetical distribution.