GENERAL FISHERIES COMMISSION FOR THE MEDITERRANEAN

ISSN 1020-7236

STUDIES AND REVIEWS

No. 73 2004

MANUAL OF FISHERIES SAMPLING SURVEYS: METHODOLOGIES FOR ESTIMATIONS OF SOCIO-ECONOMIC INDICATORS IN THE MEDITERRANEAN SEA

STUDIES AND REVIEWS No. 73

GENERAL FISHERIES COMMISSION FOR THE MEDITERRANEAN

MANUAL OF FISHERIES SAMPLING SURVEYS: METHODOLOGIES FOR ESTIMATIONS OF SOCIO-ECONOMIC INDICATORS IN THE MEDITERRANEAN SEA

by

Evelina Sabatella
Instituto Richerche Economiche per la Pesca e l'Acquacoltura
Salerno
Italy

Ramón Franquesa
Gabinete de Economía del Mar
Universitat de Barcelona
Spain

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS
Rome, 2004

ISBN 92-5-105093-7

PREPARATION OF THIS DOCUMENT

This manual was born in part from work commissioned by the General Fisheries Commission for the Mediterranean (GFCM) Scientific Advisory Committee (SAC) Sub-Committee on Social Science on the "Feasibility Assessment for a Database on Socio-Economic Indicators for Mediterranean Fisheries" and in part from the work carried out by Istituto Ricerche Economiche per la Pesca e l'Acquacoltura (IREPA) on the "Manual sampling for fisheries" used by their statistical recorders under the national multi-annual programme for the Italian Monitoring Programme for fisheries.

The Working Group on Socio-Economic Indicators (Salerno, 11–13 march 2002) reviewed draft guidelines on sampling methodologies for building socio-economic indicators and this manual is the result of this process and the efforts made by IREPA and Gabinete de Economía del Mar (GEM) supported and endorsed by Cooperation Networks to facilitate Coordination to Support Fisheries Management in the Western and Central Mediterranean (COPEMED), Scientific Cooperation to Support Responsible Fisheries in the Adriatic Sea (ADRIAMED) and Assessment and Monitoring of the Fishery Resources and Ecosystems in the Strait of Sicily (MEDSUDMED).

The manual was prepared by Evelina Sabatella and Ramon Franquesa and recommended for publication by the SAC (Rome, 1–4 July 2002).

ACKNOWLEDGEMENTS

This manual on the sampling methodology for use in the construction of socio-economic indicators was initiated by the Sub-Committee on Social Science, Scientific Advisory Committee of the GFCM. It was carried out in collaboration with IREPA (Italy) and GEM (Spain).

The authors would like to thank in particular Vincenzo Placenti from IREPA, the Development Planning Service (FIPP), as well as the Members of the Sub-Committee on Economic and Social Sciences (SCESS) network for their comments and contribution to the manual.

Authors contact:

Evelina Sabatella
Irepa Onlus, Istituto Ricerche Economiche per la Pesca e l'Acquacoltura
Via S. Leonardo, Trav. Migliaro
84131 Salerno, ITALY
Email: esabatella@irepa.org

Ramón Franquesa
GEM, Gabinete de Economía del Mar
Universitat de Barcelona
Email: ramon@gemub.com

Sabatella, E.; Franquesa, R.
Manual of fisheries sampling surveys: methodologies for estimations of socio-economic indicators in the Mediterranean Sea.
Studies and Reviews. General Fisheries Commission for the Mediterranean. No. 73. Rome, FAO. 2003. 37p.

ABSTRACT

This manual on sampling methodologies for the development of socio-economic indicators was initiated by the Sub-Committee on Economic and Social Sciences (SCESS) of the General Fisheries Commission for the Mediterranean Scientific Advisory Committee. It was prepared jointly by the Istituto Ricerche Economiche per la Pesca e l'Acquacoltura (IREPA) and the Gabinete de Economía del Mar (GEM) within the framework of the ad hoc Working Group on Socio-Economic Indicators of SCESS.

The manual is aimed at all decision-makers who may need to collect data to build socio-economic indicators. At an international level, it can be used to facilitate and simplify reporting under international conventions and agreements on matters relating to the sustainable development of the world's fisheries. Regional fisheries bodies and stakeholders involved in fisheries decision-making, such as the fishing industry, other user groups, certification bodies, local communities and non-governmental organizations, may also draw upon this manual to assist in meeting societal goals for fisheries.

The methodologies can be applied to fisheries at many different levels, from individual fisheries and coastal management units to a global level. The manual aims to encourage consistent use of statistical methods in data collection. Governments may also wish to adapt the manual to the specific requirements of their national fisheries.

CONTENTS

Page

1. INTRODUCTION 1

1.1 Socio-economic indicators 1
1.2 Study of the existing situation 1
1.3 Geographical and technical fleet segmentation 3
1.4 The scope of this manual 3

2. POPULATION AND SAMPLING **4**

2.1 Potential methods for data collection 4
2.2 Advantages of sampling methods 5
2.3 Some statistical terms 5

2.3.1 Mean 5
2.3.2 Variance and standard deviation 6
2.3.3 Normal distribution and confidence limits 6

2.4 The role of sampling theory 7
2.5 Probability sampling 8
2.6 Alternatives to probability sampling 8
2.7 Bias and its effects 8

3. THE PRINCIPAL STEPS IN A SAMPLE SURVEY 9

3.1 Objectives of the survey 9
3.2 Population to be sampled 9
3.3 Data to be collected 10
3.4 Degree of precision desired 10
3.5 The questionnaire and the choice of the data collectors 10
3.6 Selection of the sample design 11
3.7 Sampling units 11
3.8 The pre-test 11
3.9 Organization of the field work 11
3.10 Summary and analysis of the data 11
3.11 Information gained for future surveys 12

4. SURVEY METHODS AND QUESTIONNAIRE DESIGN **13**

4.1 Simple random sampling 13
4.2 Stratified random sampling 13
4.3 The estimation of sample size and allocation across strata 14

4.3.1 Sample size with more than one item 15

4.4 Estimation of parameters 16
4.5 Sources of error in surveys 17

4.5.1 Sampling errors 17
4.5.2 Effects of non-response 18

5. REFERENCES CONSULTED 19

APPENDIXES

A – List of socio-economic indicators 21

B – Fleet segmentation and parameters of the socio-economic structure for the operative units 26

C – Questionnaires for the collection of socio-economic data 27

D – Implementation of the Bethel method (SAS ® HML) 34

E – Optimal sample size and the differences between Neyman and Bethel methods – an example 36

1. INTRODUCTION

1.1 Socio-economic indicators

Indicators are data or combination of data collected and processed for a clearly defined analytical or policy purpose. That purpose should be explicitly specified and taken into account when interpreting the value of an indicator. Fisheries indicators should provide practical and cost-effective means for the evaluation of the state and the development of fisheries systems and the effects that policy changes have on those systems.

The importance attached to socio-economic information on the fisheries sector has greatly increased for several years. Most bodies ask themselves what information they should be collecting in order to establish indicators which are representative of the sector.

Most effort to date has been on developing indicators related to the ecological sustainability of fishery systems. There is a large and established literature on the use of a wide range of indicators to assess the relative abundance and health of individual fish stocks. This is done through such concepts as target and limit reference points, biomass indexes, fishing mortality and effort measures, and so on. In the meantime, relatively little attention has been paid to the set of potential indicators that could be used to assess the economic and social aspects of fisheries and the interaction with the pursuit of sustainable development objectives. The growing demand for social and economic indicators from policy makers is a result of this perceived imbalance.

Indicators are not an end in themselves. They are a tool to help make clear assessments of and comparisons between fisheries, through time. They describe in simple terms the extent to which the objectives set for sustainable development are being achieved. The main purpose in developing a set of sustainability indicators is to assist in assessing the performance of fisheries policy and management and to stimulate action to better pursue sustainability objectives.

1.2 Study of the existing situation

At its eighty-third session on 7-9 April 1999, the **OECD** Committee for fisheries decided, as a part of its Programme of Work 2000-2002, to conduct a study on Fisheries Sustainability Indicators. The study "will seek to develop fisheries social and economic indicators to be used as tools in policy analysis" [AGR/FI/M(99)1].

At its eighty-fifth session on 20-22 March 2000, the Committee agreed that: "the overall goal for this Study should be to contribute to improvement in the measurement of economic and social dimensions of sustainable development of fisheries, and where possible, relate these to the resource and environmental dimensions".

The Spanish and Italian representatives (among others) have given progress reports on their work on this subject. The document presented by Spain (AGR/FI(2001)12/PART1) is a sub-set of another document presented in the context of the GFCM, mentioned below. This paper includes a general discussion of the use of indicators, the relationship between economic and environmental indicators and the requirements to be fulfilled for indicators to be useful management instruments. Several methodological and data difficulties have arisen during this project.

Italy's document, entitled "Italian Fisheries: Implementation of a Monitoring System for Techno-Economic Data and the Evaluation of Socio-economic Parameters. Part I – Methodology" (AGR/FI(2001)12/PART3) describes the new survey methodology to be used in fisheries data collection in Italy. A second paper is envisaged to present and discuss empirical findings.

The **European Commission (Fisheries DG and Eurostat)** has been pursuing work and studies to improve knowledge of the sector for several years.

The Scientific, Technical and Economic Committee for Fisheries (STECF) issues useful proposals for managing the Common Fisheries Policy. Annex IV to Council Regulation (EC) No 1543/2000 of 29 June 2000 establishing a Community framework for the collection and management of the data needed to conduct the common fisheries policy, the detailed rules for the application of which are laid down in Commission Regulation N° 1639/2001 of 25 July 2001, also sets out a list of (socio-)economic indicators drawn up by this Committee.

The document entitled "STECF needs for socio-economic indicators" presents a general set of economic and social fisheries stability indicators. The underlying notion is that to be economically and socially sustainable a fishery must be capable of being exploited profitably at some biologically sustainable level. The purpose of indicators must therefore be to show whether a fishery is currently sustainable, economically, socially, and biologically, and if not whether it is capable of being exploited sustainably at all and if so at what levels of capital, and labour employed and of fish stock.

To support the implementation of the Code of Conduct for Responsible Fisheries, the **FAO** has published a document addressed to all parties concerned by the sustainable development of fisheries. Entitled "Indicators for sustainable development of marine capture fisheries" this document, is part of the "FAO Technical Guidelines for Responsible Fisheries No 8, Rome, FAO, 1999, 68p". These guidelines provide general information on the issue of sustainable development focusing on a system of indicators, called the Sustainable Development Reference System (SDRS).

The definition of sustainable development adopted by FAO can be considered a very general framework for fisheries sustainable development. This definition establishes five main components: the multiple resource in its environment; social and economic human needs; the technology; and the institutions. While the first two must be conserved, the others need to be respectively satisfied, controlled and established through the general management process. The guidelines also provide information on the type of indicators and related reference points needed. However, it is recognized that it is difficult to generalize, and that there is a need to agree on common conventions for the purpose of joint reporting at national, regional and global level, particularly in relation to international fisheries, or transboundary resources.

Taking into consideration the conclusions reached at the meeting of the former Working Party on Fishery Statistics and Economics of the General Fisheries Commission for the Mediterranean (**GFCM**), the newly established Scientific Advisory Committee (SAC) of the GFCM invited its Sub-Committee on Social Science to initiate work to set up a database on socio-economic indicators. It recommended to start with a pilot assessment study for a management unit. The Alboran Sea was chosen for the pilot study as it encompasses stocks shared between a developing (Morocco) and a developed (Spain) country, as well as a wide range of fishing operations typical of what can be find elsewhere in the Mediterranean. This would allow elaboration of a socio-economic indicator methodology that would be applicable across each of the GFCM management units and the Mediterranean as a whole.

In 2001, the GFCM published a "Feasibility assessment for a database on socio-economic indicators for Mediterranean fisheries" (Studies and Reviews No. 71). This represents the outcome of the above pilot study on compiling socio-economic indicators initiated by the GFCM's Subcommittee on Social Sciences (Scientific Advisory Committee).

Based on the same methodology this pilot study is being followed by a study on the Gulf of Gabès (Tunisia) fisheries. It is also likely that similar work is going to be organised for the Adriatic Sea fisheries (through the ADRIAMED project).

The Working Group on socio-economic indicators (WGSEI, Salerno 11-13 March 2002) reviewed four studies which have been launched during the intersession. They cover respectively: the Gulf of Gabès (Tunisia); the Adriatic Sea; the Tyrrhenian Sea; and the Gulf of Lion (France). In this respect, the WGSEI reiterated the importance of defining Local Operational Units (LOUs: fleet segments/ports) and recommended the formal adoption of this concept. The WGSEI adopted 3 basic social indicators (average age of fishers; length of activity; share in capital structure) and recommended to pursue works toward identifying other basic social indicators at the level of each GFCM geographical area.

In appendix A lists of socio-economic indicators developed by the previous agencies are presented

1.3 Geographical and technical fleet segmentation

One important consideration in the development of indicators is selection of the geographical "units" for which indicators will be reported. These units should reflect the geographic scale of ecological processes that reasonably define ecosystem boundaries (recognising that boundaries are always open for aquatic ecosystems), fishery resources and fishing activity, and political jurisdictions. While commitments have been made for national reporting, units at a regional scale (either within a nation or for shared resources of several nations) will be more appropriate in some cases. It may be useful to have indicators at finer scales (e.g. individual fisheries or sub-national regions).

Moreover, the economic indicators should complement the tools used in biological assessment of resources, to clarify the consequences for society of resource degradation. The decision-maker’s regulations (on fishing schedules, licenses, taxes, etc.) are usually aimed at specific fleet groups. Therefore, a proper fleet segmentation is essential in the construction of indicators.

Management authority regulations (on fishing schedules, licences, taxes, etc.) are normally binding upon specific fleet groups. That is why a correct fleet segmentation is essential in the construction of the indicators; otherwise they would prove useless. The vessel categories should be flexible enough to cover the whole of the fishing fleet operating in the Mediterranean Sea. At the same time, however, they should be precise enough to yield operative (meaningful) answers to the management units. For this reasons the concept of “Operating Unit” has been developed.

In the Mediterranean Sea context, an important issue was to reach agreement on the number of segments that should be established. The Working Group on socio-economic indicators (WGSEI, Salerno, 11-13 March 2002) reviewed in detail various possible fleet segmentation frames (i.e., from : IREPA; AER; U.E (Regulation No 1639/2001)/EUROSTAT/STEFC), and from the SCESS pilot study on the Alboran Sea). It concluded that these segmentations could be harmonized taking as reference the E.U segmentation. On this basis, the Group elaborated and agreed upon a segmentation to be submitted to SCESS for adoption. This entails 9 segments crossing 3 vessel/LOA groups (see appendix B). It further recognized that more detailed segmentation would need to be identified with regard to the small-scale fleet (< 12 m.).

1.4 The scope of this manual

Indicators need to be underpinned by data. Data availability and costs are major issues in the selection of indicators and their adoption. Data availability and their quality and quantity vary greatly between fisheries and countries.

Much of the data needed for socio-economic indicators are often already being collected by different agencies or ministries. However, the availability of data is uneven across disciplines and countries. More data is available on biological and environmental aspects than on socio-economic ones. Data availability is also uneven between developed and developing countries and it may be necessary to agree on a common minimum set of information to be collected if the objective is to assess progress towards sustainable development at regional or global levels.

The Working Group on socio-economic indicators (WGSEI, Salerno, 11-13 March 2002) reviewed and commented a draft guideline on sampling methodologies for building socio-economic indicators. The Working group also recommended that the draft guidelines be consolidated and finalised, taking into consideration the guideline used by IREPA. The present manual is the final outcome of this process.

This manual is aimed at all decision-makers who may need correct data to build socio-economic indicators. At an international level, the manual can be used to facilitate and simplify reporting under international conventions and agreements on matters relating to the sustainable development of the world's fisheries. Regional fisheries bodies and stakeholders involved in fisheries decision making, such as the fishing industry, other user groups, certification bodies, local communities and non-governmental organisations (NGOs) may also draw upon this manual to assist in meeting societal goals for fisheries.

This manual can be applied to fisheries at many different levels, from individual fisheries and coastal management units to a global level. It aims to encourage consistent usage of statistical methods in data collection. Governments may also wish to adapt the manual to the specific requirements of their national fisheries

Finally, although much care has been taken to integrate in this manual existing knowledge and experience in other sectors in statistical data collection, the experience available from the fishery sector is limited to few Mediterranean countries. As a consequence, these guidelines are intended to be flexible and capable of evolving as experience is gained and constructive suggestions accumulate. The present document is the first version of the manual and will be revised and completed (additional methodological annexes will be included) as required in the future.

2. POPULATION AND SAMPLING

2.1 Potential methods for data collection

Data collection can be classified into two general forms: census and sample.

A census is not a survey per se, as it involves collecting data from all individuals in the target population. Several European logbook programmes could be considered a census as they (theoretically) require all vessels that meet certain characteristics to provide the required data. The key advantage of a census is that (assuming perfect compliance) the results are known with certainty.

The principle disadvantage of a census is the considerable cost involved in collection and the subsequent compilation of all the data collected. In the case of fisheries, the cost of interviewing every fisher to collect the data would be prohibitive. Logbook programmes require fishers to complete the data themselves and provide the completed forms to the appropriate authority. Provision of such data is mandatory for the target population of vessels, and is enforced through legislation that enables prosecution and penalisation of individuals who do not comply or deliberately provide incorrect or misleading information.

While such an approach has considerable appeal, regular provision of such data would place an increased burden on both fishers and administrators. The benefits of more precision in the resulting values of the key indicators would need to outweigh the additional costs for such an exercise to be worthwhile.

Sample surveys involve the collection of data from a sample of the target population rather than all individuals in the target population. The key advantage of the sample survey is that less data need to be collected and analysed.

A key assumption of the sample survey is that the sample is representative of the target population as a whole. A range of sampling methods can be employed to improve the likelihood that the sample is representative (see next paragraphs), although a risk always remains that the sample estimates are biased due to the sample being different in some way to the target population as a whole. However, as the standard error decreases with sample size, the optimal sample sizes can be determined based on the desired level of precision of the data (see paragraph 4.3).

2.2 Advantages of sampling methods

- Reduced cost

If data are secured from only a small fraction of the aggregate, expenditures are smaller than if a complete census is attempted. With large populations, results accurate enough to be useful can be obtained from samples that represent only a small fraction of the population.

- Greeter speed

For the same reason, the data can be collected and summarized more quickly with a sample than with a complete count. This is a vital consideration when the information is urgently needed.

- Greater scope

In fisheries inquiry trained personnel or specialised equipment, limited in availability must be used to obtain the data. A complete census is impracticable: the choice lies between obtaining the information by sampling or not at all. Thus surveys that relay on sampling have more scope and flexibility regarding the types of information that can be obtained.

- Greater accuracy

Because personnel of higher quality can be employed and given intensive training and because more careful supervision of the field work and processing of results becomes feasible when the volume of work is reduced, a sample may produce more accurate results than the kind of complete enumeration that can be taken.

2.3 Some statistical terms

This document is not a text book of statistics many of which describe the statistical concepts related to sampling. However, knowledge of some basic statistical terms is required to a better understanding of the next sections.

2.3.1 Mean

The arithmetic mean or the *mean* of a set of *N* numbers $X_1, X_2, X_3, \ldots, X_N$ is denoted by $\overline{X}$ (read "X bar") and is defined as:

$$\overline{X} = \frac{X_1 + X_2 + X_3 + \ldots X_N}{N} = \frac{\sum_{j=1}^{N} X_j}{N} \qquad (1)$$

2.3.2 Variance and standard deviation

The standard deviation of a set of N numbers $X_1, X_2, X_3, \ldots, X_N$ is denoted by σ and is defined as:

$$\sigma = \sqrt{\frac{\sum_{j=1}^{N}\left(X_j - \overline{X}\right)^2}{N}} = \sqrt{\frac{\sum x^2}{N}} \qquad (2)$$

where x represents the deviation of each of the numbers X_j from the mean $\overline{X}$.

Sometimes the standard deviation for the data of a sample is defined with (N-1) replacing N in the denominators of the previous expression because the resulting value represents a better estimate of the standard deviation of a population from which the sample is taken. For large value of N (certainly N>30) there is practically no difference between the two definitions. Also, when the better estimate is needed we can always obtain it by multiplying the standard deviation computed according to the first definition by $\sqrt{N/(N-1)}$.

The variance of a set of data is defined as the square of the standard deviation and is thus given by σ^2.

When it is necessary to distinguish the standard deviation of a population from the standard deviation of a sample drawn from this population, we often use the symbol s for the latter and σ for the former. Thus s^2 and σ^2 would represent the *sample variance* and the *population variance* respectively.

Finally, we define the coefficient of variation as:

$$CV = \frac{\sigma}{\overline{Y}} \qquad (3)$$

The coefficient of variation does not depend on the measurement unit and it gives an indication of the importance of the standard deviation with respect to the mean.

2.3.3 Normal distribution and confidence limits

The normal distribution is a bell-shaped distribution which is used most extensively in statistical applications in a wide variety of fields. Its probability density function is given by:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] \quad -\infty < x < +\infty \qquad (4)$$

Its mean is μ and its variance is σ^2. When x has the normal distribution with mean μ and variance σ^2, we write this compactly as $x \approx N(\mu, \sigma)$.

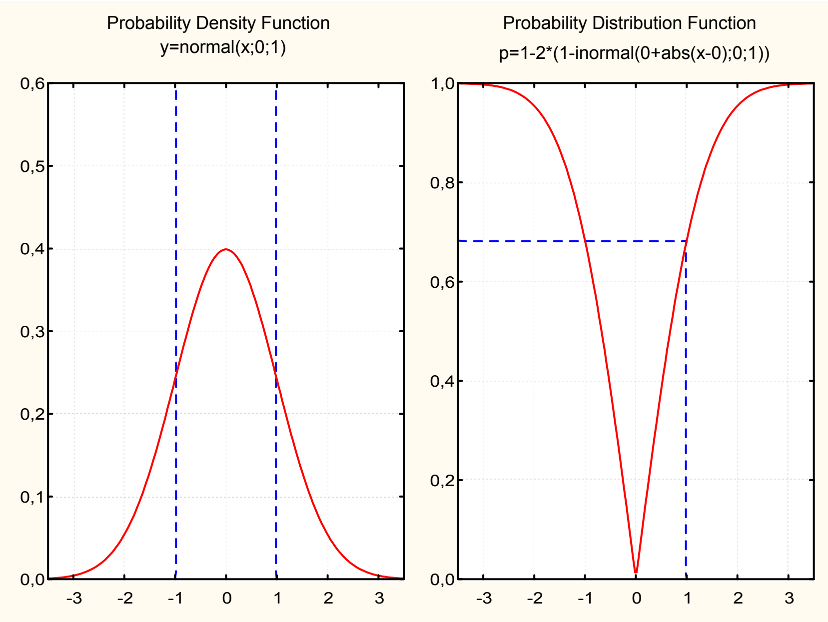
When the variable x is expressed in terms of standard units, $z=(x-\mu)/\sigma$, the previous equation is replaced by the so called standard form:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \qquad (5)$$

In such case we say that z is normally distributed with mean zero and variance one.

A graph of this standardised normal curve is shown in figure 2.1. In this graph we have indicated the area included between z=-1 and +1 as equal to 68.27% of the total area which is one. The areas included between z=-2 and +2 and z=-3 and +3 are equal respectively to 95.45% and 99.73%.

Figure 2.1 - Standardized normal curve



For example the "99% confidence" figure implies that if the same sampling plane were used many times in a population, a confidence statement being made from each sample, about 99% of these statement would be correct and 1% wrong.

2.4 The role of sampling theory

Sampling theory is a study of relationships existing between a population and samples drawn from the population. It is of great value in many connections. For example it is useful in estimation of unknown population quantities (such as population mean, variance, etc.), often called population parameters or briefly parameters, from a knowledge of corresponding sample quantities (such as sample, mean, variance, etc.), often called sample statistics or briefly statistics.

The purpose of sampling theory is to make sampling more efficient. It attempts to develop methods of sample selection and estimation that provide, at the lowest possible cost, estimates that are precise enough for our purpose. In order to apply this principle, we must be able to predict, for any sampling procedure that is under consideration, the precision and the cost to be expected.

Sampling theory is also useful in determining whether observed differences between two samples are actually due to change variation or whether they are really significant. The so-called tests of significance and hypothesis are important in the theory of decisions.

In general, a study of inferences made concerning a population by use of samples drawn from it, together with indications of the accuracy of such inferences using probability theory, is called *statistical inference*.

2.5 Probability sampling

The sampling procedures have the following mathematical properties in common.

1. We are able to define the set of distinct samples S1,S2,…Sv, which the procedure is capable of selecting if applied to a specific population. This means that we can say precisely what sampling units belong to S1, S2, and so on.

2. Each possible sample Si has assigned to it a known probability of selection πi.

3. We select one of the Si by a random process in which each Si receives its appropriate probability πi of being selected.

4. The method for computing the estimate from the sample must be stated and must lead to a unique estimate for any specific sample.

For any sampling procedure that satisfies these properties, we are in a position to calculate the frequency distribution of the estimates it generates if repeatedly applied to the same population. The term probability sampling refers to this situation.

In practice we seldom draw a probability sample by writing down the S_i and π_i outlined above. This is too laborious with a large population, where a sampling procedure may produce billions of possible sample. The drawn is most commonly made by specifying probabilities of inclusion for the individual units and drawing units, one by one in groups until the sample of desired size and type is constructed.

2.6 Alternatives to probability sampling

The following are some common types of non-probability sampling.

1. The sample is restricted to a part of the population that is readily accessible.

2. The sample is selected without conscious planning.

3. With a small but heterogeneous population, the sampler inspects the whole of it and selects a small sample of “typical “ units – that is units that are close to his impression of the average of the population.

4. The sample consists essentially of volunteers, in studies in which the measuring process is unpleasant or troublesome to the person being measured.

In some cases and under the right conditions, any of these methods can give useful results. They are not, however, open to the development of a sampling theory, since no element of random selection is involved. These methods, moreover, are unable to predict from the sample the accuracy to be expected in the estimates.

2.7 Bias and its effects

Sample bias largely arises as a result of inappropriate sample selection (i.e. the average of the selected group differs in characteristics from the true average of target population). As most sample surveys are completed on a voluntary basis, bias may also be introduced through non-response. In such cases, bias may arise if the individuals who do not participate have different characteristics to the target population as a whole. As information is not subsequently derived from these individuals, it is impossible to determine the extent of any bias that may be introduced.

Moreover, in sample survey theory it is necessary to consider biased estimators for two reasons.

1. In some of the most common problems, estimators that are convenient and suitable are found to be biased.

2. Even with estimators that are unbiased in probability sampling, errors of measurement and non response may produce biases in the numbers that we are able to compute from the data.

The use of a stratified random sample approach (see paragraph 4.2) reduces the potential for sample bias, but required additional information on the target population prior to the sample selection. Where the complete sample for particular fleet segments cannot be achieved due to non-response, bias can be reduced through assigning weights to the individual sample responses to re-balance the data. The potential bias arising directly from non-response can be reduced through replacement of boats with similar characteristics, on the assumption that the similar boat is as representative of the boat that failed to respond.

3. THE PRINCIPAL STEPS IN A SAMPLE SURVEY

As a preliminary to a discussion of the role that theory plays in a sample survey, it is useful to describe briefly the steps involved in the planning and execution of a survey.

The principal steps in a survey are grouped somewhat arbitrarily under 11 headings.

3.1 Objectives of the survey

The first step when assessing a sample survey is to well identify the general objectives of the survey. Without a lucid statement of the objectives, it is easy in a complex survey to forget the objectives when engrossed in the details of planning, and to make decisions that are at variance with the objectives.

One of the principal choice is between average values (mean of the population) or total values. In fact, depending on this choice, techniques for the optimal sample size and estimators factors are different.

A number of measures exist that have been used by various agencies to measure the economic significance of fisheries to the regional economy. In addition, a number of performance indicators also exist that can be used to assess the performance of fisheries management in achieving its economic objectives (see chapter 1 and related annexes).

3.2 Population to be sampled

The word population is used to denote the aggregate from which the sample is chosen. The definition of the population may present some problems in the fishing sector, as it should consider the complete list of vessels and their physical and technical characteristics.

The population to be sampled (the sampled population) should coincide with the population about which information is wanted (the target population). Some-times, for reasons of practicability or convenience, the sampled population is more restricted than the target population. If so, it should be remembered that conclusions drawn from the sample apply to the sampled population. Judgement about the extent to which these conclusions will also apply to the target population must depend on other sources of information. Any supplementary information that can be gathered about the nature of the differences between sampled and target population may be helpful.

For example, let us consider the Italian statistical sampling design for the estimation of "quantity and average price of fishery products landed each calendar month in Italy by Community and EFTA vessels" (Reg. CE n. 1382/91 modified by Reg. CE n. 2104/93). Aim of the survey is to estimate total catches and average prices for individual species. Therefore, the sampling basis consists of the more than 800 landing points spread over the 8 000 km of Italian coasts. It is not however feasible to consider the list of the landing points as the list of elementary units. To overcome these difficulties, a sampled population, distinct from the target population but including units in which the considered phenomenon takes place, has been considered. In synthesis, the elementary units considered are the landings of the vessels belonging to the sampled fleet. Thus, the list from which the sampling units are extracted is constituted by all the vessels belonging to the Italian fishery fleet.

3.3 Data to be collected

It is well to verify that all the data are relevant to the purposes of the survey and that no essential data are omitted There is frequently a tendency to ask too many questions, some of which are never subsequently analysed. An overlong questionnaire lowers the quality of the answers to important as well as unimportant questions.

3.4 Degree of precision desired

The results of sample surveys are always subject to some uncertainty because only part of the population has been measured and because of errors of measurement. This uncertainty can be reduced by taking larger samples and by using superior instruments of measurement. But this usually costs time and money. Consequently, the specification of the degree of precision wanted in the results is an important step. This step is the responsibility of the person who is going to use the data. It may present difficulties, since many administrators are unaccustomed to thinking in terms of the amount of error that can be tolerated in estimates, consistent with making good decisions. The statistician can often help at this stage.

3.5 The questionnaire and the choice of the data collectors

There may be a choice of measuring instrument and of method of approach to the population. The survey may employ a self-administered questionnaire, an interviewer who reads a standard set of questions with no discretion, or an interviewing process that allows much latitude in the form and ordering of the questions. The approach may be by mail, by telephone, by personal visit, or by a combination of the three. Much study has been made of interviewing methods and problems.

A major part of the preliminary work is the construction of record forms on which the questions and answers are to be entered. With simple questionnaires, the answers can sometimes be pre-coded, that is, entered in a manner in which they can be routinely transferred to mechanical equipment. In fact, for the construction of good record forms, it is necessary to visualise the structure of the final summary tables that will be used for drawing conclusions.

Information may be collected using a number of different survey methods. These include personal interview, telephone interview or postal survey. The questionnaire design needs to vary based on the approach taken.

Personal interviews involves visiting the individual from which data are to be collected. The interviewer controls the questionnaire, and fills in the required data. The questionnaire can be less detailed in terms of explanatory information as the interviewer can be trained on its completion before starting the interview process. This type of survey is best for long, complex surveys and it allows the interviewer and fisher to agree a time convenient for both parties. It is particularly useful when the respondent may have to go and find information such as accounts, log book records etc. The personal interview approach also allows the interviewer to probe more fully if he/she feels that the fisher has misunderstood a question, or information provided conflicts with other earlier statements.

Data collectors are usually external to the phenomenon that is being examined and, moreover, they are often part of some public structure, in order to avoid possible influences due to personal interests. However, on the basis of the experience acquired in this field by Irepa, it has been demonstrated (Istat, Irepa 2000) that it is essential to have data collectors belonging to the fishery productive chain in order to obtain correct and timely data. Therefore, data collectors should belong to the productive or management fishery sectors.

During meetings on socio-economic indicators partners involved presented several questionnaires. These questionnaires are aimed to collect the information required to calculate the socio-economic indicators and some of them are reported in appendix C.

3.6 Selection of the sample design

There is a variety of plans by which the sample may be selected (simple random sample, stratified random sample, two-stage sampling, etc.). For each plan that is considered, rough estimates of the size of sample can be made from a knowledge of the degree of precision desired. The relative costs and time involved for each plan are also compared before making a decision.

3.7 Sampling units

Sample units have to be drawn according to the sample design.

To draw sample units from the population, several methods can be used, depending on the type of the chosen sample strategy:

- sample with equal probabilities
- sample with probabilities proportional to the size (PPS).

In the first case, each unit of the population has the same probability to take part of the sample, while in the case of a PPS sample each unit has a different probability to be sampled and this probability is proportional to the following measure: $P_i = X_i/X_h$, where, i = a generic vessel, h = stratum, X= a size parameter, for example the overall length of a vessel.

3.8 The pre-test

It has been found useful to try out the questionnaire and the field methods on a small scale. This nearly always results in improvements in the questionnaire and may reveal other troubles that will be serious on a large scale, for example, that the cost will be much greater than expected.

3.9 Organization of the field work

In a survey, many problems of business administration are met. The personnel must receive training in the purpose of the survey and in the methods of measurement to be employed and must be adequately supervised in their work.

A procedure for early checking of the quality of the returns is invaluable.

Plans must be made for handling non-response, that is, the failure of the enumerator to obtain information from certain of the units in the sample.

3.10 Summary and analysis of the data

The first step is to edit the completed questionnaires, in the hope of amending recording errors, or at least of deleting data that are obviously erroneous. The check on the elementary data to eliminate non-sampling errors can be achieved by means of computer programmes implemented to correct the erroneous values and to permit statistical data analysis. These programmes are mainly based on graphical analysis of elementary data.

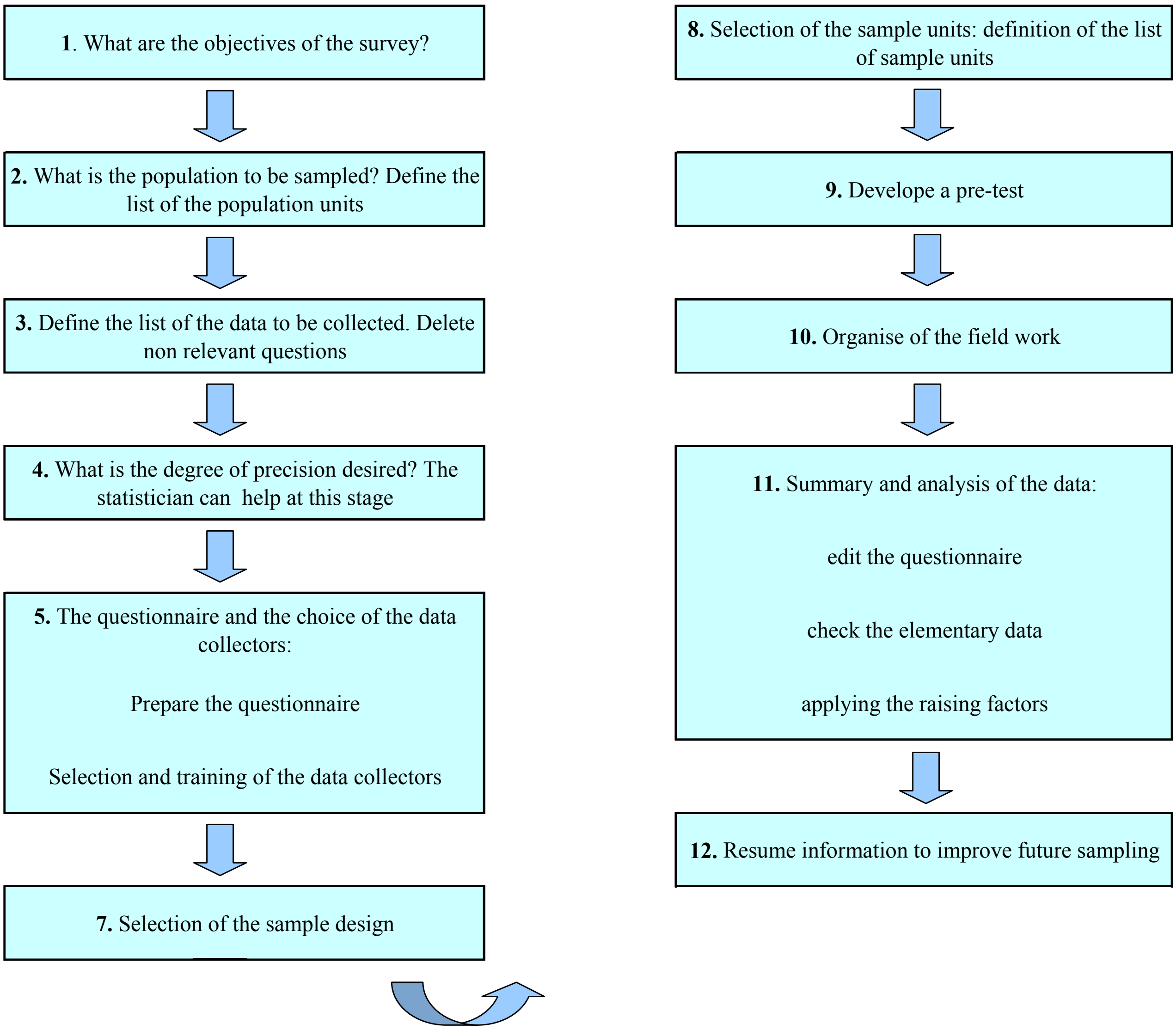
Thereafter, the computations that lead to the estimates are performed. Different methods of estimation may be available for the same data.

In the presentation of results it is good practice to report the amount of error to be expected in the most important estimates One of the advantages of probability sampling is that such statements can be made, although they have to be severely qualified if the amount of non-response is substantial

3.11 Information gained for future surveys

The more information we have initially about a population, the easier it is to devise a sample that will give accurate estimates. Any completed sample is potentially a guide to improved future sampling, in the data that it supplies about the means, standard deviations, and nature of the variability of the principal measurements and about the costs involved in getting the data. Sampling practice advances more rapidly when provisions are made to assemble and record information of this type.

Figure 1: The principal steps in a sample survey



4. SURVEY METHODS AND QUESTIONNAIRE DESIGN

A number of methods exist that can be used to select a sample. In the next paragraphs the most common sample designs are described.

4.1 Simple random sampling

Simple random sampling is a method of selecting n units out of the N such that everyone of the ${}_NC_n$ distinct sample has an equal chance of being drawn. In practice a simple random sample is drawn unit by unit. The units in the population are numbered from 1 to N. A series of random numbers between 1 and N is then drawn, either by means of a computer programme that produces such table. Random samples are particularly useful when little is known about the target population.

When a number that has been drawn is removed from the population for all subsequent draws, the method is also called random sampling *without replacement*. Random sampling with replacement is entirely feasible: at any draw, all N members of the population are given an equal chance of being drawn, no matter how often they have already been drawn.

4.2 Stratified random sampling

In stratified sampling the population of N units is first divided into subpopulations of $N_1, N_2,...,N_H$ units, respectively. These subpopulations are non-overlapping, and together they comprise the whole of the population, so that:

$$N_1 + N_2 + ... + N_H = N \qquad (6)$$

The subpopulations are called strata. To obtain the full benefit from stratification, the values of the N_h must be known. When the strata have been determined, a sample is drawn from each, the drawings being made independently in different strata. The sample sizes within the strata are denoted by n_1, n_2, ..., n_H, respectively. If a simple random sample is taken in each stratum, the whole procedure is described as stratified random sampling.

Given the above definition, we can state that stratified random samples take advantage of additional information of the fishery. For example, if boat length information was available, the survey could be stratified on the basis of this variable. The idea is to group boats into (what are considered to be) relatively homogeneous groups. For example, the boats could be grouped into several strata on the basis of the region they belong to.

Stratification is a common technique. There are many reasons for this; the principal ones are the following.

1. If data of known precision are wanted for certain subdivisions of the population, it is advisable to treat each subdivision as “population” in its own right.

2. Administrative convenience may dictate the use of stratification; for example for socio-economic indicators in the Mediterranean stratification per management units could be advisable.

3. Sampling problems may differ markedly in different parts of the population.

4. Stratification may produce a gain in precision in the estimates of characteristics of the whole population. It may possible to divide a heterogeneous population into subpopulations, each of which is internally homogenous. This is suggested by the name strata, with its implication of a division into layers. If each stratum is homogeneous, in that the measurement vary little from one until another, a precise estimate of any stratum mean can be obtained from a small sample

in that stratum. These estimates can then be combined into a precise estimate for the whole population.

5. The potential for sample bias is reduced through using a stratified random sample. Since the population has been stratified, a balanced sample can be chosen that is more likely to be representative of the population than a purely random sample. Consequently, smaller samples can be selected than using a purely random sample

The theory of stratified sampling deals with the properties of the estimates from a stratified sample and with the best choice of the sample size n_h to obtain maximum precision. The problems of how to construct strata and of how many strata there should be are presented in the next paragraph.

4.3 The estimation of sample size and allocation across strata

In the planning of a sample survey, a stage is always reached at which a decision must be made about the size of the sample. The decision is important. Too large a sample implies a waste of resources, and too small a sample diminishes the utility of the results. The decision cannot always be made satisfactorily; often we do not possess enough information to be sure that our choice of sample size is the best one. Sampling theory provides a framework to solve these problems.

The principal steps involved in the choice of a sample size are as follows.

1. There must be some statement concerning what is expected of the sample. This statement usually is in terms of desired limits of error.

2. Some equation that connects n with the desired precision of the sample must be found. The equation will vary with the content of the statement of precision and with the kind of sampling that is contemplated. One of the advantages of probability sampling is that it enables this equation to be constructed.

3. This equation will contain, as parameters, certain unknown properties of the population (for instance the variability of the investigated phenomenon). These must be estimated in order to give specific results.

4. Finally, the chosen value of n must be appraised to see whether it is consistent with the resources available to take the sample. This demands an estimation of the cost, labour, time and material required to obtain the proposed size of sample.

In the case of simple random sampling, the formula for the definition of the sample size, n, is the following:

$$n \geq \frac{N^2 S^2}{\varepsilon^2 Y^2 + N S^2} \qquad (7)$$

where:

N is the number of units in the population,

ε is equal to $\theta/2$, where θ is the maximum error that is accepted for the final estimates,

S^2 is an estimate of $V(Y)$, total variance of the phenomenon in the population and is given by:

$$S^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2 \qquad (8)$$

with

$$\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i \qquad (9)$$

In the case of single stage stratified sampling, and in the hypothesis of extracting the sampling units with equal probability and without re-pooling, the formula to calculate the sample size, for a maximum error of 2θ, and with a probability equal to P=95%, is the following:

$$n \geq \frac{\left(\sum_{h=1}^{H} N_h S_h\right)^2}{\theta^2 + \sum_{h=1}^{H} N_h S_h^2} \qquad (10)$$

where, H is the total number of strata, N_h is the population size in the stratum h, S^2_h is the estimate of the variance in the stratum h.

Once the sample size has been obtained, the allocation among strata has to be defined. The allocation across strata can be assessed on the basis of two different methods:

- the proportional criterion, in each stratum we have the same number of units:

  $n_h = n/H$

- the Neyman criterion, by which a variable percentage of elements is drawn from each stratum in order to minimise the value of the variance. In this case, the formula is:

$$n_h = n\frac{N_h S_h}{\sum_{h=1}^{H} N_h S_h} \qquad (11)$$

In order to apply the previous formulas No. 7, 10 and 11 a pre-estimate of the $S^2_{(h)}$ variances is required; in other words, the variances of the target variables of the survey must be known. For this purpose, the results of previous sampling survey can be used. In the case this information is not available, a pilot study can be developed to have an indication of the variability of the investigated phenomenon.

4.3.1 Sample size with more than one item

In most surveys information is collected on more than one item. Sometimes the number of items is large. In particular, fisheries surveys are generally multivariate, that is, the variables investigated are more than one (revenues, costs, employment, etc.).

If a desired degree of precision is prescribed for each item, the calculations lead to a series of conflicting values of n, one for each item. It may happen that the *n*'s required are all reasonably close. If the largest of the *n*'s falls within the limits of the budget, this *n* is selected. More commonly, there is a sufficient variation among the n's so that we are reluctant to choose the largest, either from budgetary considerations or because this will give an over-all standard of precision substantially higher than originally contemplated. In this event the desired standard of precision may be relaxed for certain of the items, in order to permit the use of a smaller value of n.

However, nowadays with progress in sampling theory and software technology some methods are available to menage the problem of the sample size with more than one item.

In particular, we give a brief description of the Bethel method that is the application of Neyman's method to the multivariate case. The approach used by this method is to transform the analysis into a linear programming model that allows the identification of the sample size and the allocation across strata, minimising the variances of all variables simultaneously (see also Bethel, 1989).

The optimal allocation across strata for multi-scope studies has been solved by Bethel using the Kuhn-Tucker theorem and then deriving the expressions for the optimal allocation in terms of the LaGrange multipliers. The Bethel method has been implemented on SAS basis (the implementation algorithm is reported in appendix D). Appendix E reports a comparison of the results and relative costs of the two different statistical sampling techniques (Neyman and Bethel).

4.4 Estimation of parameters

In the previous chapters we saw how sampling theory can be employed to obtain information about samples drawn at random from a known population. From a practical viewpoint, however, it is also very important to be able to infer information about a population by use of samples drawn from it. Such problems are dealt with in statistical inference, which uses principles of sampling theory.

One important problem of statistical inference is the estimation of values referred to the population (such as population totals, means, variances, etc.) from the corresponding sample data.

In this manual, the Horvitz-Thompson estimator is reported, but a lot of other estimators exist that have been studied by the sampling theory.

The Horvitz-Thompson estimator is used to estimate the total values (for examples total catches, total revenues, total employment, total costs, and so on).

In the case of simple random sampling, and in the hypothesis of extracting the sampling units with equal probability and without re-pooling the formula is:

$$\hat{Y}_{HT} = \sum_{i}^{n} w_i y_i \qquad (12)$$

where $w_i = N/n$, and y_i are the observed values from the sample units.

In the case of a stratified random sampling, and in the hypothesis of extracting the sampling units with equal probability and without re-pooling, the formula is:

$$\hat{Y}_{HT,h} = \sum_{i=1}^{n_h} w_{hi} y_{hi} \qquad (13)$$

with:

n_h sample size in the stratum h;

$w_{hi} = N_h/n_h$:

y_{hi}, sample data of the unit i in the stratum h.

4.5 Sources of error in surveys

The sample theory assumes that some kind of probability sampling is used and that the observations y_i on the ith unit is the correct value for that unit. The error of estimate arises solely from the random sampling variation that is present when n of the units are measured instead of the complete population of N units.

These assumptions hold reasonably well in the simpler types of surveys in which the measuring devices are accurate and the quality of work is high. In complex survey, particularly when difficult problems of measurement are involved, the assumptions may be far from true. Three additional sources of error that may be present are as follows.

1. Failure to measure some of the units in the chosen sample. This may occur because of their refusal to answer the questions.

2. Errors of measurement on a unit. The measuring device may be biased or imprecise. The respondents may not possess accurate information or they may give biased answers

3. Errors introduced in editing, coding and tabulating the results.

These sources of error necessitate to develop methods for computing standard errors and confidence limits that remain valid when the other errors are present.

4.5.1 Sampling errors

The standard deviation is used to provide information about the relative distribution around the estimates. The level of confidence is related to the amount of variation around the estimates, so it is related to the standard deviation. Also, from the Central limit Theory, the larger the sample size, the greater the confidence in the estimate. Therefore the level of confidence is related also to the size of the sample.

The standard error is an indicator of the level of confidence in the estimate. In the case of stratified random sampling, the standard error is given by

$$\hat{S}E(\hat{Y}_{h,HT}) = \sqrt{\frac{1-f_h}{n_h}} N_h S_h \qquad (14)$$

where s is the standard deviation of the sample and f_h is equal to n_h/N_h.

Standard errors are often expressed as relative standard errors in survey reports. These are the standard error expressed as a percentage of the mean. The relative standard error is given by

$$R\hat{E}(\hat{Y}_{h,HT}) = \sqrt{\frac{1-f_h}{n_h}} \hat{C}_h \qquad (15)$$

where C_h is the estimate of the variation coefficient.

Standard errors provide information about the confidence interval around the estimates. Assuming that the distribution is normally distributed around the estimate value, there is a 95 per cent probability that the true population value is within 2 standard errors of the sample value. That is, there is a 95 per cent probability that

$$\hat{Y}_{h,HT} - 2SE(\hat{Y}_{h,HT}) < Y_h < \hat{Y}_{h,HT} + 2SE(\hat{Y}_{h,HT}) \quad (16)$$

For small samples, the confidence interval is defined by the t distribution, such that

$$\hat{Y}_{h,HT} - t_{0.0025,n-1}SE(\hat{Y}_{h,HT}) < Y_h < \hat{Y}_{h,HT} + t_{0.0025,n-1}SE(\hat{Y}_{h,HT}) \quad (17)$$

where $t_{0.025,n-1}$ is the critical value of the t-statistic at the 5 per cent level of significance and $n-1$ degrees of freedom. This value can be read off any t-statistic table. For the number of observations in the sample, the critical values of the t-statistic ranged from roughly 2.1 to 2.6 depending on the number of observations in each size class. For the sample as a whole, the critical value is approximately 2. As the standard error is dependent upon the number of observations in the sample, the larger the sample, the smaller the standard error and the tighter the confidence interval around the mean. Consequently, the larger the sample, the more confidence that can be placed on the sample estimate.

4.5.2 Effects of non-response

Information can only be collected from those individuals selected willing to participate in the survey. However, a rejection may occur for a number of reasons, such as bad timing of the survey. A problem with rejection is that there is no guarantee that the individual who does not want to participate is the same as those who do participate. In some cases, there may be a correlation between rejection and the characteristics of the individual being surveyed. In such cases, exclusion of these individuals may result in a biased sample.

An advantage of the stratified approach is that another individual in the group can be selected to replace the reject. While this still may lead to some bias in the sample, the bias is not likely to be as great as in a purely random sample. This is because the reject is being replaced by another individual with similar characteristics, and the assumption is made that any individuals within the group are representative of the group. The greater the degree of stratification, the more likely this assumption is to hold.

Another way to treat the non-response is to adjust the initial weights w_{hi} on the basis of data referring to the responses (r_h) and the non-responses (s_h) of the sample (n_h). The method consists in multiplying the initial weights (w_{hi}) by a factor (d_h) equivalent to:

$$d_h = \frac{r_h + s_h}{r_h} \quad (18)$$

The hypothesis under this method is that a homogeneity of response exists within strata (Sarndal, Swensson and Wretman 1992). And in fact, the related estimator is called RHG (response homogeneity group).

5. REFERENCES CONSULTED

Bethel, J. 1989. Sample allocation in Multivariate Surveys. *Survey Methodology*, 15(1): 47-57.

Cassel C.M., Sarndal, C.E. & Wretman, J.H. 1977. Foundation of inference in survey sampling - Wiley

Cicchitelli, G., Herzel, A. & Montanari, G.H. 1992. Il campionamento statistico. Il Mulino, Bologna, 571 pp.

Cochran, G.W. 1977. Sampling techniques. Wiley, 430 pp

Concerted Action N.2. 2000. "Promotion of Common Methods for Economic Assessment of EU Fisheries", Annual Report 2000. Concerted Action N.2. Contract FAIR CT97-3541. ISBN: 90-5242-624-4, Report to the EC DG XIV, Brussels, November 2000.

Commission of the European Communities. 2002. STECF's needs for socio-economic indicators. 14th Report Brussels, 22-26 April 2002

EU Project no. 00/32. 2001. Data on economic performance of the fisheries sector. Final report. October 2001

EUROSTAT. 2002. Socio-economic Indicators for fisheries. Meeting of the working group "fishery statistics" 18-19 February 2002

EUROSTAT. 2002. National programmes of fishery statistics – Italy, proposal for a statistical sampling survey for the estimation of "quantity and average price of fishery products landed each calendar month in Italy by Community and EFTA vessels" (Reg. CE n. 1382/91 modified by Reg. CE n. 2104/93). Doc. ASA/FISH/218en

FAO. 1999. Indicators for sustainable development of marine capture fisheries. Technical Guidelines for Responsible Fisheries No.8

FAO ADRIAMED. 2001. Report of the AdriaMed Meeting on Socio-Economic Aspects of the Adriatic Sea Fishery Sector Campobasso, Italy, 28th – 29th May 2001

Fox, D.R. 1989. Computer selection of Size-Biased Samples, The American Statistician, 43(3): 168-194.

GFCM. 2001. Feasibility assessment for a database on socio-economic indicators for Mediterranean Fisheries. Studies and reviews No. 73

GFCM/SAC/SCESS. 2002. Deuxième Groupe de travail *ad hoc* sur les indicateurs socio-économiques (Salerno, Italie, 11-13 mars 2002)

IREPA, ISTAT. 2000. Proposta di una indagine campionaria per la stima delle qualità e del prezzo medio dei prodotti della pesca sbarcati sul territorio italiano da navi da pesca comunitarie e dell'EFTA, nel corso di ogni mese di calendario". Report Istat-Irepa.

IREPA. 2001. Osservatorio economico sulle strutture produttive della pesca marittima in Italia. Franco Angeli Eds. Milano, July 2001, Italy, 312 p.

IREPA. 2002. Guidelines for sampling methodologies for socio-economic indicators. GFCM - SAC – SCESS, Barcelona, May 2002

ISTAT, IREPA. 2000. Proposta di una indagine campionaria per la stima delle qualità e del prezzo medio dei prodotti della pesca sbarcati sul territorio italiano da navi da pesca comunitarie e dell'EFTA, nel corso di ogni mese di calendario". Report Istat-Irepa.

ISTAT. 1989. Manuale di tecniche di indagine, Note e Relazioni, n.1, 271 pp.

OECD. 1999. Programme of Work 2000-2002, [AGR/FI/M(99)1]

OECD. 2001. Economic and social sustainability indicators for fisheries survey and issues paper AGR/FI(2001)12

OECD. 2001. Espagne -- indicateurs socio-economiques pour la gestion des peches : methodologie et application -- contribution a l'etude sur les indicateurs socio-economiques de la durabilite des peches AGR/FI(2001)12/PART1

OECD. 2001. Italy - implementation of a national observatory for monitoring techno-economic data of the fishing fleet and the evaluation of socio-economic parameters. Part I - methodology - contribution to the study on economic and social sustainability indicators for fisheries. AGR/FI(2001)12/PART3

OECD. 2001. Italy - evaluation of socio-economic indicators for sustainable fishery in Sicily. Part 2 - case study - contribution to the study on economic and social sustainability indicators for fisheries. AGR/FI(2001)12/PART3

Placenti, V. 1997. Scientific and Technical Observatory in the Mediterranean. Contract MED 93/022. Report to the General Directorate for Fisheries (DG XIV), EC, Brussels, 210 p. + annexes, in collaboration with Cofrepeche (FR), Centre d'Etudes et de projects (FR), Cetemar, (SP), IEO (SP) and ACC (GR).

SAS. 2000. SAS/IML User's Guide, version 8, 841 pp.

Särndal C. E., Swensson B. & Wretman J. 1992. Model assisted survey sampling. Springer-Verlag: New York.

Scander Ben Salem, Franquesa, R. & El Abed, A. 2002. Indicateurs socioéconomiques pour la pêche au Golfe de Gabès (Tunisie). Étude de cas

Spagnolo, M., Placenti, V. 1998. I Sistemi di Informazione Statistica della Pesca in Italia. Franco Angeli Eds. Milano, Italy, September 1998, IREPA Working Papers 2, 149 pp.

StatSoft. 1994. Statistica, User's Guide, vol. I, II, III, StatSoft Inc.

APPENDIX A

LIST OF SOCIO-ECONOMIC INDICATORS

A.1 GFCM/SAC/SCESS Report of the Second Session of SCESS, Rome, 15-18 May 2001).

Indicators

Background data by Country	***Indicators by Country***	***Algorithm***
Import/Export weight and value ***IMW, IMV, EXW, EXV***	Apparent Consumption weight and value ***WAC, VAC***	***WAC= (LW+IMW+AQW-EXW)P*** ***VAC = (LV+IMV+AQV-EXV)/P***
Yearly interest rate ***R*** Population ***P*** Working population ***AP*** Gross National Product ***GNP***	Fish Commercial Balance ***CB*** Extraversion Rate ***DR*** Fish Coverage Ratio ***CR*** Fish Contribution to GNP ***FCG***	***CB = VEX- VIM*** ***DR = (IMV+VEX)/ LV+AQV*** ***CR = LV +AQV / (+IMV+AQV-EX)*** ***FCG = LV / GNP***
Aquaculture production W&V Weight, Value ***AQW, AQV***	Ratio Fish Employment ***RFE*** Ratio Harvesting Value ***RHV*** Ratio Harvesting Weight ***RHW***	***RFE = E / AP*** ***RHV = LV / AQV*** ***RHW = LW / AQW***
Background data by Local Operational Unit	***Indicators by Local Operational Unit***	
Fleet Data:number vessel, GT, HP ***N, GT, HP***	Vessel Physical Productivity ***VFP*** Capacity Physical Productivity ***CFP*** Power Physical Productivity ***PFP*** Per vessel Hour Physical Productivity ***HFP***	***VFP = LW / N*** ***CFP = LW / GT*** ***PFP = LW / HP*** ***HFP = LW / T***
Employment ***E*** Salary Share % ***SS*** Landings Weight **LW** Landings Value ***LV***	Man Physical Productivity ***MFP*** Average wage ***AW*** Man Productivity ***MP***	***MFP = LW / E*** ***AW = (SS. LV) / E*** ***MP = LV / E***
Vessel value ex novo ***VV***	Invested capital ***IC***	***IC = (VV. N)/2***
Time at sea ***TD, TH*** days/year; houres/day	Landing prices ***LP***	***LP = LV/LW***
Cost per Day ***CD*** Yearly Fixed Cost ***YFC***	Salary Cost **SC** Opportunity Cost ***OP*** Gross Estimated Profit ***GEP*** Net Estimated Profit ***NEP***	***SC = LV . SS*** ***OP = IC . R*** ***GEP = LV -SC-(CD.TD)-YFC-(IC.R)*** ***NEP = GEP - (IC/10)***
	Profit Rate ***PR*** Gross Added Value ***GAV***	***PR = (NEP+OP) / IC*** ***GAV = GEP + (IC.R) + SC***
	Capacity Productivity ***PGT*** Vessel Productivity ***PV*** Power Productivity ***PP*** Per vessel Hour Productivity ***PVH***	***PGT = LV / GT*** ***PV = LV / N*** ***PP = LV / HP*** ***PVH = LV / T***

A.2 FAO (1999) Indicators for sustainable development of marine capture fisheries. Technical Guidelines for Responsible Fisheries No.8

Appendix B: Examples of economic criteria and indicators

Criteria	Example of Indicator	Structure	Reference Point
Harvest	• landing • by-catch	• by species; age groups • by area • by fishery sub-sector	• MSY • historical level • policy target level
Harvest capacity	• GT (decked vessels) • No of boats (undecked ves.) • total effort (see below)	• by fleet type • by fishery segment • age composition of vessels • fishing mortality/species	• capacity or effort of MSY • policy target level
Harvest value (in constant prices)	• total deflated value (landed price)	• by species groups • by sub-sector & fishery	• Selected historical level
Subsidies	• Tax rebates • Grants	• by sub-sector • by fleets/fishery	• historical level • zero level • target level
Contrib. to GDP[11]	• Fisheries GDP/Nat. GDP	• by species groups	• historical level
Exports	• Export/Harvest value	• by species groups • by fishery segment	• historical level
Investments	• Market or replacement value • Depreciation • Fleet age composition	• by fleet type • by fishery	• historical level
Employment	• Total employment	• sub-sector • fleet/fishery	• historical level (?) • realistic policy target
Net returns	• (profit + rent) • net return/ investment • value of entitlements	• by sub-sector • by fishery	• historical level • MEY
Effort (mainly at fishery level)	• No of vessels; Fishing time • Amount of gear used • Employment	• By fishery segment • In physical or monetary terms	

A.3 Commission of the European Community (2002). STECF's needs for socio-economic indicators. Fourteenth Report Brussels, 22-26 April 2002. Annex II Comprehensive socio-economic indicators by MS and by fleet segment

Table 1. National level indicators

Indicator	Explanation	Input Data Need
CONSUMPTION		
Weight Apparent Consumption	Gross consumption of fishing products per inhabitant expressed as weight of consumed fish per inhabitant	1. Harvest Production weight 2. Aquaculture Production weight 3. Import weight 4. Export weight 5. Population
Value Apparent Consumption	Gross consumption of fishing products per inhabitant expressed as expense per inhabitant	6. Harvest Production value 7. Aquaculture Production value 8. Import value 9. Export value 5. Population
TRADE		
Fish Commercial Balance	Whether exports or imports of fishing products are higher	8. Import value 9. Export value
Fish Coverage Rate	Rate of apparent consumption covered by the national production.	6. Harvest Production value 7. Aquaculture Production value 8. Import value 9. Export value
Extraversion Rate	What extent the fishing sector of a country depends upon foreign trade, both for imports and exports.	6. Harvest Production value 7. Aquaculture Production value 8. Import value 9. Export value
SOCIAL		
Ratio Fish Employment	Ratio of employment created directly by the fishing industry	10. Total Employment 11. Fish Employment
MACROECONOMIC		
Fish Contribution to the GNP	The importance of fishing production in the Gross National Product.	6. Harvest Production value 7. Aquaculture Production value 12. Gross Domestic Product
Ratio Harvesting Value	The importance of fishing in comparison to aquaculture in terms of income.	6. Harvest Production value 7. Aquaculture Production value
Ratio Harvesting Weight	The importance of fishing in comparison to aquaculture in terms of production weight.	1. Harvest Production weight 2. Aquaculture Production weight

Table 2. National level indicators by fleet segments

Indicator	Explanation	Input Data Need by Segment
PHYSICAL PRODUCTIVITY		
Vessel Physical Productivity	the average production of each vessel in terms of weight of landings.	- Weight per vessel
Capacity Physical Productivity	the average production in terms of weight of landings for each capacity unit (GT) of the vessels.	- Weight per vessel - GT per vessel
Power Physical Productivity	the average production in terms of weight of landings for each power unit (HP) of the vessels.	- Weight per vessel - HP per vessel
Per vessel fishing time Physical Productivity	the average production in terms of weight of landings for each full fishing time. Is possible select the unit of fishing time (hour or day)	- Weight per vessel - Time fishing per vessel
Man Physical Productivity	the average production in terms of weight of landings for each man employed.	- Weight per vessel - Employment per vessel
PRODUCTIVITY		
Vessel Productivity	the average production in terms of market value in the first sale for each vessel.	- Value per vessel
Capacity Productivity	the average production in terms of market value in the first sale for each capacity unit installed (GT) in the vessels.	- Value per vessel - GT per vessel
Power Productivity	the average production in terms of market value in the first sale for each power unit (HP) of the vessels.	- Weight per vessel - HP per vessel
Per Vessel Hour Productivity	the average production in terms of market value in the first sale for each fishing hour.	- Weight per vessel - Time fishing per vessel
Man Productivity	the average production in terms of value in the first sale for each man used.	- Weight per vessel - Employment per vessel
SOCIAL		
Employment per segment	indicates the employment in a specific segment of vessels	- Employment per vessel
Average Wage	indicates the average salary obtained by each man employed.	- Salary Cost - Employment per vessel
MARKET		
Landing Prices	(LP) represents the average market price of landings.	- Weight per vessel - Value per vessel
INVESTMENT		
Capital Employed	a measure of the value of Vessel, Licence, Quota, etc. would provide information of the relative position of the industry. Values above a discounted sum of the returns they could provide would be an indication of an unsustainable industry.	- Invested Capital
Capital Investments	% of change in capital employed over time - normally a year. It indicates the future expectations of the enterprises. Often difficult to measure empirically in other ways than using the capital employed at two different points in time and subtracting them from one another.	- bis. Invest Capital in the precedent periode
COST		
Income to employees	serves to identify the return from fishing to the suppliers of labour. It may be used to compare and to estimate the effect of fishing on national and local economies. The fishing industry has traditionally been identified with low incomes. This is liable to create a shortage of skilled labour where there are alternative employment opportunities.	- Salary Cost
Opportunity	the yields that the owner could obtain should he invest his	- Invested Capital

Cost	money in National Debt instead of investing in his business. This means that the owner is relinquishing that potential income. There is a profit in its economic sense when the yields of the invested capital surpass the opportunity cost.	- Rate national debt
PROFITS		
Gross Profit	indicates the total profits obtained by the whole of the vessel owners, once the operating costs have been deducted.	- Value per vessel - Salary Cost - Cost per fishing day - Time fishing per vessel - Yearly Fixed Costs - Financial cost - Indicator on Opportunity Cost
Net Profit	profitability – would provide a direct comparison with returns available elsewhere in the economy. the total earnings obtained by the whole of the owners, once the depreciation cost has been deducted.	- Depreciation - Indicator on Gross profit
Profit Rate	indicates the percent ratio of yearly net profits plus the opportunity cost in relation with the investment. It should be borne in mind that this figure does not include the additional earnings obtained by the owner as an employee in artisanal fisheries.	- Indicator on Gross profit - Indicator on Opportunity cost - Invested Capital
Gross Added Value	expresses the Added Value that the segment in question contributes to the National Economy. This includes: salaries, profits, opportunity cost and depreciations.	- Salary Cost - Depreciation - Indicator on Gross profit - Indicator on Opportunity Cost
Contribution to the margin	output minus variable costs is a short run indicator of the incentive for the enterprise to carry on. Given the problem of sunken capital in fisheries (capital written off in the books but still capable of producing output) this is an important indicator to assist in assessing whether schemes to reduce capacity will be effective. With low liabilities and low opportunity costs of labour and capital the incentive to carry on in the long run is determined by this indicator.	- Value per vessel - Salary Cost - Cost per fishing day - Time fishing per vessel - Yearly Fixed Costs
Return on Capital	(net profit plus interest payments relative to capital employed) – provides a simple and direct comparison of the opportunity cost of capital.	- Indicator on Net Profit - Financial cost
MANAGEMENT		
Value of Fish stock size	measured in value – gives an indication of the use of the production factor that is not subject to price determination on a market. Will indicate as to whether the output of fish is a result of surplus harvesting or a result of production factor reduction.	- Biological data on biomass - Prices by species
Subsidies and taxes	provide information about the dependency of the industry on public support and about the GDP in factor prices.	- Data on subsidies per segment
Capacity utilization	calculation would require distinction between long run and short run, and knowledge about the state of the fish stocks as to whether they are overexploited or not. In the short run the measure disregards fish stock effects.	- Time fishing per vessel - Maximum number of sea day

APPENDIX B

FLEET SEGMENTATION AND PARAMETERS OF THE SOCIO-ECONOMIC STRUCTURE FOR THE OPERATIVE UNITS

Working Group on Socio-Economic Indicators (WGSEI, Salerno, 11-13 March 2002)

	Group 1	Group 2	Group 3
Non engine	all		
Minor Gear	<12 meters		
Trawl	<12 meters	12-24 meters	> 24 meters
Seine	<12 meters	>12 meters	
Long line	>12 meters		
Pelagic Trawl	>12 meters		
Tuna seine	>12 meters		
Dredge	>12 meters		
Polyvalent	>12 meters		

APPENDIX C

QUESTIONNAIRES FOR THE COLLECTION OF SOCIO-ECONOMIC DATA

Several questionnaires have been presented, during the meetings on socio-economic indicators, by partners involved.

These questionnaires are aimed to collect the information required to calculate the socio-economic indicators.

In this appendix, we report the questionnaires described in the following documents:

1. Report of the AdriaMed Meeting on Socio-Economic Aspects of the Adriatic Sea Fishery Sector - Campobasso, Italy, 28th – 29th May 2001. IREPA implementation of a national observatory for monitoring techno-economic data of the Italian fleet and the evaluation of socio-economic parameters.
2. GFCM/SAC/SCESS Deuxième Groupe de travail *ad hoc* sur les indicateurs socio-économiques (Salerno, Italie, 11-13 mars 2002). Etude sur les indicateurs socio-économiques pour la pêche dans le Golfe de Gabès
3. GFCM–SAC–SCSS Troisième Session du Sous-Comité des sciences économiques et sociales (SCSES) (Barcelona, Spain, 6-9 May 2002). IREPA - Guidelines for sampling methodologies for socio-economic indicators.

***

1. Report of the AdriaMed Meeting on Socio-Economic Aspects of the Adriatic Sea Fishery Sector - Campobasso, Italy, 28 – 29 May 2001.

IREPA implementation of a national observatory for monitoring techno-economic data of the Italian fleet and the evaluation of socio-economic parameters.

Data collection and estimates of economic parameters concerning the Italian fishing fleet is produced by IREPA (Institute for Fisheries and Aquaculture Economic Research) through a National Observatory, which dates back to the early 80’s.

Sample data are recorded by means of three specific questionnaires:

1. an annual questionnaire to record technical, dimensional and vessel – management information on the sample units and relevant socio-economic aspects (number of ship owners, their ages, their property quotas and relationships between them);
2. a quarterly questionnaire to record data on fixed and variable costs, and on social aspects of property and crew;
3. a weekly questionnaire to record information reporting activity such as fishing time and area, average number of crewmembers, gears used, quantities, prices and revenues – as per species or group of species – and trade channel for sales.

In brief, the most important annual, monthly and weekly information recorded are the following:

Annual information

✧ name	✧ gross registered tonnage (GRT)
✧ maritime district where the boat has been registered, (coastal area/sector)	✧ gross tonnage (GT) based on London Convention (Reg. EC 2930/86)
✧ first year of service (therefore, age)	✧ horsepower (kW)
✧ authorised fishing gears	✧ engine make, location and type of propeller
✧ maritime district from where the ship departed for fishing	✧ communication engine
✧ maritime district where the product is landed	✧ navigation engine
✧ type of association and year of its creation	✧ fish location engine
✧ number of shipowners, their ages, their property quotas and relationships between them	✧ conservation equipment
✧ type of association and year of its creation	✧ employment contract used
✧ length overall and length between perpendiculars	

Quarterly information

✧ name	✧ fish transport cost
✧ month	✧ other running cost
✧ maritime district where the boat has been registered (coastal area/sector)	✧ labour share, wages and social insurance
✧ fuel (total and unit value)	✧ ordinary maintenance
✧ cost of nets	✧ extraordinary vessel maintenance
✧ cost of bait	✧ extraordinary hull maintenance
✧ cordage and ropes	✧ extraordinary engine maintenance
✧ food	✧ vessel insurance
✧ boxes and ice	✧ tax and other fiscal costs
✧ commercialisation costs	✧ bank charges
✧ other running costs	✧ other vessel costs

Weekly information	
✧ Name	✧ Non fishing days for bad weather
✧ Week	✧ Non fishing days for rest, repair and other
✧ Maritime district where the boat has been registered	✧ Hulls
✧ Engine used	✧ Average time (in hours) for each single trip
✧ Gear used	✧ Minimum and maximum fishing area's distance perpendicular to coast line
✧ Average crew	✧ Maritime district from where the ship departs
✧ Fishing days	✧ Maritime district where the product is sold
✧ Total hours at sea (navigation and fishing)	✧ For each single species or group of species landed: quantity, prices, income and commercial channel (wholesaler, fish market, retail dealer, others).

It is also to be noted that the input of data for the single vessel is fully computerized; the software, specifically designed for the survey's objectives, is logically structured and also includes crosscheck programs to avoid partial or inconsistent filling of the questionnaire.

2. GFCM/SAC/SCESS Deuxième Groupe de travail *ad hoc* sur les indicateurs socio-économiques (Salerno, Italie, 11-13 mars 2002).

Etude sur les indicateurs socio-économiques pour la pêche dans le Golfe de Gabès

INSTITUT NATIONAL DES SCIENCES ET TECHNOLOGIES DE LA MER

Étude de Cas – Golfe de Gabès
Pour l'estimation des indicateurs socio-économiques de la pêche
Questionnaire

A) Données techniques des bateaux

- Nom et matricule du bateau
- Nombre de marins à bord (en général)
- Longueur du bateau (mètre)
- Quels sont les engins à bord

  Chalut (C), Senne Tournante (ST), Filet Maillant Invisible (FMI), Trémail à poissons (TP), Trémail à crevette (TC) Trémail à seiche (TS), Palangre de Surface (PS), Palangre de fond (PF), Autres (AU),
- Puissance en CV
- TJB
- Quelle est la distance maximale habituellement atteinte à partir de la côte (miles)
- Nombre d'heures de travail par sortie (en comptant les heures de travail dans le port, dans le marché et autres)
- Nombre de sorties par mois
- Si ce nombre est différent pour chaque mois, quel est le nombre de sorties approximatif par mois durant toute l'année

Jan	Fev	Mar	Avr	Mai	Jun	Jul	Août	Sep	Oct	Nov	Dec

B) Données sur les coûts

- Après la vente, quelles sont les choses déduites avant la distribution des parts: Carburant (C), Vivres (V), Glace (GL), Appât (A), Lubrifiants (L)
- Quel est le pourcentage de la part de l'équipage, en incluant le propriétaire s'il est pêcheur
- Quel est le coût d'un plein de gasoil
- Combien de sorties peut assurer un plein de gasoil
- Quels sont les dépenses par jour (par sortie) de pêche, en dehors du carburant

  Appâts

  Vivres

Lubrificant

Autres

- Quelle est la valeur approximative de votre bateau à son état actuel, y compris les engins de pêche, les équipements électroniques (GPS, Sondeur, Radar, Radio, etc.) et les équipements de pêche (Treuils, Power block).
- Quel est le coût annuel pour maintenir le bateau opérationnel (assurance, poste au port, licences, papiers, entretiens routiniers et réparations du moteur de la coque et des engins de pêche , etc.)

C) Données sur les débarquements

- Quel est la production mensuelle approximative en kg, si ces débarquement connaissent une grande variation dans l'année, indiquer l'évolution mensuelle dans le tableau ci-dessous

Jan	Fév	Mar	Avr	Mai	Jun	Jul	Aoû	Sep	Oct	Nov	Dec

- Valeur de la totalité des ventes pour l'année précédente

Nom de l'enquêteur:

Port:

Date de l'enquête:

3. GFCM–SAC– SCSS Troisième Session du Sous-Comité des sciences économiques et sociales (SCSES) (Barcelona, Spain, 6-9 May 2002).

IREPA - Guidelines for sampling methodologies for socio-economic indicators. Appendix A

A simple questionnaire has been developed based on questionnaires used in other surveys; primarily those undertaken by the IREPA.

Year	____________
Quarter	____________

Name of vessel	___________
GRT	___________
Length	___________
kW	___________
Vessel age	___________
Company type	___________

MU FAO	___________
Port	___________
Principal gear	___________

Revenues		
Description	***Value***	***Notes***
from the sale of fish		
other sources of revenue such as insurance claims, compensation and government assistance		

Fixed costs		
Description	***Value***	***Notes***
Social security contributions and charges		
Routine maintenance, hull and engine		
Non-routine maintenance, hull		
Non-routine maintenance, engine		
Non-routine maintenance (other)		
Vessel insurance		
Miscellaneous taxes and dues		
Interest charges (loans, etc.)		

Production costs			
Description	***Unit cost***	***Value***	***Notes***
Fuel			
Lubricants			
Purchase of nets			
Purchase of bait			
Ropes and warps			

Selling costs			
Description	***Unit cost***	***Value***	***Notes***
Boxes			
Ice			
Fish market or wholesaler			
Transport of catches			

Labour costs		
Number employed	*Grade of crew member*	*Average net monthly pay*
…..		

How many **crew** (*excluding* the skipper) are normally employed on the vessel during a typical trip?
………………..

Number of **owners**: ………………

ownership shares: ………………

Are you (the vessel's owner) also the **skipper** of the vessel?

Yes, full-time [ ] Yes, part-time [ ] No [ ]

What is the **market value** of your vessel (including equipment and license value)?
…………….

APPENDIX D

IMPLEMENTATION OF THE BETHEL METHOD (SAS ® HML)

```
/*    SOFTWARE TO CALCULATE THE SAMPLE SIZE WITH THE BETHEL METHOD*/
/*    THE ESTIMATES PER STRATUM - SUM - THE VARIANCES PER STRATUM - VAR AND
THE POPULATION SIZE - N - ARE REPORTED IN THE FILE NAMED DATIN */

%MACRO BETHEL(DATIN,DATOUT,STR,SUM,VAR,N,ERR,NITER,CONV);

/* RICHIEDE LA SPECIFICA DEL NOME DEL FILE DI INPUT, NOME DEL FILE  */
/* DI OUTPUT, NOME DELLA VAR CHE INDICA LO STRATO, NOME DELLE VARS  */
/* DI CUI SI CONOSCE LA STIMA, NOME DELLE VARIANZE, NUMEROSITA'     */
/* DELLA POP. PER STRATO, ERRORI ATTESI PER CIASCUNA VAR, NUMERO    */
/* MAX DI ITERAZ, CRITERIO DI CONVERGENZA                           */

PROC IML;
USE &DATIN;
READ ALL VAR { &SUM } INTO SOMME; READ ALL VAR { &VAR } INTO VARIA;
READ ALL VAR { &N }   INTO N;     READ ALL VAR { &STR } INTO STR;
ERR={ &ERR };
STATS=ERR;
AIJ=(N##2)#VARIA#(1/(((SOMME[+,]##2)#(ERR##2))+(N#VARIA)[+,]));
ALFA=J(NCOL(AIJ),1,1/(NCOL(AIJ)));
DO I=1 TO &NITER UNTIL(DIF< &CONV);
   X=1/(SQRT(AIJ*ALFA)*((SQRT(AIJ*ALFA))[+,])+1E-20);
   NALFA=(ALFA#(T(AIJ)*X)##2)#(1/(ALFA#(T(AIJ)*X)##2)[+,]);
   DIF=MAX(ABS(NALFA-ALFA));
   ALFA=NALFA;
END;
NH=CEIL(1/X);
VARI2=(N/NH)#(N-NH)#VARIA;
ERR=(SQRT(VARI2[+,])/SOMME[+,]);        * può dare errore se NH>N;
STATS=STATS//ERR;
NH=(NH<>J(NROW(AIJ),1,2))><N;
NHS=(N-NH)><NH;                         * trova il minore elemento per
elemento;
VARIA=(N/NH)#(N-NH)#VARIA;
ERR=(SQRT(VARIA[+,])/SOMME[+,]);        * [+,] somma per riga prendi tutte
le colonne;
STATS=STATS//ERR;
STATS=STATS//T(ALFA);
NH=STR||NH||NHS||N;                     * affianca tutti gli elementi
elencati;
CREATE STATS VAR { &SUM };APPEND FROM STATS;
CREATE &DATOUT VAR { &STR NH NHS &N };
APPEND FROM NH;
QUIT;
PROC TRANSPOSE DATA=STATS OUT=STATS;
DATA STATS;SET STATS;
ATTRIB COL1 LABEL='ERRORI RICHIESTI'
 COL2 LABEL='ERRORI OTTENUTI SENZA CORREZIONI'
 COL3 LABEL='ERRORI OTTENUTI CON CORREZIONI'
 COL4 LABEL='IMPORTANZA DELLA VAR. SULLA DIM. CAMP.'
 _NAME_ LABEL='NOME VARIABILE';
PROC PRINT DATA=STATS NOOBS LABEL;
VAR _NAME_ COL1 COL2 COL3 COL4;
DATA &DATOUT;SET &DATOUT;
ATTRIB &STR LABEL='CODICE DI STRATO'
 NH LABEL='NUMEROSITA CAMPIONE'
 NHS LABEL='NUMEROSITA EVENTUALE SUPPLETIVO'
 &N LABEL='NUMEROSITA POPOLAZIONE'
 _NAME_ LABEL='NOME VARIABILE';
PROC PRINT DATA=&DATOUT NOOBS LABEL;
VAR &STR &N NH NHS ;
```

```
SUM &N NH NHS;
%MEND;

OPTIONS MPRINT MTRACE SYMBOLGEN;
%BETHEL(IREPA.VARIANZ1,IREPA.NBETH31,STR,S1 S2 S3 S4 S5 S6 S7 S8 S9 S10,V1
V2 V3 V4 V5 V6 V7 V8 V9 V10,
_N,0.03 0.03 0.03 0.03 0.03 0.03 0.03 0.03 0.03 0.03,300,0.000001);
RUN;

/*    %MACRO BETHEL(DATIN,DATOUT,STR,SUM,VAR,N,ERR,NITER,CONV); */
/*    elenco elementi necessari a far girare la macro
```

APPENDIX E

OPTIMAL SAMPLE SIZE AND THE DIFFERENCES BETWEEN NEYMAN AND BETHEL METHODS – AN EXAMPLE

Suppose to define a sample survey in order to estimate the total revenues of the fleet per group of species. Our target variables are:

– revenues of crustaceans

– revenues of molluscs

– revenues of other fishes

– revenues of anchovies.

We stratify our target population per fleet segment (because we know that there is, at some extends, a correlation between segments and target groups of species). The stratification is as follow:

segment	N_h[1]
purse seines	235
dredges	835
small scale fishery	12425
multipurpose vessels	3564
trawls	2364
tuna fleet	212
midwater pair trawls	145
Total	19780

What must be the sample size in order to have a maximum error of our estimates not higher than, for instance, 5% with a confidence level of 95%? To answer this question we can apply the Neyman formula (see paragraph 4.3), but, as our target variables are four, we have to apply this method four times:

crustaceans - revenues

segment	N_h	Sh^2	$N_h*S_h^2$	S_h	N_h*S_h	n_h-5%
purse seines	235	2660,5	220824,2	51,6	4281,2	8
dredges	835	0,0	0,0	0,0	0,0	0
small scale fishery	12425	686,4	608782,8	59,0	58073,2	105
multipurpose vessels	3564	27023,6	10254979,9	405,9	145590,8	180
trawls	2364	80634,5	29624128,9	706,1	194462,0	221
tuna fleet	212	0,0	0,0	0,0	0,0	0
midwater pair trawls	145	5,5	298,0	2,3	126,8	25
Total	19780	111029,9	40711713,7	1229,2	403148,8	539

other fishes - revenues

segment	N_h	Sh^2	$N_h*S_h^2$	S_h	N_h*S_h	n_h-5%
purse seines	235	40499,6	1601939,5	390,4	15117,6	8
dredges	835	0,0	0,0	0,0	0,0	0
small scale fishery	12425	11640,5	15201704,2	300,2	352201,6	165
multipurpose vessels	3564	73240,6	26638037,8	678,3	243427,1	124
trawls	2364	110333,0	35197945,5	870,8	222044,1	113
tuna fleet	212	83411,7	5320063,7	529,3	32960,4	17
midwater pair trawls	145	800,8	37013,6	50,0	2062,9	9
Total	19780	319982,2	84004549,0	2826,5	868861,8	436

[1] All figures reported in this annex are part of an exercise developed by Irepa in 2000

molluscs - revenues

segment	N_h	Sh^2	$N_h*S_h^2$	S_h	N_h*S_h	n_h-5%
purse seines	235	78,3	6481,0	9,6	749,6	25
dredges	835	3237,0	829282,9	70,6	15548,6	28
small scale fishery	12425	3036,6	2571116,7	139,8	134984,7	194
multipurpose vessels	3564	7780,5	2792485,3	260,4	87990,1	159
trawls	2364	14379,0	2566988,5	334,0	64506,0	117
tuna fleet	212	0,0	0,0	0,0	0,0	0
midwater pair trawls	145	0,0	0,0	0,0	0,0	0
Total	19780	28523,8	8766668,2	818,6	303902,5	523

anchovies - revenues

segment	N_h	Sh^2	$N_h*S_h^2$	S_h	N_h*S_h	n_h-5%
purse seines	235	130797,7	5607901,6	905,2	34180,8	132
dredges	835	0,0	0,0	0,0	0,0	0
small scale fishery	12425	1436,6	2040314,0	69,7	105012,1	145
multipurpose vessels	3564	85,8	62171,7	16,6	10176,4	39
trawls	2364	277,1	23508,8	21,6	2196,6	8
tuna fleet	212	0,0	0,0	0,0	0,0	0
midwater pair trawls	145	30253,6	1237186,5	303,4	11978,6	46
Total	19780	162850,9	8971082,6	1316,5	163544,5	371

In this way, for each variable we obtain a different sample size. In order to respect our assumption (maximum error 5% of the total values) we must consider the maximum value for each segment:

segment	Neyman	Bethel
purse seines	132	95
dredges	28	19
small scale fishery	194	125
multipurpose vessels	180	99
trawls	221	165
tuna fleet	17	12
midwater pair trawls	46	32
Total	818	547

The Bethel method has been applied on the same data using the SAS implementation (see Appendix D) and the results are reported in the previous table. Applying the Bethel method, we will obtain a lower sample size (547 units instead of 818 units). In fact, the Bethel method considers the four variables all together and therefore it minimises the variances taking into account the constraints all together.

The Bethel method requires the same basic data as Neyman, and it can be implemented on the basis of any software (see appendix D for an example).

This manual on sampling methodologies for the development of socio-economic indicators was initiated by the Sub-Committee on Economic and Social Sciences (SCESS) of the General Fisheries Commission for the Mediterranean Scientific Advisory Committee. It was prepared jointly by the Istituto Ricerche Economiche per la Pesca e l'Acquacoltura (IREPA) and theGabinete de Economía del Mar (GEM) within the framework of the ad hoc Working Group on Socio-Economic Indicators of SCESS. The manual is aimed at all decision-makers who may need collect data to build socio-economic indicators. At an international level, it can be used to facilitate and simplify reporting under international conventions and agreements on matters relating to the sustainable development of the world's fisheries. Regional fisheries bodies and stakeholders involved in fisheries decision-making, such as the fishing industry, other user groups, certification bodies, local communities and non-governmental organizations may also draw upon this manual to assist in meeting societal goals for fisheries. The methodologies can be applied to fisheries at many different levels, from individual fisheries and coastal management units to a global level. It aims to encourage consistent use of statistical methods in data collection. Governments may also wish to adapt the manual to the specific requirements of their national fisheries.

ISBN 92-5-105093-7 ISSN 1020-7236

9 789251 050934

TC/M/Y5228E/1/02.04/1300