From **Brunelli, Lina & Cicchitelli, Giuseppe (editors). Proceedings of the First Scientific Meeting (of the IASE).** Università di Perugia (Italy), 1994. Pages 47-51. This pdf file is from the IASE website at http://www.stat.auckland.nz/~iase/publications/proc1993. Copies of the complete Proceedings are available for 10 Euros from the ISI (International Statistical Institute). See http://isi.cbs.nl/sale-iase.htm for details.

ON THE PSYCHOPEDAGOGICAL ASPECTS OF TEACHING STATISTICS AT THE PRIMARY SCHOOL

Maria Pia Perelli D'Argenzio
Dipartimento di Matematica, Università Cattolica di Brescia
Via Trieste 17, 25121 Brescia, Italy

1. Introduction

The aim of this work is to present a few ideas on how to introduce the initial basic concepts of statistics; to date, studies on the teaching of basic concepts have generally been concerned more with the intrinsic aspects of the concepts themselves than with the problems involved in understanding them.

Before seeking to identify the problems involved in laying the foundation of basic concepts, we must bear in mind that all scientific research and knowledge starts out with a preliminary observation, eventually leading the scientist to formulate a hypothesis which is then followed by an examination of its validity. This well-defined course of investigation involves various stages in the mental process in which statistics can offer a valuable contribution. The aim of teaching statistics in schools is usually to provide students with tools useful in a variety of problem-solving situations. At the same time, using these tools also helps students to acquire a capacity for understanding, assimilating and interpreting information in general.

2. Teaching children the basic statistical concepts

The first step in understanding statistics is to identify the *population* being investigated and the corresponding *statistical units* it comprises. In primary school terms, this first step presents none of the difficulties of genuine research, because the type of investigation selected by teachers must be simple and must deal with easily identifiable, not too numerous, finite universes. These conditions are not always met, but a competent and sensitive teacher can easily satisfy such requirements.

The next step is to identify the *case* and the statistical *datum*. This datum has been clearly defined by the experts in terms of its intrinsic characteristics (continuous or discrete quantities, more or less easily distinguishable attributes, etc.), but the psychopedagogical problems it

entails have been far less generally acknowledged. The process by which a datum is created has certain aspects which are not easily perceived by primary-school children: in particular, the type of abstraction views the statistical unit *not as a whole* but as an *anonymous element* lacking any feature or characteristic other than the one we are interested in. This type of abstraction is particularly difficult at primary-school age, which is a time for practical operations, not for formal reasoning; moreover, the child's intuitive faculties, which are essential to learning, typically take an all-inclusive rather than an analytical outlook. It is therefore important to help the child understand that, in the process of abstraction, the feature being dealt with is of interest in its own right, separate from all the other features which form a part of, and are combined together in, the statistical unit. For a child to understand that hair colour, age or weight, are separate statistical data which cannot and must not be identified with the individuals possessing those features calls for an abstraction which has more to do with logical processes than with practical considerations and is therefore not easy.

At the age considered in this study, it is also particularly difficult to understand that in observing certain features of each statistical unit, the intention is not to evaluate those features, but rather to draw conclusions about the presence of those features in the population as a whole, because statistical study aims to provide information on phenomena concerning the group, not the individuals.

To achieve this type of abstraction, it may be a good idea to plan a class activity similar to the following, which was used quite successfully in a third-year primary school class: each child brought a full-figure photograph of himself and was first asked to cut out the figure from the background, then to cut parts from the figure in order to separate certain features. The child sees his own image (which to his mind is representative of himself) robbed of the variety of features which combine to define it, so that it becomes an anonymous carrier of the feature and hence of the datum.

Another ability that the teacher has to establish in order for the child to consciously obtain the statistical datum is *to know how to see the alternatives* that exist in the *feature* being observed. When the feature is quantitative, and the alternatives can be expressed numerically, the task is fairly straightforward because the child has grown used to the comparison between numbers in learning arithmetic; but when the observed feature is an attribute, uniformity or diversity in the alternatives becomes a complex problem which a child is not always able to solve.

3. The problem of misinterpretation of statistical information

The second problem stems from the *tendency to re-attribute synthetic values to single statistical units.* To illustrate this point, I would like to quote a sonnet written by Trilussa, a poet in Rome at the turn of the century, which clearly expresses the layman's impression of statistical information and his inability to realise its informative value.

La statistica

Sai ched'è la statistica?
È 'na cosa che serve pe' fa' un conto in generale
del la gente che nasce, che sta male,
che more, che va in carcere e che spósa.

Ma pe' me la statistica curiosa
è dove c'entra la percentuale,
pe' via che, lì, la media è sempre eguale
puro co' la persona bisognosa.

Me spiego: da li conti che se fanno
seconno le statistiche d'adesso
risulta che te tocca un pollo all'anno:
e, se nun entra ne le spese tue,
t'entra ne la statistica lo stesso
perché c'è un antro che ne magna due.

Roughly translated, the poem reads: you know what statistics is? It's something you use to take a general count of the people who were born, who get sick or die, who go to prison or get married; but what intrigues me is statistics that have to do with percentages, because then the mean remains the same for all, even for people who have nothing. Let me explain: from the way they count in statistics these days, it appears that everyone eats a chicken a year; and, even if you can't afford it, you come into the statistics anyway, because theres' somebody else who eats two.

Trilussa identifies as characteristic of statistics that they provide a means for describing social phenomena (people who were born, people who die, etc.); though this "definition" is obviously incomplete, it is nonetheless basically true. The poet then falls straight into the usual trap of attributing a synthetic value (the mean) to the individual who, together with the other statistical units, has contributed towards generating said value.

This misunderstanding is strongly rooted and is a further and almost fundamental justification for the introduction of statistical concepts at primary school level. To see the reason for this misunderstanding, we need to compare the learning of these statistical concepts with the learning of apparently similar mathematical concepts.

Learning a mathematical concept is a lengthy process which starts with observations that may appear complex, fragmented, and somewhat chaotic at first sight. This is followed by isolating a criterion on the basis of which several distinct objects can be considered "equal" among themselves but not equal to others. For istance, in introducing the geometrical notion of a rectangle, we overlook such features of an object as its colour, mass, hardness or material composition and consider only whether or not it has a particular form, which we call rectangular. We can further introduce the mathematical concept of the square by identifying a new set of four-sided objects that are at the same time equilateral and equiangular. This gives rise to a formula of equivalence which is what ensures that all the objects considered are interchangeable in terms of this criterion. Such equivalence among four-sided objects that are equilateral and equi-angular embodies the mathematical concept of the square in that these properties belong to all squares and only to squares. At this point, those properties that are characteristic of the square (in the mathematical sense) can be reattributed to all the real objects that led to the formulation of this concept.

In the formulation of statistical concepts, the first step again calls for a process of abstraction from the statistical units to statistical data and this first step is very similar to the process described above in the formulation of a geometrical concept. If, for example, we are interested in the shoe sizes of school children, we overlook the brand, type (sandals, shoes), colour or style of the shoe and consider size as our statistical datum for each child. The similarity stops here, however: when we go on to obtain decriptive statistics, which are synthetic values derived from the entire distribution of the data, we meet with conceptual difficulties. For one thing, while the arithmetic mean of a set of measurements has a precise physical interpretation, the same cannot be said of the arithmetic mean of a set of shoe sizes. Moreover, the synthetic quantity, e.g. the arithmetic mean, can no longer be re-attributed to each element of the population from which it was generated. In spite of that, young children (and adults too, as Trilussa's sonnet demonstrates) seem to have an instinctive tendency to follow their intuition and re-attribute these values to individuals. The child tends to personalise his experience, to refer all things to his own universe. "In medio stat virtus", says the Latin proverb, and the child considers this "middle" as exactly where he stands. If he detects any contradiction between the calculated "middle" and the

position derived from his experience, he becomes convinced that the statistical processing of data produces a lie and is therefore of no use.

What is lacking here is an in-depth awareness of the fact that all statistical information, be it the synthesis or the totality of a number of individual items of information, *loses its individual connotation.* Attempting to use it in characterizing an individual is dangerous because it would involve a high degree of uncertainty. Proceeding with any inductive inference, i.e. drawing any conclusions from matters that are not directly observable, though their existence can be assumed, leads to the risk of error. Such inductive conclusions can only be interpreted correctly in probabilistic terms. Such people are confronted with decision-making situations of this type almost every day, it is essential for them to be trained to do so properly at the primary school level.

The primary school, then, is not required to train statisticians and provide high technical training; it should provide the qualitative education that would enable the child to overcome common, intuitive ideas. The final objective of a basic educational program could be, for example, to create a teaching unit capable of transmitting the two following concepts.

Statistical information is directly informative in terms of the community, not in terms of single individuals. It results from a combination of individual data which are generally variable, in that data from different cases are in general different. In addition, as soon as two diverse items of information are combined, the resulting synthetic information, by itself, cannot be reliably used to draw conclusions about the single individual observations.

But the tendency to re-attribute summary information back to individuals is common in adults and even more so in children, raising difficulties in understanding the meaning of statistical concepts, rather than in learning them. The risk inherent in teaching statistics in primary schools is that the child may tend to separate the notions learnt at school (surveys, calculation of mean values and other statistics) from what he really experiences and believes. In fact, when he discovers that what he has learned does not confirm his own experience (see Trilussa), he tends to rely with more conviction on past experience alone than on what he has "worked out" at school. He is unable to link subjective judgements and personal experience with what he is taught at school, which often leads to his accepting the theory at school (and being able to accurately calculate means, for example), but rejecting these skills for practical situations, and that means that the teaching method has failed. It is more important to implement teaching strategies designed to overcome these psychological difficulties than to provide technically valid instruction.