A Project of Merged Approach to the Teaching of Probability and Statistics in the Italian Secondary Schools

Romano Scozzafava - Rome, Italy

1. Introduction

We report on a project of combined teaching of probability and statistics in the Italian secondary schools. The aim is to bring a group of secondary school teachers up-to-date, without getting involved with complicated arguments and trying to overcome the barriers created by prevailing opinions (such as the combinatorial or frequentist approaches) concerning the teaching of probability and statistics. The project involves the study of specific teaching patterns, which aim at inculcating the idea of inductive reasoning as complementary to deductive reasoning, and not providing a kit of statistical tools for use in practice. In fact, teaching based on subjective probability and Bayesian statistics allows the coverage in a short time of some of the most significant aspects of statistical thinking. This can be done through an inferential model based on the iterative use of Bayes' theorem: hundreds of experiments can be simulated on a computer using elementary software. New statistical data can be used to evaluate current (posterior) probability; then taking this as the starting point (prior probability), further experiments produce another posterior probability, and so on.

2. An example

A file is randomly chosen among ten, each containing thousands of data concerning a given illness and referring to an equal (approximately) number of males and females; it is known that there is one file, F_0, undistinguishable from the others, whose data refer only to males (for example, levy soldiers). Six cards are drawn from the randomly chosen file and it turns out that they *all refer to male individuals* (let E denote this event): what is the probability that *the randomly chosen file is* F_0? Let H be the latter event: before the occurrence of E, we clearly have P(H) = 1/10; to compute P(H|E) we need to evaluate the other quantities which appear in Bayes' theorem.

$$(1)\qquad P(H|E) = \frac{P(H)P(E|H)}{P(H)P(E|H) + P(H^c)P(E|H^c)}\,.$$

Obviously, $P(H^c) = 9/10$ and $P(E|H) = 1$ (since it is certain that the six cards refer to males if it were known that the randomly chosen file was F_0). Moreover $P(E|H^c) = 1/2^6$, since among all the possible sequences of six cards formed by males or females (we assume equal probability of these sequences) there is only one sequence with all males. Then from (1) we get

$$(2)\qquad P(H|E) = \frac{1/10}{(1/10) + (9/10)(1/2^6)} = \frac{64}{73} = 0.877\,.$$

In conclusion, the *initial* evaluation P(H) = 1/10, that did not take into account the information acquired through the observation of E, is now updated to the *final* evaluation P(H|E) = 64/73.

A very important remark is now in order: in no way could this value of P(H|E) be expounded in terms of combinatorial or frequentist evaluations of probability. Otherwise, in contrast with reality, we should think (in combinatorial terms) as though we had drawn our file from a collection (of ten files) contained in a set of 73 possible outcomes, i.e. 73 collections such that for each one it *should* turn out that six males *would* be observed in a randomly chosen file among the ten, with 64 among these files being of the kind F_0; or (in frequentist terms) *we should actually have at our disposal* 73 collections "similar" to the given one, such that in each randomly chosen file among the ten, six males *will* be observed, and then verify (by a complete scrutiny of each card) that 64 among these files are of the kind F_0. Clearly, no one would subscribe to such intricate and involved combinatorial or frequentist arguments in order to give meaning to the value (2), which can instead be easily interpreted as "degree of belief", in the intuitive sense that even the layman gives (perhaps unconsciously) to probability evaluations.

3. Combined teaching of probability and statistics

A unitary view of probability and statistics as a *different way of thinking* (inductive and not deductive) suggests a merged approach to their teaching. This can be achieved in a short time through an inferential model based on an iterative use of Bayes' theorem: so, given certain events looked upon as *hypotheses*, their probability is interpreted as a *degree of belief* and evaluated by conditioning on another event which represents statistical data. A concrete and not stereotyped teaching approach (too often too many examples are related to gambling, offering students a biased impression of the subject!) should start from the general meaning of *event* (which goes beyond the two particular cases corresponding to "equally probable outcomes" or to "observations that can be repeated under similar conditions": so also a *statistical hypothesis* is an event), and then introduce the subjective and intuitive "real life" meaning of probability.

4. A project for the Italian secondary schools

With this object in view, we set up two years ago a project, approved by the Italian CNR (Consiglio Nazionale delle Ricerche), with the aim of bringing a group of secondary school teachers up-to-date: about thirty people (including some university professors and researchers) take part in the project, whose realisation is still in progress. To initiate these teachers in the aforementioned study, the project started by acquainting them, through meetings and seminars, with some controversial aspects of the usual approaches. In particular, it has been underlined that, when relying upon the assessment of the *probability of an event* through the *frequency observed in the past* (for events that are considered, in a sense, "equal" to that of interest), this choice of "past results" is necessarily *subjective*; and also the choice, in the "*combinatorial*" *assessment of probability*, of the possible cases (for which equal probability is assumed) is clearly subjective. In other words, it is essential to give up any artful limitation to particular events (not even clearly definable) and try to ascribe to probability a more general meaning, as a sensible way to cope with real situations.

5. Probability as "degree of belief"

The *probability* of an event E is introduced (Scozzafava, 1990) as an amount p which makes *coherent* a suitable bet on E, i.e. the choice of p *would not make the player a sure loser or winner*. Notice that this definition is based on *hypothetical* bets: the force of the argument *does not depend on whether or not one actually has the possibility or intends to bet*. In fact a method of assessing probabilities making one a sure loser or winner *if he had to gamble* (whether or not he really will so act) would be suspicious and unreliable *for any purposes whatsoever*.

Notice that *no restriction* (such as equally probable, or "repeatable", or anything else) *has been imposed on the kind of events involved*. In other words, *we have at our disposal the most general interpretation of probability*.

To face the problem of choosing, for each event H, a suitable value of its probability p *among all coherent evaluations*, it is not enough directing our attention toward the event H; we need to take into account also *other events* which contribute in determining our information on H. To this end a fundamental tool is that of *conditioning event*: the ensuing concept of *conditional probability* and simple arguments involving only coherence lead to *Bayes' theorem*, that allows, given an event E (usually representing observed data), to evaluate the conditional probability P(H|E) of another event H (often called *hypothesis*). This is the procedure that has been followed in discussing the example in Section 2.

It is possible to give a different *probability evaluation* of H for each different "state of information" expressed by E: almost all Bayesian statistical procedures are, essentially, extensions of this way of thinking.

6. Concluding comments

Teaching is not a "linear" process: a merged approach to probability and statistics fosters a continuing education of the students to a versatile and correct way of thinking. In the traditional teaching there are some controversial aspects that may be dangerous: in this respect it may not be surprising that empirical research done at Brunel University shows (as noted by Flavia Jolliffe in one of her papers at this Conference) that those students who had received instruction in probability gave a larger number of incorrect responses than those students who had had no instruction!

References

De Finetti, B (1973) Bayesianism : its unifying role for both the foundations and the applications of statistics. *Bull. ISI, Proc. 39th Session (Vienna)*, 349-368.

De Finetti, B (1976) Probability : beware of falsifications! *Scientia* **111**, 282-303.

Scozzafava, R (1980) Bayesian inference and inductive 'logic'. *Scientia* **115**, 47-53.

Scozzafava, R (1983) Come introdurre alcuni concetti fondamentali di probabilitá e statistica nelle scuole secondarie. *Period. Matem.* **59(3-4)**, 3-15.

Scozzafava, R (1990) A merged approach to stochastics in engineering curricula. *Europ. J. Eng. Educ.* **15(3)**, 241-248.

Steinbring, H and Von Harten, G (1982) Learning from experience - Bayes' theorem : a model for stochastic learning situations. In: D R Gray, P Holmes, V Barnett and G M Constable (eds) *Proc. First Intern. Conf. on Teaching Statistics*, Sheffield, UK, 701-713.