Conceptual, Computational and Didactic Aspects of Teachers' Training in Probability and Statistics

Talma Leviatan
Tel Aviv University, Department of Statistics and OR
School of Mathematical Sciences
Tel Aviv, 69978, Israel
talmale@post.tau.ac.il

In determining the objectives and contents of a secondary school teachers' program in "Stochastics" (Probability and Statistics), one has to take into consideration that the program should supply a sound basis for 10-20 years of a professional career. It is also important to acknowledge the importance of probability - as a discipline by itself, for a better understanding of the world. So the first stage should be a sound probability course- taught as the relevant mathematical theory for dealing with chance situations (and not as a special case of measure theory).

There are several theoretical approaches to probability theory. Many find it confusing, some going as far as arguing which approach is best (see [Shaughnessy (1992)]). A good program should emphasize the need to get acquainted with the various approaches, because each is relevant in its own context. There is no contradiction between them, but rather, they complement one another like a new layer on top of the previous one. To naturally combine them, we suggest starting the program by adopting a historical viewpoint, strengthened by a body of examples, which will demonstrate the practical need of each of these approaches under various chance situations.

We advocate the traditional **modeling approach** to probability. Each probabilistic model presented should be motivated by *many* "real life" examples (not "urns and balls" problems!), problem-solving situations, and analyses of **simple** chance data in everyday life. In starting such a program there is no need to look at databases for motivation; the world around us provides ample motivation and data to demonstrate any model that we want to introduce throughout the course.

When it comes to the next course in statistics, the issues are more complicated. We found it beneficial **to start** most new topics in statistics by first working on a specific set of **real life data**, using **worksheets** as guidance (for a demonstration see e.g. [Evangelista, Hemenway (2002)]), thereby "discovering" new ideas (sometimes **working individually** and other times **together in class**). A teachers' course in statistics should proceed to supply the **theoretical background** of the models presented, taking special care not to become a "cookbook course" - based solely on intuitions and recipes. This will require formal presentation of definitions, theorems and models (including discussion of the underlying assumptions).

We present a 3-4 semesters (2 h/w) program that works well in achieving the above goals.

1. Conceptual aspects

We start with a full year post-calculus probability course, followed by a one-semester statistical inference course. Given in the framework of a Mathematics program, it is necessary that the **foundations** of probability be presented rigorously, focusing on the main probabilistic **concepts** (sometimes highly non-intuitive). A whole semester is devoted to laying these foundations.

Probability theory is presented first from a **historical viewpoint,** starting from the consideration of **odds** (which are very intuitive to students), leading naturally to the "**Classical Theory**" and then proceeding to the "**Relative Frequency**" approach. While solving "real life" problems, we soon realize that the classical theory is valid and useful, but not general enough for our purposes; the relative frequency approach is useful in solving real life problems, but is problematic and cannot be turned into a proper definition. The relative frequency approach, though, is **the key motivation** for the **unavoidable** subsequent "**Axiomatic"** approach. We supplement the introductory part of the course by examples of other theories in Mathematics that started with a naïve approach and evolved into axiomatic theories. The formal axiomatic approach to probability theory is important, as they are the basis for all other probabilistic rules. At this stage the axioms are simple and natural assumptions for describing real life chance situations.

Though we finally adopt the axiomatic approach, both the classical and frequency approaches continue to play central roles throughout the course in problem solving. The **Bayesian viewpoint** to Probability is also discussed at this introductory stage, and is kept in mind throughout the course.

Later on, special emphasis is put on the concept of **conditional probability** (CP) and on independence, which are key issues in understanding probability and Statistics. We first present many chance situations in which CP can be computed using **heuristic methods**. Then we explore the limitations of these methods, which lead us to the necessity of a general formal definition. The motivation for heuristic methods and for the formal definition both come from considerations of odds and from the two historical approaches. It is important to discover in class that the previous "heuristic solutions" agree with the formal solutions. Moreover, after becoming accustomed to applying the formal definition, it is a good idea to work on "extending intuitions" on the subject of conditioning, to be able to later apply "heuristic solutions" in more general situations.

The concept of **independence** is also quite tricky for a novice. It is well known that in some situations people tend to assume false independence while in other situations they assume false dependence. Generally speaking, we found that a good context, in which the concept of independence can be formalized, is a discussion on the economical **"value of information"**. This helps avoiding the common misconception of confusing the ideas of dependence and causation.

Another important theoretical issue is that of **measures of location and variation**. The formal definitions should be accompanied by a good discussion (and well-chosen examples) comparing the three common measures of location, and their properties (additivity, etc). Knowing these properties turns out to pay dividends in practical problems too, yielding alternative computational methods, which do not require a detailed knowledge of the distribution.

From a practical point of view, the core of the course is the introduction of "**families of distributions**" (discrete and continuous). Again, the important task here is a discussion of the **underlying assumptions** of each such model accompanied by *many small* "real life" examples.

The course ends with the important **laws of large numbers**; this nicely rounds up the course, as it returns us to the relative frequency motivation for the axiomatic approach to probability.

2. Investigative projects and employment of "judgment problems"

During the first semester students are often involved in "small **theoretical investigative projects**" in which they have to explore (at home, and together in class) how far the current approach can be extended, and what the limitations of the current state of theoretical knowledge are. The first project, for example, is to try *use the classical approach* **only** to deduce the probability of:

(i) Getting, say, two white balls when drawing *together* two balls from an urn containing two *identical* white balls and three *identical* black balls;

(ii) All possible outcomes when repeatedly tossing a fair coin until it lands on "heads".

Note the lack of symmetry in the sample space of (i) and the infinite number of outcomes in (ii)

An important part of teachers' professional education is aimed at realizing how Mathematics is related to their own life. So another goal in teaching Stochastics is to expose students to chance situations in their everyday life and discuss how to **interpret and analyze** them. This may include the proper understanding and criticism of statistical information presented in the **media**, and also the ability to make **knowledgeable decisions** under conditions of uncertainty. To achieve these goals, students are regularly presented with **questionnaires** throughout the course in which "real life like **judgment problems"** are given. Students' individual replies along with the statistics of class responses are regularly discussed (see "Feedback Enhanced Methods" [Leviatan (2003)]).

3. Computational aspects - problem solving

The theory offers many useful **probabilistic models** - Binomial, Geometric, Exponential, etc. After experiencing a wide range of applications, many students get the wrong idea that it is possible to solve any real life problem just by identifying the right model (presented in class). What they really do is "force" a model on the given problem, without seriously checking the underlying assumptions. Perhaps in an Engineering course (and other "application oriented" populations) this may suffice, as indeed many real life problems fall under one of the standard models. But for prospective teachers, a good probability course should quickly shift from demonstrating more and more models to supplying tools that also help to **solve non-routine problems**. In problem solving sessions throughout the course, emphasis should be put on the **probabilistic principles** that lie behind the methods of solution, e.g. choosing an **equivalent** (more intuitive) problem, using **symmetries**, **focusing** on the event itself (disregarding irrelevant data), etc (see [Leviatan (2000)]).

Another topic that requires less routine treatment is the **expected value.** It is important to get away from the need for the definition itself for its computation (it requires knowledge of the actual distribution) and to proceed to more realistic situations in which a non-direct approach is used (requiring less information) - the use of indicators or the use of the "total expectation formula".

4. Didactic and technological aspects

Prospective teachers should obviously be acquainted with the many common **misconceptions** in the theory and with **common mistakes,** and be able to deal with them effectively. They should also be aware of some famous probabilistic **puzzles** and **paradoxes** (see [Leviatan (1998), (2002)]).

What is the role of **technology** in the process of learning probability? The frontal part of the probability course uses almost no technology. There is a website, which accompanies the course, and is used mainly as a convenient means of dealing with assignments and of distributing enrichment material (including links to interesting sites on the Web). Students are encouraged to use those links, to experience simulations of chance experiments (including simulations of famous paradoxes and puzzles.) etc. Prospective probability/statistics teachers should be able to later use and demonstrate all kinds of **statistical software** including simulations, random generators, etc. Acknowledging the fact that mathematics teachers should be able to use the multitude of resources the Web has to offer, we designed in our Teachers' college Mathematics program, a special year-long course entitled "Using Mathematical software in the teaching of Mathematics", in which such resources are presented. The Prob/Stat section of this course is carefully coordinated with the regular courses, but the fact remains that these resources are discussed in a different course with a different teacher. One can argue, of course, that mathematical technology should be presented "across the curriculum", but we have found that the above method is more practical to apply. Also, that in this way, technology (which is time consuming) does not get priority over the basic theory.

5. The teaching of statistics

The next stage in the Stochastics program is a one-semester course in statistics. The whole semester is devoted to **statistical inference**, as **descriptive statistics** (and also the concept of a p-value!) is actually cleverly interwoven into the probability course.

The issues in teaching statistical inference are more complicated. Some results (like the standard point estimators) are so natural that it is beneficial for students to "discover" them by themselves (working *individually*) starting with some **specific real life data** and using well-designed **worksheets** as guidance (see e.g. [Evangelista, Hemenway]). Other statistical concepts are, at first, quite artificial for a novice (like MSE approach, maximum likelihood estimators, significant level tests). But here again it is easier to introduce them first (working *together in class*) using specific real life data and well-designed **worksheets** as guidance. So in many cases it seems beneficial to start a new subject in statistics by first working on a specific set of real data, guiding the students to formulate the "right concepts". Of course, a teachers' course should later proceed to supply the **theoretical background**. This requires discussion of the central limit theorem, getting acquainted with new distributions (like t and c^2), discussion of properties of estimators and estimation methods, definition of the maximum power test, etc. Note that there is no need to prove the two basic theorems, the central limit theorem and the Neyman-Pearson lemma, as the proofs are not statistically instructive, but a discussion - using a graphical approach (see [Raviv, Leviatan (2002)]) of why those theorems **"must be true"**, should be supplied. Another important issue here is making sure that prospective teachers know the underlying assumptions of each statistical model, and understand the need to check them before applying any statistical procedure (e.g. t-tests).

6. Conclusion

A program in Stochastics for secondary school **teachers** should aim at the right balance of **all** the above aspects. To achieve all this, it is crucial to have well-tailored **textbooks** (see e.g. [Raviv, Leviatan (1996), (2000)]) - which may be a problem in non-English speaking countries. It is also very helpful to have a good set of worksheets to accompany the statistics course. Needless to say, it is highly desirable to have an **experienced teacher** whose expertise is Stochastics (unfortunately this is not always the case) because these theories are so different to other Mathematics subjects. Having an accompanying **website** is very helpful too, both for the reasons mentioned above, and also to enable additional interactive sessions outside the classroom for students with special needs.

Regarding a course for **secondary school students**, it is clearly not possible (and not desirable) to aim at all these goals. A school course should be tailored according to the specific level of the students and be coordinated with other Mathematics courses in school's program. It is fine to offer a course dealing solely with descriptive statistics using some real database, but it is important to ensure that students do not attempt to draw statistical conclusions from their data. It is strongly advisable that **all** secondary school students should have some exposure to probability theory (adopting at least the classical approach, but hopefully, also the relative frequency approach).

In view of the long-range goals of teachers' training programs, secondary school teachers should be prepared for any type of Stochastics courses needed in the future, be it theoretical or experimentall and be aware of the didactic aspects of both subjects (see [Leviatan (1998)]). Current fashionable instruction trends should play a limited role in such a program.

REFERENCES

Evangelista, F., Hemenway, C. (2002). The use of the Jigsaw in Hypothesis testing. Proceedings of ICTM-2.

Leviatan, T. (1998). Misconceptions, fallacies, and pitfalls in the teaching of probability, Proceedings of ICM.

Leviatan, T. (2000). Strategies and principles of problem solving in probability, Proceedings of ICME-9.

Leviatan, T. (2002). On the use of paradoxes in the teaching of probability, Proceedings of ICOTS-6.

Leviatan, T. (2003). Judgment Problems and Probabilistic Puzzles - Two Effective Tools in Teaching Probability. To appear, Journal of Statistics Education.

Raviv A., Leviatan T. (1996), (2000). Introduction to Probability and Statistics. Part I-Probability. Part II- Statistical inference. Amihai press (In Hebrew).

Raviv A., Leviatan T. (2002). Dilemmas in teaching introduction to statistical inference to economics students. Proceedings of ICOTS-6.

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. Handbook of Research on Mathematics Teaching and Learning. Macmillan, New York. (pp.115)

RÉSUMÉ

Les domaines des probabilités et de la statistique supposent un cadre mathématique et des outils pour analyser les situations de la vie quotidienne où le hasard intervient. Ces objets, relativement nouveaux dans les programmes scolaires, peuvent subir de nombreuses évolutions. Notre objectif est de proposer des bases solides pour l'enseignement dans les 10 ou 20 ans prochains. Nous proposons un programme de formation des enseignants sur 3 ou 4 semestres, qui prend en considération les aspects conceptuels, informatiques et didactiques de cet enseignement.