

Why Do Students Find Statistics So Difficult?

James B. Ramsey
New York University, Dept. of Economics
269 Mercer St.
New York, NY 10003, U.S.
James.Ramsey@Econ.nyu.edu

1 Introduction

It is a common experience for statisticians when introducing themselves to a nonstatistician at a reception to be greeted with the observation; “I took statistics once and hated it; all those formulae to memorise.” The questions are why we get this reaction and what can we do about it. In my answer below I am presuming it is useful for society at large, much less the profession itself, that nonstatisticians should have a reasonably sophisticated understanding of statistics and probability. Further, it is a depressing fact that corporations, government agencies, and other organisations are staffed with individuals, even those holding MA degrees, who do not understand what it is that they are doing and why. Again we should query why the state described is universally true.

The answer lies in the conjunction between the very special nature of the discipline and the unwillingness of instructors to recognise the implications of that fact for the teaching of statistics. Statistics cannot be taught like mathematics, or like physics. The two key concepts to be kept in mind in the sequel are that the subject matter of statistics *per se* is *acausal* and that distributions are key to the understanding of statistical concepts. Much of what I have to say will be considered obvious to my colleagues; but we forget that these matters are not obvious to the student. Indeed, it is this wide gulf between what has become intuitive to the professional statistician and the mind set of the layman that is the major source of the difficulty in communication.

Because the student finds the typical exposition so baffling, the student's defence is to rely solely on the memorisation of formulae. The memorised formulae are then inserted into “problems” for which rules of thumb have been developed, often aided by the instructor, to know which formula is to be plugged into which problem. This is not a bad strategy for getting a good grade in the course. However, once the memorisation strategy has been espoused, learning by the student is now entirely lost.

The remainder of this paper is dedicated to elaborating the key intellectual distinctions between the subject matter of statistics and other seemingly related disciplines. Building on these insights I attempt to deduce the implications for the teaching of statistics together with practical suggestions for their implementation.

2 The Unique Characteristics of Statistics

Students and the lay public are familiar with simple notions of causality; indeed, they can be said to order their lives on the basis of looking for causality to “explain” every event, especially important events with negative consequences. In particular, for events that do not obviously have a physical interpretation, they look for the role of a

human agent as the causal factor; the universal idea is that some person is to be blamed or thanked, but in any event credited with the observed outcome.

Statistics and probability theory are essentially acausal. This simple fact is the first major stumbling block to the layman's understanding of random phenomena. The shift from disciplines with an all pervading causal interpretation to one that is inherently acausal represents a major fundamental shift in viewpoint and one that cannot merely be dismissed as an alternative "explanation." It is easier for a student to accept alternative and often contradictory theories coming from different disciplines' explanations of some sequence of events than it is to accept the notion that the events are random. Notwithstanding everyone's exposure to what is regarded by the statistician as overwhelming evidence of randomness, that notion represents a deep insight on the part of the student that is in fact difficult to achieve. This is the reason why teaching statistics is not like teaching mathematics and not like teaching physics either.

Arising out of this insight is the concept that distributions and their links to physical experiments are crucial to the understanding of statistical reasoning. The very notion of a distribution is mysterious to the student. The idea that individual phenomena can be independently distributed, but the "distribution" of large collections of these random events can be described mathematically with great regularity, is at best a mystery wrapped in an enigma. Independence is the quintessential example. Many psychological experiments have indicated that people expect samples, even small samples, to be representative. Consequently, the ability to reconcile independence with a given distributional shape is a challenge. A common misunderstanding is captured by: "If ten heads have been thrown in a row, the next few tosses have to be tails for the results to represent the distribution."

The second source of difficulty is due to the fact that statistical reasoning is very abstract. Even though we attempt to relate our probability theory to observed events, the exercise is one that is quite foreign to the average student. The connection between theory and observation is not easily made; no matter how many times we, as statisticians, profess it. Events occur and the student neither knows how nor why. The link between specific observation and general inference is opaque to the student. Our presentations are not aided by the decrease in training in abstract thinking that is the hallmark of conventional education.

Consider a simple pair of examples. In explaining Hooke's Law, it is easy for the student to visualise the process of extending a spring and measuring the force needed. Alternatively, imagine an equally simple statistical explanation of the outcomes from the toss of a die; for the student there is no explanation at all. The outcomes from the experiment are as mysterious after the statistician's "explanation" as before it. The student looks for explanations that rely on the coin tosser's inability to control his toss precisely, or the coin tosser's deliberate efforts to alternate setting heads and tails when initiating the toss, and so on.

Finally, the most difficult aspect is that of tying statistical reasoning to physical cognate disciplines such as physics, chemistry, biology, and economics. Because statistics is acausal and the cognate disciplines are inherently causal, the link between

the two is difficult for the student to fathom. The difficulty is greater in those cases where the formulation of the cognate discipline's theory must itself be couched in terms of distributions, as is true for statistical mechanics, biology, or economics.

3 The Implications for the Teaching of Statistics

The comments above indicate a broad strategy for the teaching of statistics.

The first task is that the concept of randomness must be dealt with immediately, directly, and extensively. The student has to recognise that she is being asked to consider phenomena that are very different to those that she believes that she understands. Illustrating random outcomes through a variety of physical mechanisms is best. The student needs to be able to observe directly the mechanism generating the random numbers and if she can manipulate the parameters of the system herself that is better still. The major lesson to be learned is that while individual outcomes are unpredictable, the collections exhibit regularities that require explanation. It is best for the students to discover on their own these twin, seemingly irreconcilable characteristics, of random data. Further, it is critical that the student recognises that the same experiment repeated yields the same distribution, but that (statistically) different experiments yield different distributions. Repeated physical examples, including computer generated experiments under the control of the student, seem to be the best way to achieve the desired outcome; lecturing is of little help. The greater the variety of experiments that are illustrated the better.

The next essential element underpinning the student's understanding of the process of statistical analysis is that the theory underlying the formal calculations is absolutely essential to the interpretations that are made. More than any other discipline is it true that without the supporting theory, there is very little that we can say about any observed situation itself, much less the attempt to make inferences about unobserved events. From a relatively small number of observations, we presume to infer for a broad class of events, the relevant distribution or stochastic process, its parameter values and to postulate "optimal" decision making procedures. To the student and the layman, this borders on magic, or academic hubris at the worst.

A key precursor to a successful statistics course is the mathematical training of the student; this is even, perhaps more, important for supposedly "non-mathematical" courses. The problem to be addressed is this. The mathematical training of far too many students leaves them unprepared for dealing with abstractions in general and algebra in particular. This is even true for students that have had one or two calculus courses; they have memorised a string of formulae, but have little understanding of basic principles. I always begin my courses with laboratory sessions that make sure that the students are trained in the use of algebra and are comfortable with sigma notation. I have discovered that the way to do this is to begin with the concrete and then demonstrate the analytical advantages of abstraction; the key is to begin with actual numbers and then to demonstrate the role that algebra plays. It also pays to review with the students the principles that underlie calculus. In an introductory statistics course the fundamental concepts are most important, not the formalism in terms of which calculus is usually taught. A corollary aspect is to introduce no concept, function, or formula that cannot at that moment be given at least an intuitive

explanation. For example, I do not introduce the statistic s^2 , as opposed to the regular second moment, until I have explained expectation and bias. At that time the students can discover the rationale for s^2 themselves.

While it is a truism that students only learn through doing exercises; not much thought has been put into designing exercises to meet specific objectives. I note that there are three basic functions that exercises play. And it is neither efficient nor effective to try to combine them into the same exercises.

The first is the simplest of providing the student with the opportunity to gain mechanical skills in manipulating data and calculating; students need to practice actually calculating means, variances, correlation coefficients, Student's t and F ratios, and all our other formulae. The second task is to help the student explore the properties and characteristics of the statistics that they are using; I call this "exploring the tools." The idea is for the student to examine the behaviour of the statistics under various special cases, to discover their limitations, and to achieve a sense of what it is that the calculations are providing. For example, students should be able to "guesstimate" the effect on the variance of adding observations that are bigger, smaller, same size, or zero. Similarly, students are asked to determine the effect on the correlation coefficient from adding and subtracting data of different relative sizes to a prior calculation; to be able to guess the order of magnitude and sign of a correlation by observing a scatter diagram; to guess the sign and magnitude of the higher moments by an examination of histograms; and so on.

The third task, to be tackled only after the others have been successfully completed, is called "Applications." But even here more thought needs to be put into what is done. I believe that there is nothing more destructive to the interest of the student than to be asked repeatedly to calculate means and variances and write down 5% confidence intervals, even if the description of the postulated context varies. The application questions should presume that this mechanical part of the answer has been done successfully; the meat of the question should concern the all-important interpretation and possible *use* of the calculated statistics. The key question is: "What did we learn about the process that generated the data?" Calculating statistics is a means to an end, not an end in itself.

RÉSUMÉ

Après avoir constaté le manque d'intérêt par le non-spécialiste pour l'analyse statistique, nous en discutons les raisons. L'analyse statistique est unique de par sa nature marquée par l'absence de lien de causalité et de par son caractère abstrait pour les étudiants. Les statistiques ne peuvent donc être enseignées comme les mathématiques ou la physique. Par conséquent la notion et les différentes facettes du concept d'aléatoire doivent être découvertes par l'étudiant lui-même. Il en est de même pour le fait que l'interprétation d'événements aléatoires dépend fortement de la théorie sous-jacente. Les implications de cette approche pour l'enseignement des statistiques sont indiquées. Dans ce contexte, pour être efficaces, les exercices doivent être adaptés spécifiquement à chaque point d'étude.