Informal Probability in the First Service Course

K. L. Weldon
Department of Mathematics and Statistics
Simon Fraser University
Burnaby, BC. Canada. V5A 1S6

e-mail: weldon@sfu.ca

Author's Footnote: K.L. Weldon is Associate Professor, Department of Mathematics and Statistics, Burnaby, BC, Canada. V5A 1S6 (e-Mail: weldon@sfu.ca).

Abstract

This paper suggests a way to include probability ideas in an introductory "service" course in statistics. Many modern service courses focus on data analysis and inference to the exclusion of exposure to the consequences of randomness in real life. It is argued here that a very elementary understanding of probability is all that is required to understand some important features of sports, lotteries, education, investment, medicine, politics, and insurance. For students taking only one course in statistics, these "randomness" topics are more useful than confidence intervals and hypothesis tests, and time saved on a reduction of formal inference procedures can be reallocated to randomness phenomena. This randomness material can complement some descriptive data analysis, to provide a really useful introductory course in statistics.

KEY WORDS: Service course; Simulation; Inference; Sports; Investment; Lotteries.

1. INTRODUCTION

Traditional introductory courses in statistics focussed on inferential procedures, especially estimation and statistical testing. Recent developments have increased the emphasis on descriptive statistics, especially graphical methods, at the expense of some inferential techniques (Cleveland(1993), Cobb(1993), Hogg(1994), Singer and Willett (1990), and Griffiths, Stirling and Weldon (1998)). To quote Hogg (1994), who chaired a curriculum review committee for an engineering statistics course:

"...all of us agreed that most of the material concerning the formal laws of probability, joint distributions, and reliability should be eliminated. Moreover, the more mathematical parts of testing hypotheses, like Type I and Type II errors along with power functions (or OC curves) should be scrapped in this course."

Another modern trend is to include more discussion of study design. (Lawrence(1996), Ledolter(1995), Weldon(1986), and the text by Moore and McCabe(1993)). Both of these modern trends, more data analysis and more about study design, have tended to displace the more traditional topics of inference and probability models. It does seem that a lot of data analysis and inference can be taught assuming a small set of probability models: the normal and binomial models suffice for introductory courses. However, data analysis and inference are not the only things one can do with probability, and the de-emphasis of probability may have inadvertently led to the omission of the key to an understanding of some very important aspects of modern life. If we pass up the chance in our first course to convey important practical information while we are describing randomness, we are missing an opportunity to demonstrate the relevance of our discipline to a general audience.

The influential CHANCE course resources available on the world wide web emphasize experiential learning of statistical concepts based on articles in the popular press (Snell (1992)). In addition to articles involving data analysis and inference, many of these articles involve probability, and some even include mind-bending paradoxes . The data-analytic experience of the CHANCE course certainly encourages the student to think seriously about the hazards of cavalier data analysis. In fact, the student may learn that data analysis is a skill requiring great sophistication, and that one course does not develop enough skill for practical applications. Whether or not the student takes more courses in statistics, a realization of the broad skills needed for good data analysis will be a valuable lesson. However, it is likely that the student will not become a good data analyst or a probabilist as a result of the one course. Are there other goals that might be achieved in a first course in statistics, that might have a more direct value to students who only take the one course in statistics?

An appreciation for the impact of randomness in everyday life can be gained from a thorough introduction to applied probability. However, many students taking a single course in statistics will not have the aptitude and the inclination to master the mathematical approach to probability. Many older service courses in statistics were criticized as containing a probability portion that was too difficult for the general audience. Probability was usually presented mathematically, and the connection between the symbols and the concepts was baffling to many students. But any student who can understand the sense in which a fair-coin toss is a repeatable process with certain long run characteristics, can understand a wide variety of applications of models of randomness. A first course in statistics can inform the student, through simulations, about the role of randomness in sports, lotteries, education, investment, medicine, politics, insurance, and many other common features of everyday life. These applications carry general insight into the influence of randomness, and at the same time inform the student about particular applications that they are bound to encounter in real life.

During the last forty years, the first service course in Statistics has evolved from a purely math applications course to a broader-based statistics course. In the sixties and seventies, as the demand for an understanding of statistics grew among application disciplines, the mathematics was reduced and the courses emphasized the procedures required by standard inferential techniques. As computers became more important in statistical practice, the areas of descriptive data analysis and data collection were expanded in the first course. To accommodate this expansion, the discussion of probability was reduced to a bare minimum. Although there has been some reduction in the emphasis on hand-calculation formulas for inference procedures, the time saved by this has tended to be replaced by the additional inferential procedures: regression, contingency tables, and nonparametric tests.

The result of all this change is that current service courses in statistics teach a long list of techniques for data analysis and inference; However there is some doubt that they successfully convey an understanding of random variation and its relevance to the real world. Lipson (2000) showed that the idea of a sampling distribution was understood by only a small fraction of her subjects, even after an intensive and pedagogically sophisticated course aimed at accomplishing this understanding. Anecdotal evidence from statistics instructors certainly confirms this finding.

My proposal is to offer this first service course in statistics with P-values and confidence intervals omitted. The course would concentrate on informal probability (to be described in the following sections) and descriptive data analysis. The logic of P-values could be included in an informal way in one or two simple scenarios. For example the statement "When something happens that is unusual under ordinary circumstances, then this is evidence that the circumstances are not ordinary" contains the basic logic and does not leave the world of everyday language. No new jargon is needed. Similarly, the idea that a sample mean will vary from the population mean by a factor $1/\sqrt{n}$ of the variability of the sample is possible to convey without the more complex idea of an interval estimate of a population mean. In this sense the jargon and rituals of P-values and confidence intervals can be avoided without omitting the ideas entirely.

One text closely aligned with this approach is "Probabilities in Everyday Life" by McGervey (1986). This book is aimed at a lay audience and omits a lot of material that would considered as basic for an introduction to probability and statistics, such as the central limit theorem, and simulation. But it does include a discussion of lotteries, sports betting, political polls, insurance, investment, and various games involving randomness. The university student should be exposed to this material but may need a bit more theory than the book offers.

Before proceeding with my proposed course outline, I would like to review what I consider to be the "basics" of probability and statistics: the theory that contains the building blocks on which the discipline of probability and statistics is based, and the tools which enable the theory to be used.

2. WHAT ARE THE BASICS?

'Inference' in its broadest meaning is clearly a central component of our discipline. But inference taken to mean parameter estimation and hypothesis testing is too narrow. The high prestige of parametric inference as the central theme of statistics is based on the idea that investigators must arrive at conclusions about parameter values from individual data sets. The reality is that most data-based studies are only suggestive of new knowledge, and anything really interesting must be confirmed in different ways by different people with different data. Data-based research is rarely final. Statistical theorists who concentrate on inferential outcomes of statistical analysis are ignoring the iterative nature of scientific development. More energy needs to be expended on information retrieval from fuzzy data, and less on the requirements of incontestable proof of certain parametric properties in the face of unexplained variation. We need courses which concentrate more on finding 'signals' and less on discounting 'noise'. A more realistic view of the process of inference from data broadens the perspective to include a legitimate stage of hypothesis exploration and conjecture, and these may or may not be expressible in parametric form. The bases of skill development in this area include guided experience with the tools of data exploration and also a deep and intuitive appreciation of the effects of randomness on observable outcomes.

In course revisions over the last few years, it has become fashionable to increase the emphasis on data analysis (especially graphics and descriptive statistics) at the expense of probability and probability models. The argument that graphical displays and descriptive statistics are more useful to the general student than combinatorics formulas and probability calculus is compelling. However this same argument does not apply so clearly to certain informal probability phenomena that can be conveyed without mathematical structures. Variability phenomena associated with sports, lotteries, education, investment, medicine, politics, and insurance, amongst others, can be taught via practical scenarios explored with simulation of very simple models. These applications help to relate both the descriptive tools and the variability phenomena to problems that the student will meet in other courses and in life experiences. In the following I will argue that the "data analysis + informal probability" course is a better choice than the "data analysis + inference" that is commonly provided in service courses. I will also argue that this first course is a better preparation for the serious statistics student than the usual mainstream introductory course based on calculus and formal probability models.

The basic tools needed to understand the impact of variability in the applied areas mentioned are: Distributions, Probability, Sampling, Models, Simulation, Averaging, Variation, and Relationships. These concepts can be demonstrated from a basis of the very simplest probability ideas. Even though only the simplest probability ideas are covered, the applications to which they apply are an important part of a general education. The applications are not merely examples to illustrate the theory, although they do illustrate some of the theory underlying statistical thinking.

I am not underestimating the novelty to many college students of the concept of probability. As Gal and Garfield (1997, p 11) say "[numerous studies] reveal that many students and adults find [probability] topics difficult to learn and understand in both formal and everyday contexts..." However, rather than avoid the subject in introductory statistics courses, my suggestion is to spend more time with probability in its very simplest form: a list of outcomes and their associated probabilities.

The argument will proceed as follows: I will outline the presentation of a series of applied topics (sport, lotteries, etc) using simple probability simulations. For each topic I will suggest the practical importance of familiarity with this particular context, and also suggest, by way of an example, the role of the context in advancing the student's understanding of the underlying conceptual material. After all these examples are described, I will summarize the way in which they do cover the basic statistical concepts. The contexts and examples can vary according to the instructor's preference, as long as the complete list of concepts is covered. The questions to ask of these demonstrations are: 1. Is this accessible and interesting to a general audience? and 2. Is this context of practical importance to a general audience? When all the topics are covered we can ask 3. Are all the basic concepts of statistics conveyed by this series of examples?

The ability to demonstrate the concepts using simulations frees the instructor from much of the complication of a traditional probability-based approach. The novel feature of this approach is that the goal is to convey a lasting intuitive understanding of randomness rather than a familiarity with formal inferenceFormal parametric inference is postponed to future courses. If the student has no future courses, they would not be expected to understand formal inference. However, I am not suggesting replacement of descriptive data analysis, only the formal inferential material. , from the traditional course. Some informal inference will be covered incidentally. , and this is certainly part of understanding randomness.

3. DEMONSTRATIONS OF THE CONSEQUENCES OF RANDOMNESS

Some of the negative feelings directed toward compulsory service courses in statistics are the result of a technique-based presentation, when the students served are more interested in useful information. A technique like the calculation of standard deviation is not useful information until it solves a problem relevant to the student. Fortunately, there are several application contexts which can serve to introduce statistical ideas while being of general importance and interest at the same time. In this section we give a brief outline of several important practical contexts which include

unexplained or uncontrolled variation, that can be simply described, and for which the key to understanding the outcomes is the concept of a probability. The real life utility of the examples helps to maintain student interest. Even students who do not find sports, lotteries etc interesting may appreciate that the examples contain information that will be useful for them in everyday life.
1. Sports:

Ignorance of the effects of variability is especially obvious in the business of sports. Commentators will make a point of noting three wins in row as evidence of a "turn-around" of a basement team. Additional millions of dollars will flow to teams having a successful season, even if there is no valid evidence that the best team is really better than the worst one. Simulations of league outcomes when each contest is "fair" (e.g. P(win) = 0.5) will demonstrate the folly of ignoring chance as an explanation for apparent quality in sports teams. Any student who understands how a computer can simulate tosses of a fair coin will understand this league simulation. One can show that the point spread in a season's actual league results approximates what one would get if each game in the season was determined by the toss of a fair coin. Allowance for the tendency of the masses to ignore chance variability can provide an opportunity for the enterprising gambler - if team quality is an illusion, betting the underdog is always the best strategy! The assumption here is that favorable odds would be given to the gambler betting that the underdog would win. Any favorable odds would therefore benefit the gambler in the long run.

The claim that betting the underdog is the best strategy is based on two important assumptions:
1. That all teams are of equal quality in the sense that each has a 0.5 chance of winning any game.
2. That a bet to win on the team with the poorer record would be given favorable odds.
While I have argued that these assumptions may be reasonable in certain situations, anyone using this stratgey must satisfy himself that these assumptions apply. Caveat Emptor!

Here is an example of how a league simulation might be constructed.

Arrange that a team gets 1 point for a tie and 2 points for a win. A simple simulation of tosses of a fair coin would yield results such as are shown below. In this demonstration, 16 teams are in the league and each team plays each other team twice, once at home and once away. In the following we assume that all teams have the same probability of winning a game and that all teams have a probability of .2 of tying in a given game. (This tying probability is typical of a local soccer league in my area.) In other words, we assume for each team in each game that the number of points won by that team is 0 with probability .4, 1 with probability .2, and 2 with probability .4.

Final standings for three simulated seasons: 16 teams play each other home and away.

team ID	Season Pts	team ID	Season Points	team ID	Season Pts
5	38	4	38	13	39
9	38	7	37	15	37
10	38	14	37	4	35
14	33	1	36	5	34
1	31	3	36	8	32
16	31	16	35	10	32
2	30	2	32	14	32

6	30	8	31	6	30
13	30	9	31	9	29
3	29	5	28	11	27
15	28	10	26	12	27
4	27	11	26	1	26
12	26	13	26	2	25
7	24	6	21	3	25
11	24	15	21	7	25
8	23	12	19	16	25

The lesson in simulations like these is that a large variation in points for a team can be the result of random variation, and that the apparent quality difference, of the magnitude suggested by the above outcomes, is an illusion. Would results such as those in the second simulated season noted above induce a gambler to give long odds in favor of team #4 against team #12, in a subsequent game? If so, then the opposing gambler would have a favorable bet, if the simulation assumption of equal quality teams were correct. Simulations will show that the median ratio of highest team points to lowest team points is about 2, and ratios of 3 or 4 are not uncommon, in this situation in which all teams have equal chance to win. In other words, when the variation of points across the league is similar to the variation shown in these simulations, we have no evidence of a differential in team quality. On the other hand, suppose just one team (# 16 say) were superior, with a probability of winning against the other teams of .5 instead of .4, (so W/T/L probabilities are .5/.2/.3) then typical league standings would look like this:

team ID	Season Pts	team ID	Season Points	team ID	Season Pts
11	40	14	45	11	38
10	37	9	35	**16**	38
16	35	2	33	2	37
3	34	5	32	6	37
1	33	1	30	1	34
4	32	4	30	7	34
9	32	6	30	4	33
12	31	13	30	5	32
13	30	3	29	9	30
5	29	**16**	29	12	28
15	28	7	28	8	27
2	26	10	28	10	26
14	26	15	28	3	23
8	25	12	26	13	22
6	24	8	24	14	21
7	18	11	23	15	20

Note that the superior team # 16 typically does not emerge as the season winner. Although team #16 plays 30 games with a probability of winning of 0.5 and tying with a probability of 0.2, this margin of superiority is not enough to be revealed in a one-round league. It turns out that a team as superior as team #16 would win a one-round season (each team plays each other team at home and away) with probability of only .32. If the superiority of team 16 over all the others were such that it would win a game with probability 0.6, so the W/T/L probabilities were 0.6/0.2/0.2, the result would be more clear cut: in this case the probability that the best team wins the league title is .76.

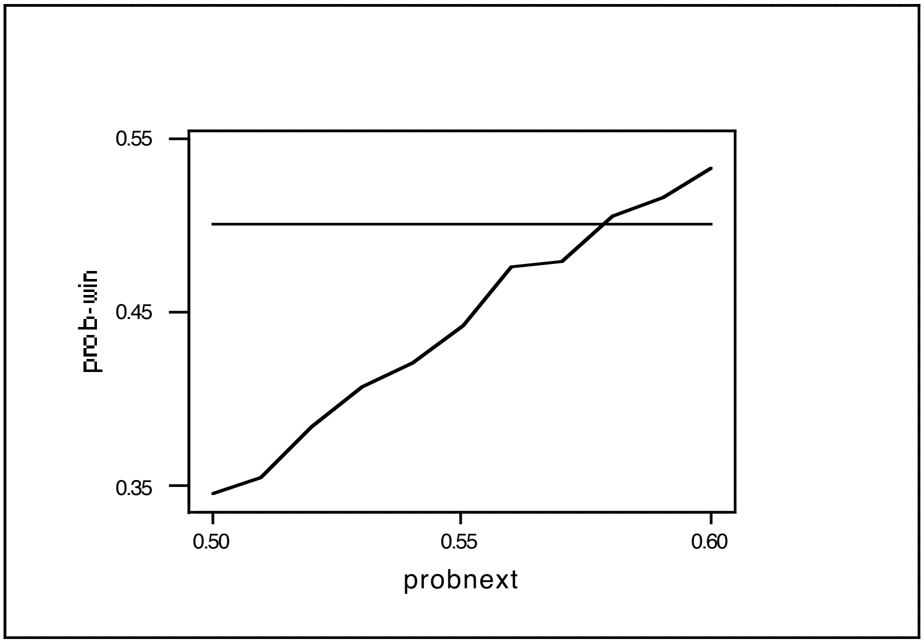
Since a superiority of this degree is extremely unlikely in a league, the overall conclusion from this is that the best team will not usually earn the most points in a season, and will usually not win the league title!

Another interesting mental exercise is to consider the influence on the variability of points across the league, of a home-team advantage. Considering what would happen if there were a very strong home team advantage, that the home team would almost always win, leads one to the correct inference that a home team advantage would decrease the variability of the team scores. Simulation shows the size of this effect to be rather small for realistic home-team advantages.

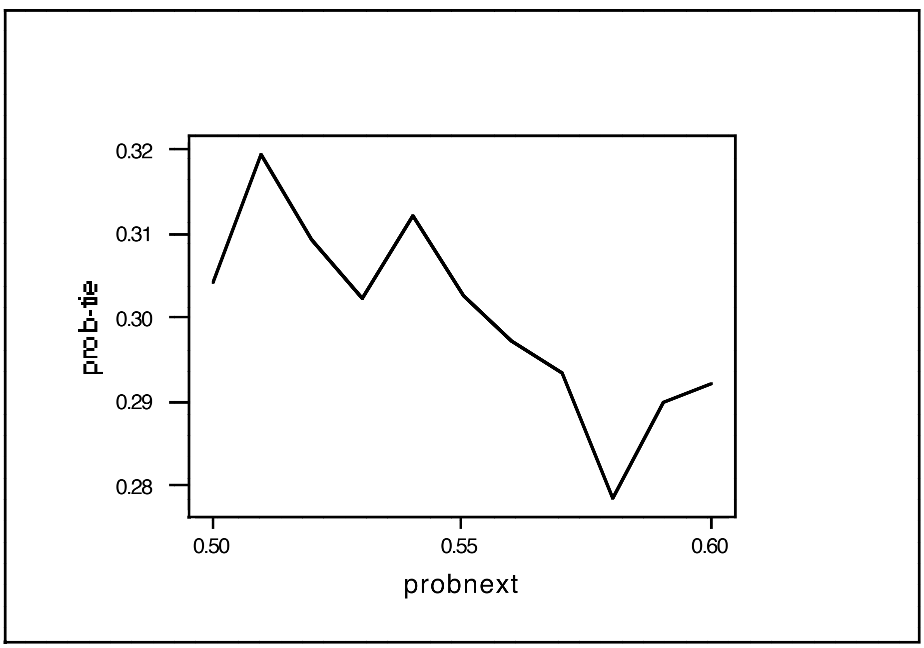
The strength of this particular example is that it produces information of some practical use, it is information that is not generally known, and it has applicability far beyond the particular situation described. The cumulative effect of many "random" influences does not necessarily "average out in the long run". The lesson of the true meaning of "the law of averages" is clearly one of the most important. The discussion by Freedman, Pisani, and Purves (1998, Chapter 16) is particularly clear, and might be used as an introduction to this concept, to be followed by this league-standings example.

Another aspect of this team-quality context is the relationship between the quality of a team and its chance of winning a particular game. Consider ice hockey: it has the simplifying feature that each goal re-starts the play the same way. Suppose there are only two teams being compared. We can represent team quality by the probability that the team gets the next goal, given that another goal is scored before the game ends. Let p be the quality of team A, so 1-p is the quality of team B. For a more advanced class one could let the number of goals in a game have a Poisson distribution but for the simplest demonstration, it could be assumed that 6 goals are scored. The only issue is which team scores the goals.

The simulation of the 6 biased-coin tosses (i.e. Bernoulli trials) is easy to describe and to simulate. An interesting question is, what is the relationship between p and the probability of winning the game? A simulation would produce graphs such as are shown below. Of course, the binomial calculation can also be done, but this requires more probability theory than is necessary for the rest of the course. Moreover, the simulated outcome is more informative from the pedagogic point of view, as will be seen. The following graph shows how the probability of winning the game relates to the probability of scoring the next goal, in games with a total of six goals scored.



The instructor can point out that, in order to win the game with a better than even chance, one has to get the next goal about 58% of the time. Of course, ties will be fairly common when there are a total of six goals: The simulation used 5000 games for each of p=.50, .51, .52, …, .60 to produce these graphs.



Note that the probability of a tie appears at first to be less precisely estimated than the probability of a win, but that this is an illusion. By paying attention to the vertical scales, one can see that the reverse is true (which of course is what theory confirms). An important point in comparing computer-generated graphs is that the default scaling always fills the screen, and the scaling is chosen to produce this - consequently one has to either control the scaling or take careful note of the scaling. These observations contain important ideas for students to consider. Moreover, they would be missed in a non-simulation demonstration of the phenomenon.

Once the fine scaling is noticed, a conclusion might be that the probability of a tie is not very sensitive to the team quality departure of "p" from 1/2, although as anticipated it does decline as p moves away from 1/2. When p increases by 20 percent, the probability of a tie only decreases by about 10 percent.

2. Lotteries:

It is a rare individual that has not been tempted by advertising for public lotteries. But how many really understand their chances of winning, or their long-run prospects in return to loyal participation in a public lottery? Again, a simple example using an artificial lottery with perhaps only one prize can make the relevant points. How many tickets in how many lotteries would one have to buy to have a ten percent chance of a major prize in one's lifetime? A simple calculation (or simulation) shows that if one buys lottery tickets every week and every year for 40 years, and one wishes a 10% chance of winning a major prize, then the probability of a win in one particular week must be at least .00005. This would require about 50 tickets in a lottery in which the probability of a major prize were $^{1}/_{1,000,000}$, which is typical. One would therefore have a 90 % chance of spending a lifetime buying 50 tickets each week, and ending up never having won a major prize. On the other hand, this loyalty would have a 10% chance of being rewarded by at least one major prize. Of course the cost of the 100,000 tickets over the period represents a significant risk!

Other aspects of lotteries that one can simulate include the strategy of quitting as soon as a moderate prize is won, or doubling the purchase amount each year of no prize, and so on. Another important detail is that only a portion of the income from tickets is returned in prizes, and that the very-long-run return on investment in lottery tickets is not attractive.

There is a practical aspect to this example: lottery tickets should not be viewed as an ordinary investment, where one has a reasonable expectation of a return on investment. One is buying hope more than a serious expectation of gain. This understanding achieved will prepare students to make wise decisions about lottery participation that fit their financial situation.

The example also helps to reinforce the idea of randomness, and the fact that those "theoretical" expectations may never be realized, and even when they are realized, the result may be surprising.

3. Education:

Although there are many contexts one could use here, an area close to the student's heart is the variability of marks. Consider a student assigned a sequence of letter grades for various components of a course. Assume that each letter grade is based on a numerical grade in rough accordance to a percentile guideline, and that the final course letter grade is based on a weighted combination of the component numerical grades. More specifically, suppose the letter grades are assigned in each component, and in the aggregate grade, according to the following percentile guidelines: top 20% - A, next 30% - B, next 40% -C , and the last 10% - D or F. Further, suppose there are ten equally-weighted components and that the aggregate score is just the simple average of the numerical component scores. Since the aggregate score will have a smaller standard deviation than the component scores, a numerical grade that is near the top of the B range for the component scores will certainly be in the A range if it were also obtained for the average score. In other words, a student getting B in every component of the course should not be surprised to get an A overall on the course. Likewise, a student who gets C on every component could end up failing! The assumption underlying these claims is that there is at least some degree of independence between component marks. Of course the effect is lessened by the correlation among scores of an individual student, and the crucial idea of dependent observations can be introduced here. Simulations can easily demonstrate this concentration phenomenon due to averaging. A simulation in which the ten component scores have a 0.5 correlation with each other will yield a contraction of the SD for the aggregate by a factor of about .75 instead of $1/\sqrt{10} = .31$ that would have occurred if the scores were independent. This simulation can be easily set up with iid normals if one of the normals is added to each of the others. The added normal represents the "quality" of the student performance that does not vary from one component to the next.

A preliminary to the important idea of the sampling distribution is an understanding of the relative stability of averages. This grade accumulation phenomenon should help to arouse interest in this concept. However, the perspective of letter grades as the result of an academic competition is realistic, and the role of consistency in this competition deserves emphasis. If a student begins to think of statistical concepts as ideas that relate to their general academic achievement, they may be motivated to absorb these concepts.

4. Investment:

How can one explain the widespread ignorance about financial risk in investment? Investment advisors use standard deviation to measure risk - even though they know that money put in a sock has zero standard deviation but will almost certainly become worthless over time, a risky strategy indeed! The distinction between the risk of financial loss and the variability of the value of an investment is a simple idea. The important area of investment for retirement leads one to consider very-long-term returns to investment, and short-term variability has little to do with it. But those who do not understand what standard deviation measures will end up believing a savings account has less "risk" than a diversified portfolio of equities. This is a case of less variability being misinterpreted as less risk.

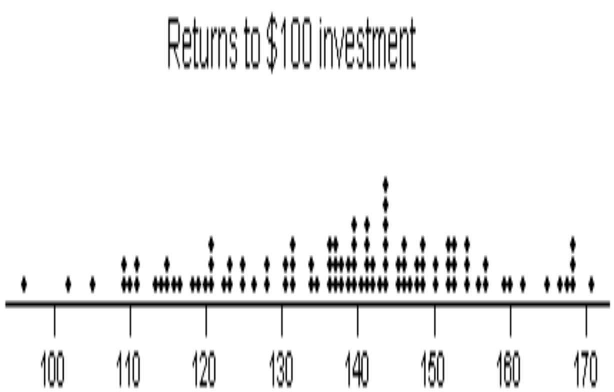
There is another message to make in connection with investments - an important consequence of the relative stability of averages, the $\frac{\sigma}{\sqrt{n}}$ formula. A portfolio of a large number of very risky investments can produce a very stable and high long term gain. The demonstration of the central limit theorem need not involve mathematics - simulation models will tell the story adequately. Again, the simple notion of a discrete probability model is all that is needed here.

Suppose I hold 1 shares in each of 100 companies, and suppose all the companies have shares priced at $1.00, so my total investment is $100. Further suppose that each company has the following prospects over the following year. The amount under "prospect" below is the possible share market value at the end of the year.

Prospect	Probability
$0.00	0.25
$0.50	0.25
$1.00	0.25
$4.00	0.25

A company with these prospects is certainly a "risky" investment. Each of these companies has a 75% of doing worse than one could obtain by simply putting the money in a savings account in the bank. But there is a 25% chance that the company will make a good profit. The average return for each company's share is 0.25 x 0 + 0.25 x .50 + 0.25 x 1.00 + 0.25 x 4.00 or in other words $1.38, which would be a 38% return, a very good return over one year. Now the wonderful thing about a diversified portfolio of high-risk stocks can be seen: The 100 companies will mostly do very badly, but will do well on average. The service of mutual funds is to give the small investor access to this kind of portfolio.

In fact, if this whole portfolio is simulated 100 times, the year-end return would look like it did in the particular simulations summarized below: the return varies from $96 to $171 and the average return is $138. This portfolio has both low risk and an excellent return! Each dot in the graph below records the outcome of one year's investment in 100 of these risky companies.



Typical returns are between 15% and 55%, enough to satisfy the most optimistic investors, let alone the conservative ones. No savings account, term deposit, or bond will yield this high a return, even though their safety is no greater than this portfolio of risky investments.

We must admit at this point one crucial assumption of the above simulation: that the companies' successes or failures are independent of one another. But we can strive for this in selecting a diverse portfolio of company stocks. To the extent that our companies have similar good or bad luck, the variability of the portfolio will increase, and the outcome would be less predictable. However the overall effect still obtains. A secondary lesson for students here is that the idea of independence is important for understanding portfolio diversification.

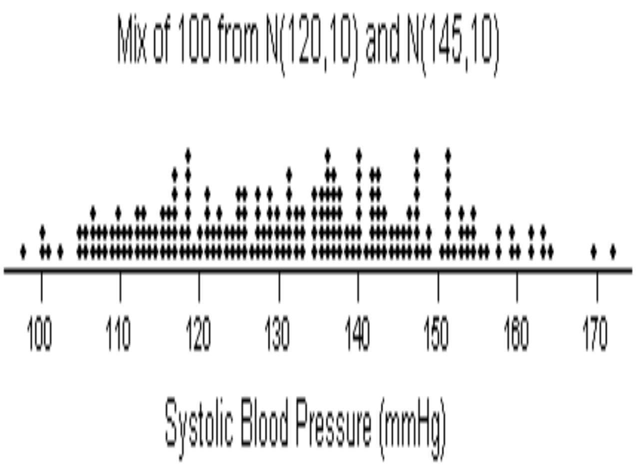
Investment notions are thought of as sophisticated not because they really are, but because the powerful concept of probability is not widely understood. The lesson demonstrated in the above example is of great practical importance to anyone with investment concerns. It may be argued that students are not usually investors. However, many students will have influence and interest in family wealth, some will be directed to a career in the finance industry, and others will win the lottery or be e.com millionaires. All but the most pessimistic students expect they will have money to invest!

The investment paradigm is another example which brings into a familiar context the ideas of the sampling distribution. This should help to break down the wall between rote knowledge of probability phenomena and the needs for understanding of variability in everyday life.

5. Medicine:

One reason doctors are so reluctant to share with a patient the patient's biochemical data is the lack of appreciation among the public of variability in those biochemical tests. A transient high blood pressure or cholesterol level, once revealed to the patient, may cause harmful behaviors on the part of the patient. An understanding of standard deviation is not quite enough here - one also needs to know about the sampling framework of the measurement model. In other words, one needs to know the difference between within-individual variation and between-individual variation. But this is not beyond the scope of a first course. Again, a simulation model can demonstrate the effect.

Biochemical tests tend to be quantitative measurements, so a great convenience here would be to use the normal distribution for simulations. If one wanted to avoid the complications of explaining a continuous density function, on could demonstrate the empirical nature of the $N(\mu,\sigma)$ generator by drawing histograms of its output. Having done this one could simulate the systolic blood pressure measurement for two hypothetical populations from N(120,10) and N(145,10). The N(120,10) population would be the medical "normals", while the N(145,10) population could represent the moderate hypertensives requiring medication. A clinical population for a specialist might contain roughly equal numbers from the two populations. Combined random samples of 100 from each one can be simulated and will typically look like this:



A common medical problem is to decide on the basis of one of these measurements whether the patient has a high blood pressure problem or not. By showing the populations separately and together, one can make a convincing argument that identification of the group from the measurement will be difficult over a wide range of individual measurements. Another observation is that the bimodal nature of the combined distribution is not that clear, and this could also lead to a simulation exercise. To be convincing, it would be possible to have students measure their own blood pressure at the machines common in many pharmacies, at intervals of a week or so. This would enable the standard deviation to be estimated realistically both for within-individual and between-individual. Naturally the data would be collected anonymously.

The practical lesson here is that medical measurements are subject to measurement error, intra-personal and inter-personal variation, and this complication should explain why individual measurements are inconclusive. One theoretical lesson is that sources of variation must be considered before measurements can be properly interpreted. Another is that overlapping distributions make classification problematical. These are lessons worth learning.

6. Politics:

A few years ago, the word statistics would suggest "batting averages" or "yards gained" to the North American public. These days the mention of statistics provokes comments about the "19 times out of 20" so often quoted in the results of opinion polls. Variability of sample proportions is something that can be simply explained using simulation or even urns with colored beads. Questions like: "A political opinion poll based on a random sample of 400 individuals estimates a candidate will obtain 60 percent of the vote, how likely is it that the candidate will receive a majority of votes on election day, assuming the population sampled in the poll is the same as the population of voters?" should prompt the right kind of thinking. By simulating the distribution of sample proportions when the population proportion is only .5 (on the edge of majority), one can show that a sample proportion of .6 is not very common, and by either simulation or logic deduce that for population proportions less than .5, a sample proportion of .6 would be even less likely. A reasonable conclusion in this case is that the election majority for the candidate would very likely follow. One could similarly show that a sample proportion of only .53 , based on this same sample size of n=400, would not be so comforting to the candidate wanting a majority. Again, a simple understanding of probability and sampling is all that is needed for this application. And an understanding of the power and limitations of opinion polls is surely useful for anyone who votes. They will understand that the opinions of a small sample can give valuable information about a large population, but that sample information is not perfect and a slim majority in a poll may not reflect a majority in the population sampled. It is common to ban poll-taking near election day so as to prevent undue influence on voters. This suggests that understanding polls may actually affect voter behavior is some instances.

Service courses in statistics often use the survey context as a natural model for the concept of random sampling -- it is the concrete nature of this sampling scheme that makes it a natural introduction. If a stratified sample is considered, for example from graduates and undergraduates at a university, the difference between simple random sampling and "fair" sampling can be illustrated. By "fair" sampling is meant that each population member has an equal chance of being chosen into the sample. But if the graduate:undergraduate ratio were 1:4 say, the stratified sample with proportional subsamples in the same ratio would be "fair" without being "simple random". From this example students can get the idea that unbiasedness is achieved by simple random sampling, but that prior knowledge of populations can lead to improved random sampling schemes. The growing business of opinion polling is one that students should be aware of, and have some understanding of why they work, and why their methodologies are not trivial. So this is another example with current practical importance as well as being a good introduction of basic statistical theory.

7. Insurance:

Every household needs at least one person who understands the basics of insurance. Both life insurance and casualty insurance can involve a very large amount of money over a long period, and in order for a buyer to get value from this expenditure, they must understand how the various products relate to their personal needs. Young students may be more interested in casualty insurance than life insurance, but the same principles of distribution of risk apply in both areas.

The idea of distribution of risk can be portrayed with a simple example. Consideration of a cohort of 20 year-old car-owners, the annual auto insurance premiums they pay, and the frequency of large pay-outs by the insurance company, during the year, should be a process students can relate to. The unlikely event that premiums will exactly equal financial needs of the insurance company can be demonstrated by simulation. The idea of business risk would be demonstrated as a bonus. All this can be explained using simple Bernoulli sampling experiments simulated by computer, and without reference to any formulas at all.

In listing these applications, I am attempting to show that some important, real-world phenomena of randomness can be demonstrated without the complications of formal probability, statistical inference or the analysis of time series. Furthermore, the student does not have to be fascinated with the whole subject of statistics to learn enough to deal intelligently with randomness in real life. Insurance, investment, sport, etc. are things that everyone *must* deal with, even if they never get close to formal inference in their careers.

I am not claiming that these examples are the ones that must be used. I hope that they illustrate the kind of examples which will be interesting to students, have practical value for their lives, and also convey an appreciation of the surprising consequences of randomness, and its power in explaining phenomena which the naïve student would assume was inexplicable. One strength of the approach taken here is that with simulations the instructor and the student will know exactly the model underlying the data - a disadvantage of real data demonstrations is that one is never sure of the underlying model, and there will always be a lot of argument about the various factors that might have influenced the response. Simulations present a simpler task, and put more emphasis on the underlying randomness concepts.

4. IS BASIC THEORY COVERED?

Traditional instructors may worry that such a problem-based course relying on exploration by simulation would fail to convey the logical structure of the basic theory of statistics. However, once the ideas contained in the applications are understood, a review of the basics in a logical progression could be quickly covered.

Populations, sampling, probability, distributions, models, averaging, variability, simple graphical displays, and simulation are basic concepts in univariate statistics. Except for the concepts of parametric inference, experimental design, and bivariate statistics (including simple regression and correlation), these are all amply demonstrated in the particular examples chosen. Here is a table showing the strongest correspondences with the examples described:

Application	Populations & sampling	Probability	Distributions	Models	Averaging	Variability	Graphics	Simulation
Sports	√	√	√	√		√	√	√
Lotteries	√	√				√		√
Education	√		√		√	√	√	√
Investment		√	√	√	√	√	√	√
Medicine	√		√			√		
Politics	√		√			√		√
Insurance	√	√	√	√	√	√		√

Consider now what is omitted: parametric inference, experimental design, and bivariate statistics. Some of this may be included in the course as "data analysis", but without further probability development. Regression and Correlation can be considered as descriptive measures, a way of describing bivariate data. (See Weldon (2000)). Experimental design can be touched on in an exploratory way, without getting into analysis of variance: for example, a comparison of the hang time of certain models of paper planes can illustrate the idea of comparison of distributions, blocking (by thrower, for example), and randomization (of order of testing, for example). Parametric inference is the only really important omission here. Many departments will require it in their service courses, responding to the demand of the served departments. However, the counter-arguments need to be made: students do not understand parametric inference based on a single course; they rarely make serious use it in the lower level courses anyway since they are neither producing nor evaluating data-based research at this stage; they will only get an appreciation of the implications of randomness through applications that interest them; and they need the information in the non-inference course described here for real life. Less formal inference can make the first course in Statistics more useful, more lasting, and provide better motivation for additional courses that rely on an understanding of statistical strategies.

5. RESOURCES FOR THE PROPOSED COURSE:

A good understanding of probability, experience with data analysis, and a general knowledge of the application areas mentioned are all that are really needed by the instructor to teach this course. Of course, familiarity with some statistical software providing easy access to simulation routines is also essential. The author has constructed MINITAB macros for many applications, but other popular programs such as XLStats (Carr(2000)), DDXL (Data Description Inc(2000)) or JMP (SAS Institute(2000)) can also be used. Some books that provide similar ideas are the ones listed below by McGervey (1986), Tanur(1989), Orkin (1991), and Schaeffer (1996), although no one of these would serve as a text for this course.

6. DISCUSSION AND CONCLUSION

The basic ideas of probability and probability models have a widespread applicability to everyday living. In recent years, these aspects have been excised from the introductory statistics curriculum in order to provide more emphasis on data analysis, and a more thorough explanation of methods of inference. But the utility of the formal inference to the beginning university student may be rather slight, compared to the usefulness of the random phenomena discussed here. More people have to understand the role of variability in investment, insurance, and sports, than have to understand P-values and confidence intervals for assessment of research.

Furthermore, students actually requiring inference skills for their later studies will find these techniques much more comprehensible once a good intuitive grasp of probability has been attained. Students in conventional courses often find the probability material the most difficult to grasp.

Courses in probability per se have existed for a long time. But these courses are seldom considered service courses, and are likely to include rather heavy mathematical symbolism and sophistication. Even the students who are comfortable with mathematics are unlikely to really understand the implications of the probability formulas to the real world, unless time is taken to detail these connections. Such students rely on the symbolism and there is a danger of ignoring the underlying concept if it is not needed for the course itself. Students with a good mathematics background are no less in need of a course like the one proposed here than would be the typical "service course" enrollee. However, it is not clear that the practical motivations would be as effective for the mathematics student as for the typical service course student.

There is certainly a practical problem of implementation of the proposed course. There is no question that the prevailing wisdom is to include some formal inference in the first course in statistics, and many undergraduate programs have these built in as compulsory components. The shift to include more probabilistic phenomena in the first course and less inference would have to occur gradually. My aim has been to suggest a goal, and the implementation of this shift would depend on local interest and context.

In conclusion, my suggestion is to have this random phenomena-based course, along with some descriptive data analysis, as the first course and the formal inference courses as a follow-up for those who need or want them.

7. REFERENCES

1. Cleveland, W.S. (1993) _Visualizing Data_: Summit, N.J.: Hobart Press.

2. Cobb, G.W. (1993) Reconsidering Statistics Education: A National Science Foundation Conference. _Journal Of Statistics Education_ volume 1, number 1 (July 1993)
See http://www.amstat.org/publications/jse/ .

3. Gal, I. and Garfield, J. (1997) _The Assessment Challenge in Statistics Education_. Burke, Va. USA. IOS Press.

4. Griffiths, D., Stirling, W.D. and Weldon, K.L. (1998) _Understanding Data: The Principles and Practice of Statistics_, New York, N.Y. John Wiley & Sons.

5. Freedman, D., Pisani, R. and Purves, R. (1998) _Statistics_ Third Edition, Norton, NY.

6. Hogg, R.V. (1994) A Core in Statistics for Engineering Students. _American Statistician_48:285-287.

7. Lawrance, A.J. (1996) A Design of Experiments Workshop as an Introduction to Statistics._ American Statistician_ 50:156-158.

8. Ledolter, J. (1995) Projects in Introductory Statistics Courses.(1995) _American Statistician_ 49: 364-367.

9. McGervey, John D. (1986) _Probabilities in Everyday Life_ New York. Ivy Books

10. Moore, D.S. and McCabe, G.P (1993) Introduction to the Practice of Statistic*s*_, New York. Freeman.

11. Orkin, M. (1991) _Can You Win?: The Real Odds for Casino Gambling, Sports Betting, and Lotteries_. New York. W.H. Freeman

12. Singer, J.D. and Willett, J.B.(1990) Improving the teaching of statistics: putting the data back into data analysis. _American Statistician_ 44:223-230.

13. Snell, L. (2000) The Chance Course. This continuously updated resource may be viewed via the www URL: http://www.dartmouth.edu/~chance/index.html

14. Tanur, J.M. et al (eds) (1989) _Statistics: A Guide to the Unknown_ 3rd Edition. Wadsworth.

15. Weldon, K.L. (1986) Statistics: _A conceptual approach._ Englewood Cliffs, NJ Prentice-Hall.

16. Weldon, K.L. (2000) A simplified introduction to correlation and regression. Journal of Statistical Education: Vol 8, No 3 (November). See http://www.amstat.org/publications/jse/ .

17. Carr, R. (2000). XLStatistics 5.28. XLent Works, Australia. The URL is http://www.man.deakin.edu.au/rodneyc/xlstats.htm

18. JMP - SAS Institute Inc. - www.JMPdiscovery.com

19. DDXL - Data Description Inc. - www.datadesk.com

20. Lipson, K. (2000) PhD thesis at Swinburne University, Melbourne. "Determining the relationship between the concept of sampling distribution and the development of understanding of inferential statistics".