BAYESIAN APPROACHES TO TEACHING ENGINEERING STATISTICS

J. Stuart Hunter
Princeton, N.J., U.S.A.

The reduction of uncertainty enhances knowledge.
Probability is a measure of uncertainty
Statistics provides the logic and art for combining the two.

Most students know that probability is a quantitative measure of uncertainty, a number lying between zero and one. In repeated trials they recognize that the number of "successes" divided by the total number of trials estimates the probability of a success on the next trial. Further, most are willing to estimate the probability that Mr. G will win the next election, although repeated trials are not here appropriate. And many will be willing to guess the weight of the teacher by providing a set of limits within which the student is confident that the true answer will lie.

Less common is the knowledge that other metrics exist for measuring uncertainty. For example, odds:

$$\text{Odds} = (\text{Probability})/(1 - \text{Probability})$$

and log-odds. Thus, if the probability a horse will win a race is 0.05, the odds it will win the race are "one to nineteen" or simply 0.0526, and the $\log_{10}$ - odds equals -1.2788. We note that:

$$0 \leq \text{Probability} \leq 1,$$

$$0 \leq \text{Odds} \leq \infty ,$$

$$-\infty \leq \text{log-odds} \leq +\infty$$

All are measures of uncertainty.

Log-odds can be particularly useful. Many students will readily equate "impossible" and "certain" to the log-odds equivalents of $\pm \infty$. Log-odds have been called "evidence": if the log-odds are to the base ten one has so many "bels" of evidence in favour of an hypothesis, to the base e so many "napiers", and to the base two so many "bits". Log-odds are particularly useful in illustrating Bayesian methods, that is, for demonstrating how a rational person reduces or increases his personal uncertainty (belief) in an hypothesis. Postponing the theory, the concept is best illustrated by an example.

Consider two urns: Urn A consisting of 70% green items and 30% blue, while Urn B has 50% of each. The student is presented with one of these urns. What is the probability it is Urn A? In the absence of any information the student, who may not peek into the urn, will usually state that the probability it is Urn A is: Pr(A) = 0.5. The Odds it is Urn A are then:

Odds(A)=1.0 and the evidence it is urn A is Ev(A)=log-odds=log(1.0)=0. At this stage of the student's knowledge, there is zero evidence that it is Urn A.

To learn something about the urn the teacher withdraws one item. It is green. What is the probability it is Urn A now? Most would agree that since it is easier to get a green from Urn A then from Urn B the probability should now be somewhat larger than 0.5 in favor of Urn A. This probability can be exactly computed using Bayes' postulate. (Further details of the Postulate are given in the Appendix.) Symbolically we have: (the vertical slash stands for the word "given"):

$$Pr(A|g) = [Pr_o(A)\ Pr(g|A)]/[Pr_o(A)\ Pr(g|A) + Pr_o(B)\ Pr(g|B)]. \quad (1)$$

$$Pr(A \text{ given Green}) = [(0.5)(0.7)]/[(0.5)(0.7) + (0.5)(0.5)] = 0.5833$$

The Odds are: $Odds(A|g) = Pr(A|g)/[1 - Pr(A|g)] = 1.40$.

The evidence is: $Ev(A|G) = \log_{10}(1.40) = 0.1461$ bels.

A second item is now drawn. It is also green. What is the probability of Urn A now? (We assume an infinite number of items in each urn. If not infinite, then our computations will have to be slightly adjusted. The point here is to demonstrate the learning process in as simply a context as practicable.) Substituting in (1) gives:

Newest $Pr(A|g) = [(.5833)(0.7)]/[(0.5833)(0.7)+(0.4167)(05)] = 0.6621$.

The new odds are (0.6621)/(0.2279) = 1.9597, and the new log-odds equal 0.2922.

It is important now to note that the second green added 0.1461 bels of evidence. In fact, each successive green adds 0.1461 bels in favor of Urn A. In dealing with uncertainty, <u>evidence adds</u>. A little arithmetic will show that a single blue item will contribute -0.2218 bels of evidence in favor of A since:

$$Pr(A|b) = Pr_o(A)\ Pr(b|A)/[Pr_o(A)\ Pr(b|A) + Pr_o(B)\ Pr(b|B)].$$

An image of Dame Justice emerges, blindfolded and holding aloft a two-pan balance. Each green item, placed in one pan, adding evidence for A while each blue item, placed in the second pan, subtracting (adding negative evidence on behalf of A. The pans tilt and eventually one may learn enough to take a decision.

A practical application in industry concerns the problem of acceptance sampling. A large batch of items appears at the receiving wharf and an engineer is asked to determine whether the batch is acceptable. How does one proceed?

The process manufacturing the batches of items is considered to have two states: A and B. Let Process A be the OK process identified as producing matches with only θ = 1% bad items and let Process B be the undesirable process with θ = 3% bad items. Let the initial probability the items were produced by Process A be $Pr_0(A)$ = 0.5, that is, let there be zero prior evidence as to whether the engineer is dealing with batches from A or B.

The engineer now takes an item from the batch and discovers it is good. How much evidence has been generated on behalf of the hypothesis that the unknown process is actually process A? Evoking Bayes Postulate we get:

$$Pr(A|g)) = (0.5)(.99)/[(0.5)(0.99) + (0.5)\ (0.97)] = 0.50510$$

or equivalently, EV(Ag) = 0.00886 bels. If the item is bad then:

$$Pr(A|b) = (0.5)(0.01)/[(0.5)(0.01) + (0.5)(0.03)] = 0.25000$$

or $Ev(A|b)$ = -0.47712 bels. With each successive item the engineer learns more about the process, the evidence in favor of A grows or declines in a simple additive fashion.

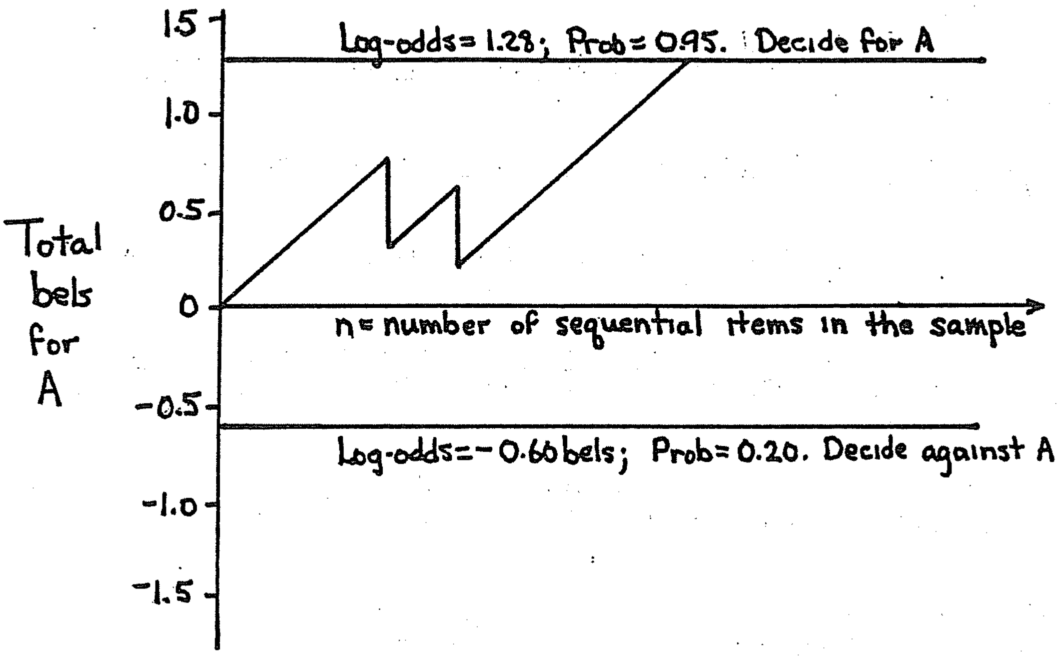
The engineer now adopts the following two rules.

1) If the probability the process is A ever gets to $Pr(A) > 0.95$ he will decide it is A. He thus requires 1.27875 bels for A, and since evidence adds, he will require at least (1.27875)/(0.00886) = 145 good items in succession.

2) If the probability it is process A ever falls to $Pr(A) < 0.20$ he will decide it is not a batch manufactured by process A and reject the batch. He thus needs -0.60206 bels to reject the batch.

Note, in this example it will take more than one bad item to reject.

The engineer may now construct the graphical sequential acceptance sampling scheme as illustrated in the following figure. As each good item is identified he adds 0.00886 bels to his total evidence and subtracts 0.47712 for every bad item. The accumulation of evidence is displayed in the Figure.



Appendix

In the metric of probability Bayes Postulate is:

$$\left\{\begin{matrix}\text{New Probability}\\ \text{of H given e}\end{matrix}\right\} = \frac{\left\{\begin{matrix}\text{Initial Probability}\\ \text{of H}\end{matrix}\right\} \times \left\{\begin{matrix}\text{Probability of e}\\ \text{given H is true}\end{matrix}\right\}}{\{\text{Probability of e}\}}$$

where H stands for the hypothesis in question and e for some observed event. In more terse notation we have:

$$Pr(H|e) = Pr_o(h)\ Pr(e|H)/[Pr_o(h)\ Pr(e|h) + Pr_o(h)\ Pr(e|h)]$$

where the vertical slash | stands for the word "given", where $Pr_o(H)$ is the probability on the hypothesis before the event e, and where $Pr_o(h)$ = $[1 - Pr_o(H)]$.

In the metric of log-odds Bayes Postulate becomes:

$$\left\{\begin{matrix}\text{New log-odds}\\ \text{on H given e}\end{matrix}\right\} = \left\{\begin{matrix}\text{Old log-odds}\\ \text{on H}\end{matrix}\right\} + \log\left\{\frac{Pr(e|H)}{Pr(e|h)}\right\}$$

or simply:

$$\left\{\begin{matrix}\text{Total evidence}\\ \text{on H given e}\end{matrix}\right\} = \left\{\begin{matrix}\text{Original evidence}\\ \text{on H}\end{matrix}\right\} + \left\{\begin{matrix}\text{Evidence provided}\\ \text{by the event e}\end{matrix}\right\}$$

Bibliography

Box, G.E.P., & Tiao, G. (1973). Bayesian inference in statistical analysis. Addison-Wesley. The grand tour.

DeGroot, M.H. (1970). Optimal statistical decisions. McGraw Hill. The most read of all.

Guttman, I., Wilks, S.S., & Hunter, J.S. (1982). Introductory engineering statistics. John Wiley. Chapter 9a is special.

Lindley, D.V. (1965). Introduction to probability and statistics from a Bayesian viewpoint. Cambridge University Press. The masterwork and everyone's reference.

Pratt, J.W., Raiffa, A., & Schlaifer, R. (1965). Introduction to statistical decision theory. Cambridge: Harvard University Press. The theory well exposited.

Schlaifer, R. (1959). Probability and statistics for business decisions. McGraw Hill. Read especially Chapters 39 and 40.

Schmitt, S.S. (1969). Measuring uncertainty. Addison-Wesley. A beginner's text, entirely readable.

Tribus, M. (1969). Rational descriptions, decisions and designs. Pergamon Press. An engineer evokes Bayes.

Zellner, A. (1972). An introduction to Bayesian inference in econometrics. John Wiley. Interesting to non-econometricians too.