Prof. David Draper
Department of Applied Mathematics and Statistics
University of California, Santa Cruz

AMS 131: Quiz 4

Name:
Jeremy Lafond

- *(Fact 1)* As a broad generalization (which you can verify empirically), statisticians tend to have shy personalities more often than economists do — let's quantify this observation by assuming that 80% of statisticians are shy but the corresponding percentage among economists is only 15%.
- *(Fact 2)* Conferences on the topic of *econometrics* are almost exclusively attended by economists and statisticians, with the majority of participants being economists — let's quantify this fact by assuming that 90% of the attendees are economists (and the rest statisticians).

Suppose that you (a physicist, say) go to an econometrics conference — you strike up a conversation with the first person you (haphazardly) meet, and find that this person is shy. The point of this problem is to show that the (conditional) probability p that you're talking to a statistician is only about 37%, which most people find surprisingly low, and to understand why this is the right answer. Let St = (person is statistician), E = (person is economist), and Sh = (person is shy).

(a) Using the St, E and Sh notation, express the three numbers (80%, 15%, 90%) above, and the probability we're solving for, in unconditional and conditional probability terms.

$Pr(Sh|St) = 0.80$ 80% of statisticians are shy

$Pr(Sh|E) = 0.15$ 15% of economists are shy

$Pr(E) = 0.90$ 90% of those attending are economists

$Pr(St) = 0.10$ 10% of those attending are statisticians

$Pr(Sh) = 0.15 * 0.90 + 0.80 * 0.10 = 0.135 + 0.08 = 0.215$

or 21.5% of attendees are shy.

$Pr(St|Sh) = \frac{Pr(Sh|St)*Pr(St)}{Pr(Sh)} = \frac{0.8*0.1}{0.215} \approx 0.372$ or

If you're talking to a shy person, there's a 37.2% chance they're a statistician.

(b) Briefly explain why calculating the desired probability is a good job for Bayes's Theorem.

Bayes's Theorem is effective in calculating the desired probability because it takes into account that some given event has occured. This "evidence" narrows the calculation to a subset of the original probability space before it was given and is accounted for using this method.

(c) Briefly explain why the following expression is a correct use of Bayes's Theorem in odds form in this problem.

$$\begin{array}{ccccc} \left[\dfrac{P(St\,|\,Sh)}{P(E\,|\,Sh)}\right] & = & \left[\dfrac{P(St)}{P(E)}\right] & \cdot & \left[\dfrac{P(Sh\,|\,St)}{P(Sh\,|\,E)}\right] \\ (1) & = & (2) & \cdot & (3) \end{array}$$

This form of Bayes's Theorem represents the odds as ratios of 3 parts to show that any Bayesian analysis of a conditional event includes both the prior and posterior odds as well.

If we divide the posterior odds(1) by the prior odds(2) then multiply the Bayesian factor's denominator; also known as the likelihood ratio,(the bottom of 3) we are left with Bayes Theorem in terms of posterior and prior odds.

(over)

(d) Here are three terms that are relevant to the quantities in part (c) above:

- (Prior odds in favor of St over E)
- (Bayes factor in favor of St over E)
- (Posterior odds in favor of St over E)

Match these three terms with the numbers $(1), (2), (3)$ in the second line of the equation in part (c).

Posterior odds in favor of St over E(1) = Prior odds in favor of St over E(2) * Bayes factor in favor of St over E(3)

(e) Compute the three odds values in part (d), briefly explaining your reasoning, thereby demonstrating that the posterior odds value o in favor of St over E is $o = \frac{16}{27} \doteq 0.593$.

$\frac{0.372}{0.628} = \frac{0.10}{0.9} * \frac{0.8}{0.15}$

We are essentially showing the odds of finding out someone is shy and a statistician divided by the odds of finding out someone is shy and an economist(the posterior odds in favor of St) is equal to the ratio of attendees (St/E the prior odds ratio) * the ratio of shy statisticians versus shy economists ((Sh — St)/(Sh — E) the Bayes factor).

(f) Use the expression $p = \frac{o}{1+o}$ to show that the desired probability in this problem — the conditional probability that you're talking to a statistician — is $p = \frac{16}{43} \doteq 0.372$.

0.593/1.593 = 0.372 shows us that:

$$\frac{\textbf{Posterior Odds in favor of St}}{1+\textbf{Posterior Odds in favor of St}}$$

This is essentially Bayes Theorem in terms of Posterior odds.

(g) Someone says, "That probability can't be right: 80% of statisticians are shy, versus 15% for economists, so your probability of talking to a statistician has to be over 50%." Briefly explain why this line of reasoning is wrong, and why p should indeed be less than 50%.

p should be less than 50% merely based on the fact that of the 21.5% of shy attendees, more than half are NOT statisticians. This is mainly because of the 90-10 split in attendance. This certain someone may have confused their posterior odds with their Bayesian ones.