Task 2

Adam Morris – Computational Social Cognition Bootcamp, July 2017

After your stunning success in Task 1, Scotland Yard has given you the day off. You make your way over to a local pub ("The Horse and Buggy") for a cool, cool beer. As you sit, you observe a series of well-dressed people playing pool in the room next door. You wonder: How good are they?

Congratulations on completing Task 1. If you want to compare your code to mine, my solution is available at: http://mprlab327.webfactional.com/amorris/elementary/lie_detector.m.

In this task, you will infer the posterior probability of a continuous parameter – how good is someone at pool? – after watching them win or lose a series of games. Assume that everyone has a fixed internal parameter, θ, that represents their skill at pool. θ is between 0 and 1. For now, assume that θ directly translates into their probability of winning games: If a person's $\theta = 0.9$, then they have a 90% chance of winning each game they play.

Your goal here is to compute the posterior probability function over θ, given their win/loss record in a series of games – i.e. $Prob(\theta \mid record)$. Accomplishing this, however, is trickier than before. Before, we were only dealing with discrete hypothesis sets, and so the denominator in Bayes' rule was a sum over those hypotheses. But now, we're dealing with a continuous parameter. The reason that continuous parameters like θ are tricky is that the denominator in Bayes' rule becomes an integral, not a sum. Using Bayes' rule, the posterior probability over θ is:

$$Prob(\theta \mid record) = \frac{Prob(Record \mid \theta) * Prob(\theta)}{\int_{\theta=0}^{1} Prob(Record \mid \theta) * Prob(\theta) * d\theta}$$

Let's break this up into parts.

Part A

Write a function that computes the numerator in the equation above, for any θ and win/loss record. Assume you have a uniform prior over θ. (Hint: Your likelihood function should be similar to before.)

Part B

Now comes the tricky part. We need to compute that denominator. To do it, we're going to use a Riemann sum approximation to the integral. (Google this if you don't know what it is!) Instead of finding the exact area under the curve from $\theta = 0$ to $\theta = 1$, we'll break the interval up into little rectangles, find the area of each, and add them all together. Mathematically:

$$\int_{\theta=0}^{1} Prob(Record \mid \theta) * Prob(\theta) * d\theta \approx \sum_{\theta=0:stepSize:1} Prob(Record \mid \theta) * Prob(\theta) * stepSize$$

where *stepSize* is some small number, like 0.01.

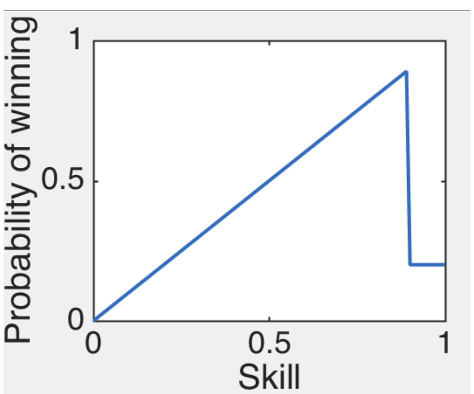
Using this approximation, write a function that computes the posterior probability for any θ and win/loss record. (Note that technically, you're computing the probability *density function* over θ, not the probability of any one value of θ. Hence, these “probabilities” can be greater than 1. The area under the density function tells you the probability that θ is between any two values.)

Plot the posterior over $\theta = 0\colon 0.01\colon 1$, assuming that the person won 2 games and lost 1. Now, replot it assuming that the person won 1 game and lost 9. Describe both results. In each case, what θ value is the peak of the distribution (called the *maximum a posteriori* estimate, or MAP)?

Part C

Now, let’s complicate things. You think about it more, and you realize that θ might not always translate directly into a person’s probability of winning a game. Sometimes, there are pool sharks – people who are amazing at pool, but purposefully lose in order to lure unsuspecting opponents into making ill-placed bets on future games.

How do we incorporate this belief into our Bayesian model? Assume that, from $\theta = 0$ to $\theta = 0.9$, θ translates directly into a person’s probability of winning games. But if a person is *really* good – i.e. $\theta > .9$ – then they only win 20% of their games (because they’re purposefully losing). In other words, the relationship between skill and win probability looks like this:



Redo Parts A and B with this change. (Hint: You should only have to change your likelihood function.) Plot the posterior, assuming that the person won 1 game and lost 9. What does the posterior look like now? Why?

Part D

Now, suppose you believe that pool players also experience “beginner’s luck” – i.e. very bad players consistently get lucky and win often. Redo Part C, incorporating this change into the likelihood function. *Now* what does the posterior look like? Why?

Parts C and D are meant to illustrate the power and flexibility of the Bayesian approach. By altering the prior and likelihood functions, you can model a vast set of beliefs that people might bring to bear on their judgment. What kinds of social-psychological situations could this approach be used to model?

If you want to compare your Task 2 code to mine, my solution is available at: http://mprlab327.webfactional.com/elementary/pool_sharks.m.