Assignment 1

Harry Bendekgey
Math 153: Bayesian Statistics

January 23 2018

1. I own two coins, each looking identical, however, one is known to have
   $P(H) = .5$ while the other has $P(H) = .4$. One of these two coins goes
   missing, and I am curious as to which coin remains (I don't think it is
   more likely for one coin to go missing more than the other).

   (a) Write down a prior for the coin I am in possession of, in terms of
       $\theta = P(H)$.

   Because we believe it is equally likely that we have either coin, we can say
   that $P(\theta = .4) = .5$ and $P(\theta = .5) = .5$.

   (b) I flip the coin 10 times and observe 3 heads. Write down the *likelihood*
       of the data as a function of θ. (coin flips are independent)

   Because this is a Bernoulli distribution, we can use its pmf to calculate
   likelihood, getting $p(x) = \theta^3 (1 - \theta)^7$

   (c) Using Bayes' Rule, write down the posterior distribution of θ given
       the outcome of the 10 flips.

   We apply Bayes' Rule:

   $$P(\theta = .5|H = 3) = \frac{.5 \cdot P(H = 3|\theta = .5)}{.5 \cdot P(H = 3|\theta = .5) + .5 \cdot P(H = 3|\theta = .4)}$$

   $$P(\theta = .5|H = 3) = \frac{.5^{10}}{.5^{10} + .4^3 .6^7} = .35$$

   Therefore $P(\theta = .4|H = 3)$ must be equal to .65

   (d) I bought the biased coin for \$10 because I will use it to cheat at a
       simple gambling game. The game consists of a \$100 dollar bet, where
       I win if I correctly call the outcome of a coin flip (and I will call tails
       because I we are playing with the coin I brought). I will only play
       this game once. Is it cost effective to buy a new coin or should I play
       with the one I have? Justify in terms of expected values.

First let us examine the expected value if you buy a new biased coin. The expected value is given by:

$$EV = .6 \cdot \$100 + .4 \cdot -\$100 - \$10 = \$10$$

Now, if you don't buy a new coin, the expected value is:

$$EV = .35 \cdot \$0 + .65 \cdot (.6 \cdot \$100 + .4 \cdot -\$100) = \$13$$

Here I take the shortcut of noting that if you have the unbiased coin, the expected value should be $0, as the coin toss is fair. From these calculations, we can see that it would not be cost effective to buy a new coin; just use the one you have.

2. I would like to estimate the chance that a standard normal random variable lives between 0 and 2, where the standard normal pdf is

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}.$$

Suppose that I have the ability to generate random draws from a uniform(0,1) distribution (random number between 0 and 1), where the pdf is just 1 in the unit interval and 0 outside of it (hence the nickname of the box distribution). Using the law of large numbers (see end of notes 2), write down a scheme for estimating this integral.

We want $\int_0^2 f_z(z)dz$. We have a function that can generate a random number between 0 and 1 with equal probability of any value. If we double the value generated, we have an equal probability of getting any value between 0 and 2. Written as a pdf, this is defined as: $\begin{cases} \frac{1}{2} & 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$. Let this be $f(x)$ for the purpose of the law of large numbers. We want to rephrase our integral such that we can identify $f(x)$ being multiplied by some other function $g(x)$. We see that $\int_0^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \int_0^2 \frac{1}{2} \frac{\sqrt{2}}{\sqrt{\pi}} e^{-x^2/2} dx$. Thus we see that our $g(x) = \frac{\sqrt{2}}{\sqrt{\pi}} e^{-x^2/2}$. Therefore if we generate numbers from our random generator, then double them, and plug the results into $g(x)$ as defined above, then average our results, the value will converge to this integral.

3. Other than moments, another way we will summarize the posterior is by returning intervals that contain a nominal amount of probability under the posterior as an analog to confidence intervals. As with confidence intervals, for a fixed confidence level, these intervals are more informative if they are smaller. Make an argument justifying that the interval (-1.96, 1.96) for a standard normal distribution is the *highest density interval* (HDI), i.e. minimizes $b - a$ such that $\int_a^b f(z)dz = .95$. Give conditions for which the HDI will look similar to this one, i.e. easy to find.

Given the interval (-1.96, 1.96) suppose we wanted to shift it slightly to the right. We know that $\int_b^{b+\Delta} f(z)dz < \int_{b-\Delta}^{b} f(z)dz = \int_a^{a+\Delta} f(z)dz$ for any small value of Δ. We know the equality is true because the normal distribution is symmetric, and we know the inequality above because the normal distribution always decreases when moving away from the mean (which is 0). Thus any shifts to the right will decrease the area in that interval, and a and b would have to be spread further apart to keep the same area. Symmetric logic can be used to show that shifts to the left don't work either. Thus this is the highest density interval.

I conjecture that any function that is symmetric around its mean and always decreasing from its mean will have an extremely easy to find HDI: namely, it will always be centered at the mean, and the only thing you need to find out is how wide does it need to be to meet the threshold.