Derivation of Laplacian Smoothing

Ernest Davis
May 11, 2021

Historically, this argument was originally advanced by Pierre-Simon Laplace to calculate the probability that the sun will rise tomorrow, given that it has been observed to rise N mornings in a row, for some large N. His answer was $(N+1)/(N+2)$ on the following argument:

Flipping a coin

Suppose that you flip a coin which has some unknown weighting, and it comes up heads n times. What is the probability that it will come up heads on the next flip?

Let us assume that the weighting of the coin W is uniformly distributed between 0 and 1. Let D be the observed event that the coin has come up heads n times and let H be the event that the next flip will be heads.

Then:

$$P(H|D) = \int_0^1 P(H, W=p|D)\,dp = \int_0^1 P(H|W=p)\cdot P(W=p|D)\,dp =$$

$$\int_0^1 P(H|W=p)\cdot P(D|W=p)\cdot P(W=p)/P(D)\,dp =$$

$$\frac{1}{P(D)}\int_0^1 P(H|W=p)\cdot P(D|W=p)\cdot P(W=p)\,dp$$

Now $P(H|W=p)=p$, $P(D|W=p)=p^n$, and $P(W=p)=1$ so

$$\int_0^1 P(H|W=p)\cdot P(D|W=p)\cdot P(W=p)\,dp = \int_0^1 p^{n+1}dp = \frac{1}{n+2}$$

$$P(D) = \int_0^1 P(D|W=p)\cdot P(W=p)\,dp = \int_0^1 p^n dp = \frac{1}{n+1}$$

So $P(H|D) = (1/(n+2))/(1/(n+1)) = (n+1)/(n+2)$ which is the Laplacian smoothing with $\delta = 1$.

Consider now the case where you have flipped the coin n times and it has come up heads k times, in some specified order. In this case $P(D|W=p) = p^k(1-p)^k$, so using the same analysis as before, we get

$$P(H|D) = \frac{\int_0^1 p^{k+1}(1-p)^{n-k}\,dp}{\int_0^1 p^k(1-p)^{n-k}\,dp}$$

Now if a and b are integers,

$$\int_0^1 p^a(1-p)^b\,dp = \frac{a!\cdot b!}{(a+b+1)!}$$

You can find a proof in the Wikipedia article on the Beta function. Using this on the two integrals above we get

$$P(H|D) = \frac{(k+1)!(n-k)!/(n+2)!}{k!(n-k)!/(n+1)!} = \frac{k+1}{n+2}$$

which again is the Laplacian smoothing with $\delta = 1$.

Rolling a multi-sided die

Now, consider the case where you are rolling a die with q sides n times and a particular side — call it "heads" — comes up k times. What is the probability that the next roll will be heads?

The first question to answer is, what is the equivalent of a uniform distribution? In particular, if weights are uniformly distributed across the sides, what is the probability distribution for the weight on heads?

We use the following model. Draw a circle of circumference 1. Drop q cut points uniformly and independently, so now you have divided the circle into q pieces. Choose a piece at random. Pretty clearly, this is a reasonable way of uniformly dividing 1 into q pieces.

Now modify this: After dropping the first cut point, we will decide that the first piece we will choose will be the one that just borders this cut counterclockwise and we will straighten out the circle into a straight line of length 1. So what we are doing now is we're dropping $q-1$ cut points uniformly in the interval [0,1] and considering the probability distribution of the length of the first piece.

What is the probability density that this first piece has length w? Well, one way that can happen is that cut #2 is at w, which has probability density 1, and the other $q-2$ cuts are all between w and 1, which has total probability $(1-w)^{q-2}$. Or this can happen with cut #3 being at w, and so on, up to cut #q. So the total probability density is $(q-1) \cdot (1-w)^{q-2}$. So we have $P(W = w) = (q-1) \cdot (1-w)^{q-2}$. It is easily checked that this has integral 1, so it is a valid probability density and that it has mean $1/q$, which obviously has to be right, so this seems OK.

We now can do the same calculation as with the die:

$$P(H|D) = \int_0^1 P(H|W=p) \cdot P(D|W=p) \cdot P(W=p)/P(D)dp =$$

$$\frac{\int_0^1 p \cdot p^k(1-p)^{n-k} \cdot (q-1) \cdot (1-p)^{q-2}dp}{\int_0^1 p^k(1-p)^{n-k} \cdot (q-1) \cdot (1-p)^{q-2}dp} = \frac{(k+1)!(n+q-k-2)!/(n+q)!}{k! \cdot (n+1+k-2)!/(n+q-1)!} = \frac{k+1}{n+q}$$

which again is the Laplacian smoothing.