# Handout 20: Bayesian Estimators

Today we cover another method for computing an estimator of an unknown quantity. This technique can also be extended to provide an alternative method of constructing confidence intervals, but we will not have time to cover these this semester.

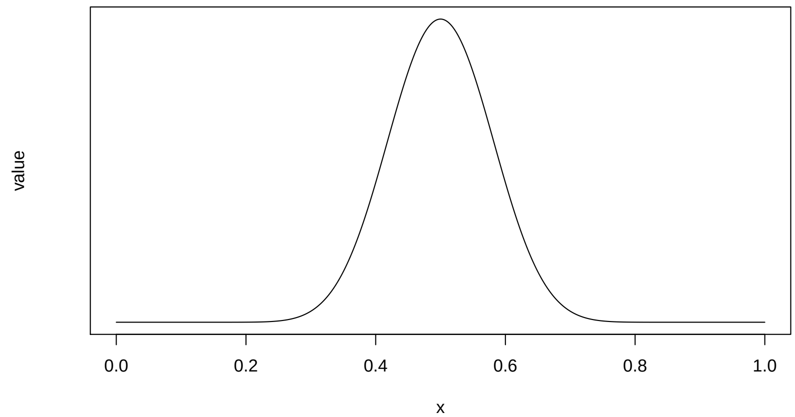
Consider being given the following bent half-dollar coin:

We want to flip the coin twice and estimate from these tosses the probability that the coin comes up heads in a random toss. The binomial distribution is a natural choice to use here, where we estimate the probability $p$ that the coin is flipped as a heads. If we observe both flips coming up heads, both the maximum likelihood and method of moments estimators yield the same result:

$$\widehat{p}_{MME} = \frac{1}{n}\sum_i x_i = 1$$
$$\widehat{p}_{MLE} = \frac{1}{n}\sum_i x_i = 1$$

Does this really make sense as an estimator? Do you really believe that the coin comes up heads *every single time*? Likely not!

What's the issue here? The problem is that we have some prior knowledge that is not being used in the estimator. We know that a value equal to or very close to 1 is highly unlikely. In the absence of any flips we might have a picture like this of how likely each value of $p$ would be:




Generally values are centered around 0.5, but there is room for being somewhat biased by the bend in the coin. Without much thought, it is unclear whether the bent will make $p$ tend to be higher or lower. How can we describe this situation use probabilities? Try this:

$$p \sim Beta(20, 20)$$
$$X|p \sim Bin(2, p)$$

I knew to fill in a $Beta(20, 20)$ because that is the density that yielded the prior plot. How could we use this to calculate an estimate of $p$? Consider the conditional probability $p|X$, which we have already calculated on a previous worksheet:

$$p|X \sim Beta(20 + \sum_i X_i, 20 + \sum_i (1 - X_i)) = Beta(22, 20)$$

This measures what we know about $p$ based on our prior belief but updated with new data. A reasonable estimator is to take the mean of the conditional probability:

$$\mathbb{E}\left[p|X\right] = \frac{22}{22 + 20} \approx 0.523$$

So the two flips increased our initial best estimate from 0.5 to 0.523, but did not make it jump all of the way to the (unreasonable) value of 1.

The method we just described yields a Bayesian estimator. The initial distribution on $p$ is called the *prior*, the distribution on the data is the *likelihood*, and the conditional distribution $p|X$ is the *posterior*. When the posterior belongs to the same family as the prior, the prior and likelihood are said to be *conjugate priors*. We will explore several of these on today's worksheet.