STAT 491 - Lecture 2

January 16, 2018

Ch.2 Introduction: Credibility, Models, and Parameters

Reallocation of probabilities

Recall the Guess Who setting, where the goal was to identify the opponents character.

In this setting, after determining that the character wore a hat and non-purple glasses we ended up with the following probability.

In this situation we were able to deduce that Donna was our opponents character.

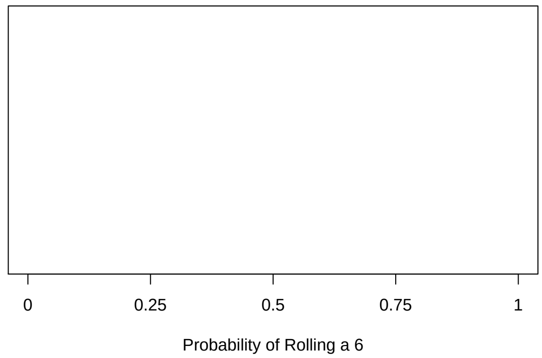
Ex. Rolling a Die

Now consider a similar example using a die. Suppose our goal is to determine the probability of the die landing on 6. Now constuct your prior belief for this die.

Note this is different from the previous example.

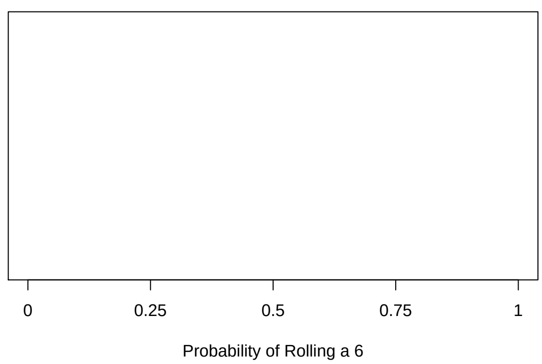
-

-

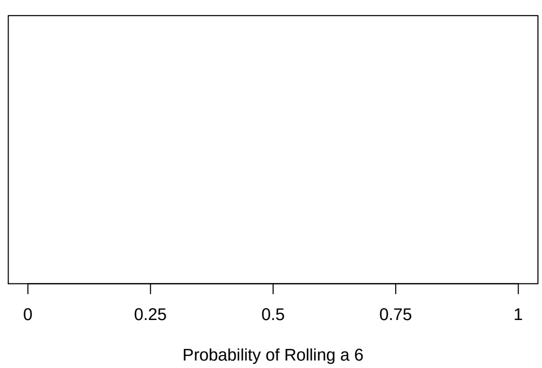


Data are noisy and inferences are probabilistic

Suppose we observe three rolls, now update your beliefs.



Finally suppose we observe 100 more rolls, now update your beliefs.



Possibilities are parameter values in descriptive models

Data resulting from the roll of dice can be characterized by a Multinomial distribution. We can formally characterize this using the probability mass function (pmf) of the random variable (more details in later sections.)

$$Pr[\text{die } = i]\frac{n!}{x_1!x_2!x_3!x_4!x_5!x_6!}p_1^{x_1}\ p_2^{x_2}\ p_3^{x_3}\ p_4^{x_4}\ p_5^{x_5}\ p_6^{x_6},$$

where p_i is the probability of rolling an i, x_i is the total count of $i's$ rolled in n total rolls.

-
-
-

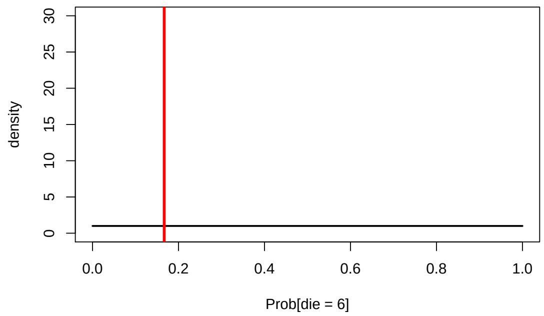
Given that our goal is to only estimate, $P[die = 6]$, we can simplify the pmf above and consider only two cases: 6 and not 6. This is known as a binomial random variable and can be written as

where p is the probability of rolling a 6 and x is the number of ones rolled out of n total rolls.

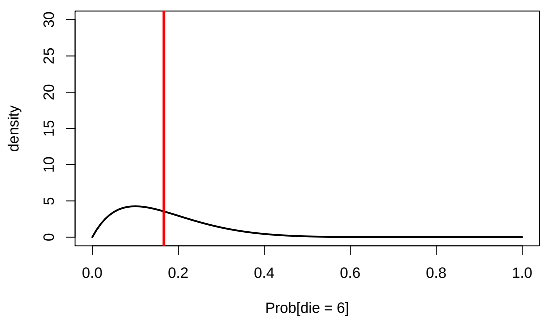
Question: So, why do we need these mathematical models?

Mathematical Notation for Coin Flipping Example

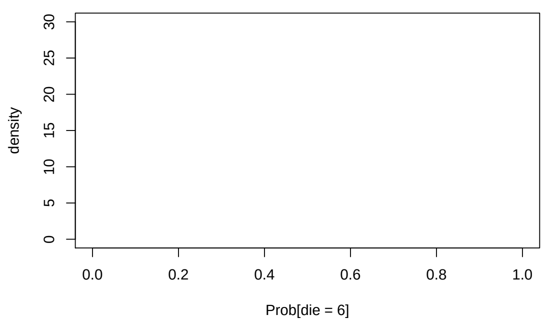
Suppose our initial prior is a uniform prior over the range of values from 0 to 1.



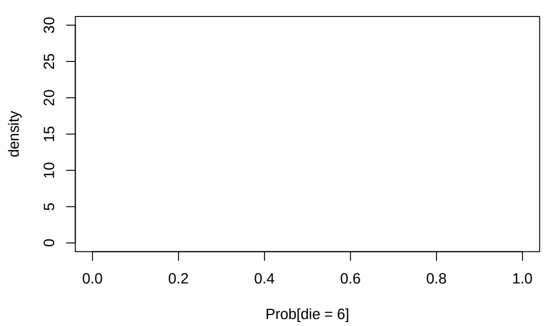
Now assume we are using a fair die and update the probabilities after the following ten rolls: 2, 5, 3, 1, 1, 5, 4, 1, 4, 6. Our distribution can be updated as



Now we observe 100 more rolls of the die, which results in 12 rolls of 1.

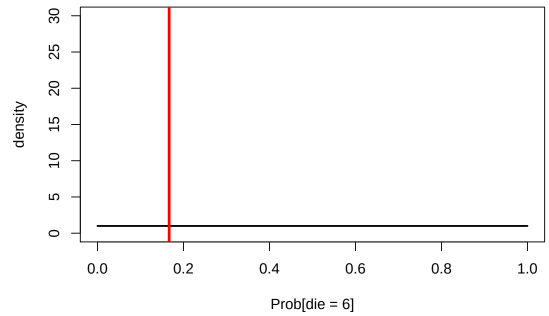


After 1000 more rolls of the die, which results in 158 rolls of 1.

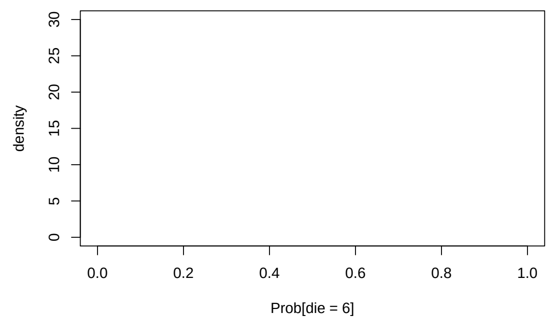


Unfair Coin

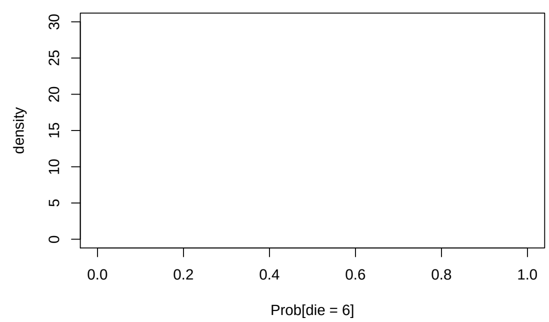
Suppose our initial prior is a uniform prior over the range of values from 0 to 1.



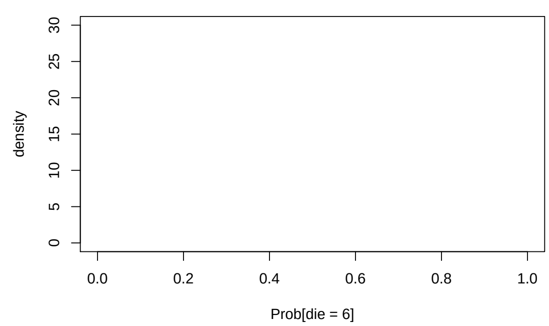
Now assume we are using a fair die and update the probabilities after the following ten rolls: 1, 6, 6, 5, 6, 6, 5, 2, 2, 6. Our distribution can be updated as



Now we observe 100 more rolls of the die, which results in 45 rolls of 1.



After 1000 more rolls of the die, which results in 517 rolls of 1.



Steps of Bayesian Data Analysis

For a Bayesian analysis we will follow these steps:

1. **Identify the data relevant to the research questions.** What are the measurement scales of the data? Which data variables are to be predicted, and which data variables are supposed to act as predictors?

2. **Define a descriptive model for the relevant data.** The mathematical form and its parameters should be meaningful and appropriate to the theoretical purposes of the analysis.

3. **Specify a prior distribution on the parameters.** The prior must pass muster with the audience of the analysis, such as skeptical scientists.

4.

5.