STAT 491 - Lecture 5

February 1, 2018

Ch.5 Bayes Rule

Compare the two probability statements:

The first probability statement considers two possible outcomes:

The second probability statement incorporates some additional information (data) into the probability statement,

Bayes rule is the mathematical foundation for re-allocating credibility (or probability) when conditioning on data.

Bayes Rule and Conditional Probability

- Recall: the conditional probability $P(A|B) = \frac{P(A \cap B)}{P(B)}$. From here we do some algebra to obtain Bayes rule.

- Either of the last two equations are called **Bayes Rule**, named after Thomas Bayes.

Bayes Rule with two-way discrete table

	COLUMN	
ROW	c	**Marginal**
r	p(r,c) = p(r|c)p(c) = p(c|r)p(r)	p(r)
Marginal	p(c)	

Recall the following two-way table:

	Hair Color				
Eye Color	Black	Brunette	Red	Blond	Marginal (eye color)
Brown	0.11	0.20	0.04	0.01	0.37
Blue	0.03	0.14	0.03	0.16	0.36
Hazel	0.03	0.09	0.02	0.02	0.16
Green	0.01	0.05	0.02	0.03	0.11
Marginal (hair color)	0.18	0.48	0.12	0.21	1.0

Previously we calculated:

- What is the probability of a person having red hair given that they have blue eyes

  We now see that this is a simple illustration of Bayes rule.

A classic example of Bayes rule focuses on diagnosing a rare disease. There are a few important values we need to state:

- Let θ be
- Let T be
- Let $Pr(\theta = Yes) = p_\theta$ be
- Let $Pr(Test = Yes|\theta = Yes) = p_{T+}$ be
- Let $Pr(Test = Yes|\theta = No) = p_{T-}$ be
- **Question:** do we need,
  - $Pr(\theta = No)$
  - $Pr(Test = No|\theta = Yes)$
  - $Pr(Test = No|\theta = No)$

Assume we are testing citizens for Extra Sensory Perception (ESP). The ultimate goal will be to determine the probability that an individual has ESP if they test positive for ESP. Mathematically this is stated as $Pr(\theta = Yes|Test = Yes)$.

First using the generic probability from the previous page, compute $Pr(\theta = Yes|Test = Yes)$.

$$Pr(\theta = Yes|Test = Yes) \quad =$$

Now to make this concrete assume:

- The rate of ESP in the population is 1 in
- The hit rate of the test is
- The false detection rate is 1 in
- **Question:** Before doing any math, what is your guess for the probability that a person receiving a positive test actually has ESP?

```
p.theta <- 1 / 100000
p.t.plus <- 9999 / 10000
p.t.minus <- 1 / 10000
p.theta.true <- p.t.plus * p.theta / (p.t.plus * p.theta + p.t.minus * (1 - p.theta))
```

It turns out that the probability that a person actually has ESP given they had a positive test is $Pr(\theta = Yes|Test = Yes) =$

This example allows us to understand the mechanisms behing Bayes rule.

-
-

Bayes rule with parameters and data

The previous example was essentially a probability exercise and we were not doing Bayesian statistical analysis per se, but rather just using Bayes rule. Bayesian statistical analysis refers to a fairly specific application of this theorem where:

-

-

- Bayes rule is used to convert the prior belief on the parameters *and* the statistical model into a **posterior** belief $p(\theta|\mathcal{D})$.

Example of Bayesian Analysis on a binary outcome

Consider estimating the probability that a die will roll a six and recall the 5 steps in a Bayesian analysis

1. Identify the data relevant to the research question.
2. Define a descriptive model for the relevant data.
3. Specify a prior distribution on the parameters.
4. Use Bayesian inference to re-allocate credibility across parameter values.
5. Check that the posterior predictions mimic the data with reasonable accuracy.

1. Identify the data relevant to the research question.

- What data do we need to determine of the probability that a die lands on a six?

2. Define a descriptive model for the relevant data.

A descriptive model denoted as $p(\mathcal{D}|\theta)$ is needed for the die rolling experiment.

- what is $\mathcal{D} = \{d_1, d_2, \ldots, d_n\}$

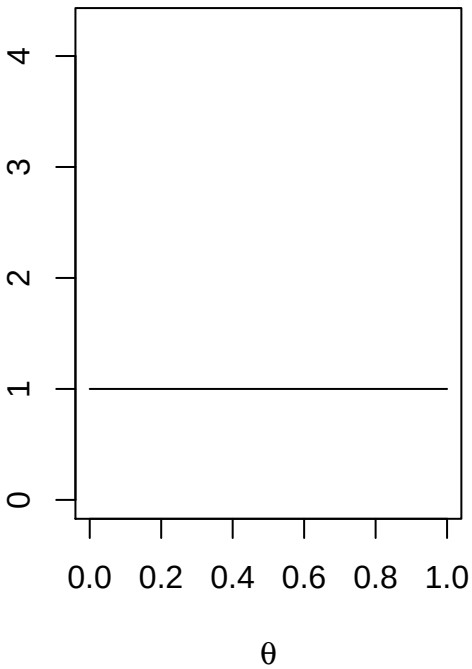
- what is θ:

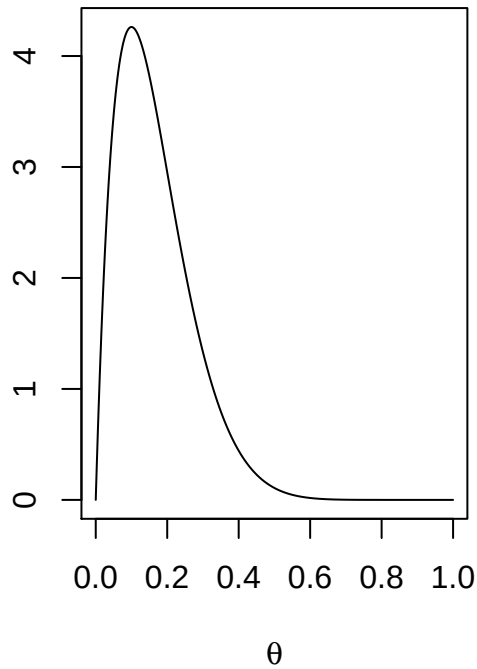
- what is a descriptive model for $p(\mathcal{D}|\theta)$,

This model is related to a binomial distribution and will be the mathematical machinery that allows updated prior beliefs in a formulaic manner.

3. Specify a prior distribution on the parameters

Here are a couple of reasonable prior distributions on θ, the probability of rolling a 6.





Discuss the implications behind each figure.

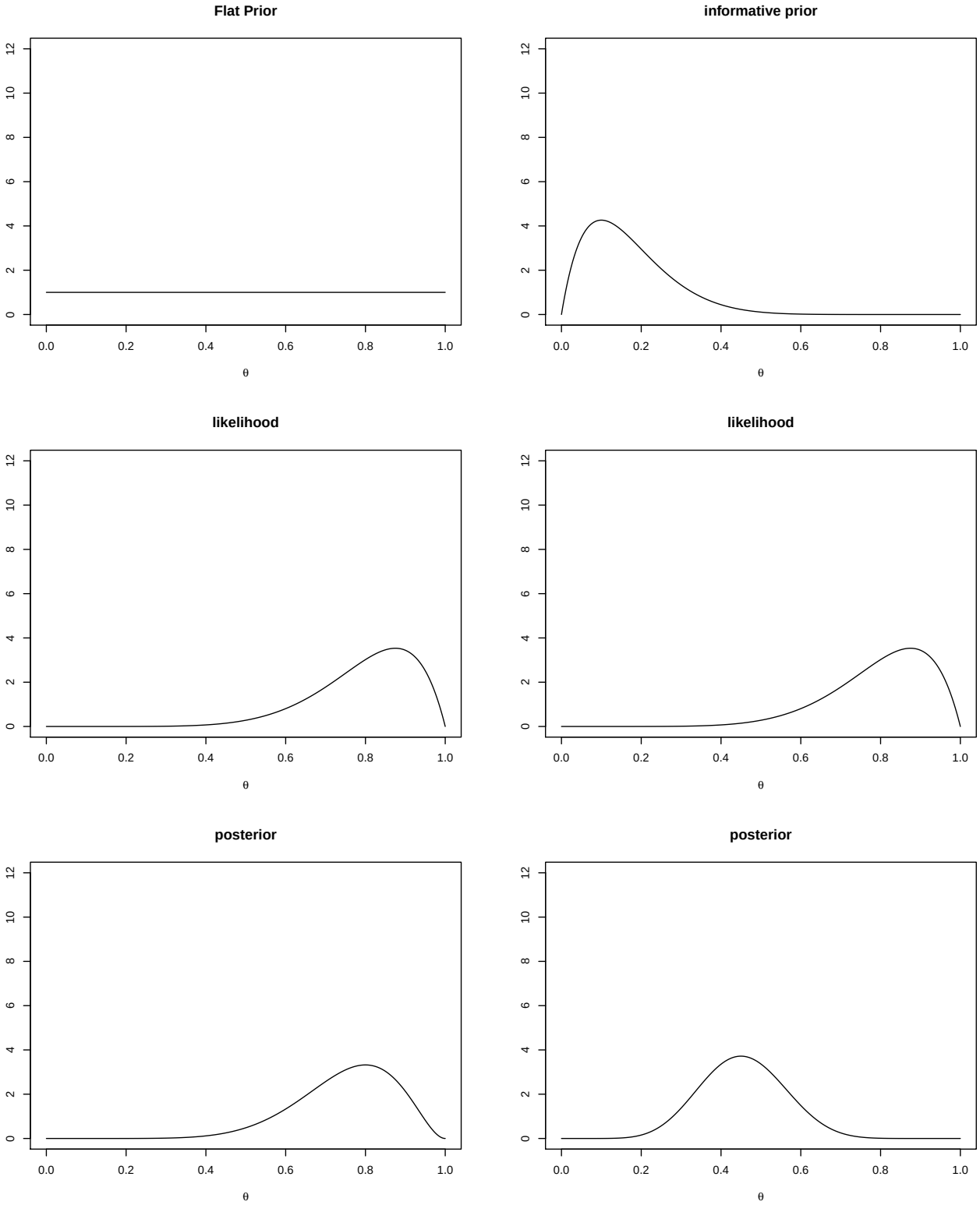
4. Use Bayesian inference to re-allocate credibility across parameter values.

Recall the goal of this analysis was to learn about θ the probability of rolling a six. Specifically, we are interested in the posterior distribution $p(\theta|\mathcal{D})$.

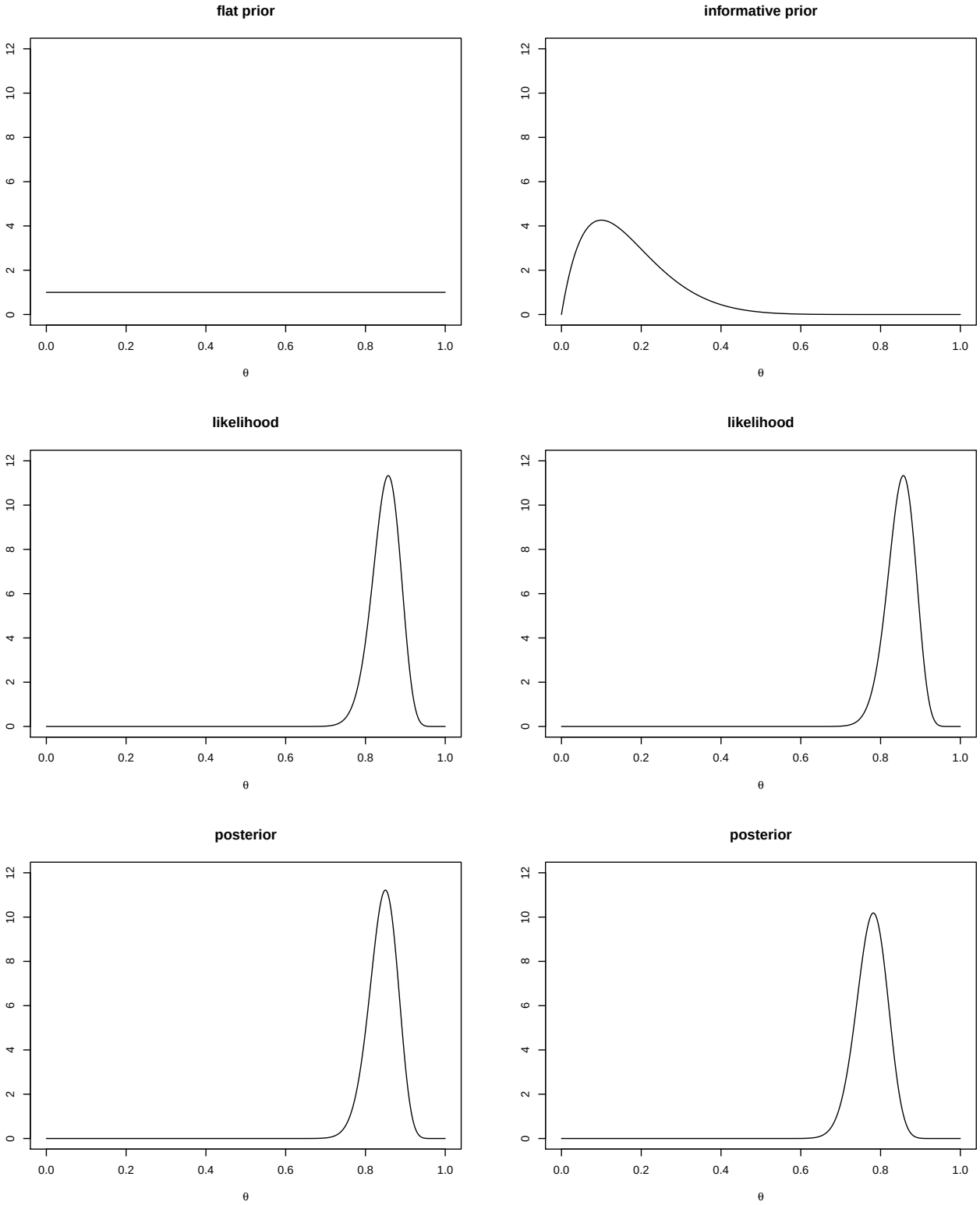
Lets assume a few data collection procedures:

1. 10 rolls of the die, with
2. 25 rolls of the die, with
3. 100 rolls of the die, with

With 10 rolls



Then with 100 rolls



influence of the sample size and prior on posterior

With Bayesian statistics there is an interplay between the strength of our prior beliefs and the amount of data collected.

- The posterior distribution can be considered as a weighted average between the prior distribution and the data.

-

-