Week Three: Video Lecture

This Week: Bayesian Estimation for binomal and multinomial data

Tuesday:

- Watch Week 3 videos and submit HW 2 (video notes)
- Week 3 activity

Thursday:

- Lab 2

Primer for Bayesian Estimation

- Maximum Likelihood Estimators are based strictly on observed data.
- Bayesian estimation incorporates prior information into estimation.

Prior Distributions

- Bayesian inference mimics (imo) human learning, where you observe information and update beliefs about the world
- Bayesian thinking is inherently distributional, prior specification requires a probability distribution
- Bayesian statistics permits statements like, "there is a 95% probability…" as opposed to "confidence'

Example:

Let's consider estimating the probability that you make it through the intersection of 19th and Main without stopping at the light. Based on your experience sketch a figure that contains this probability. Remember this should be a distribution.

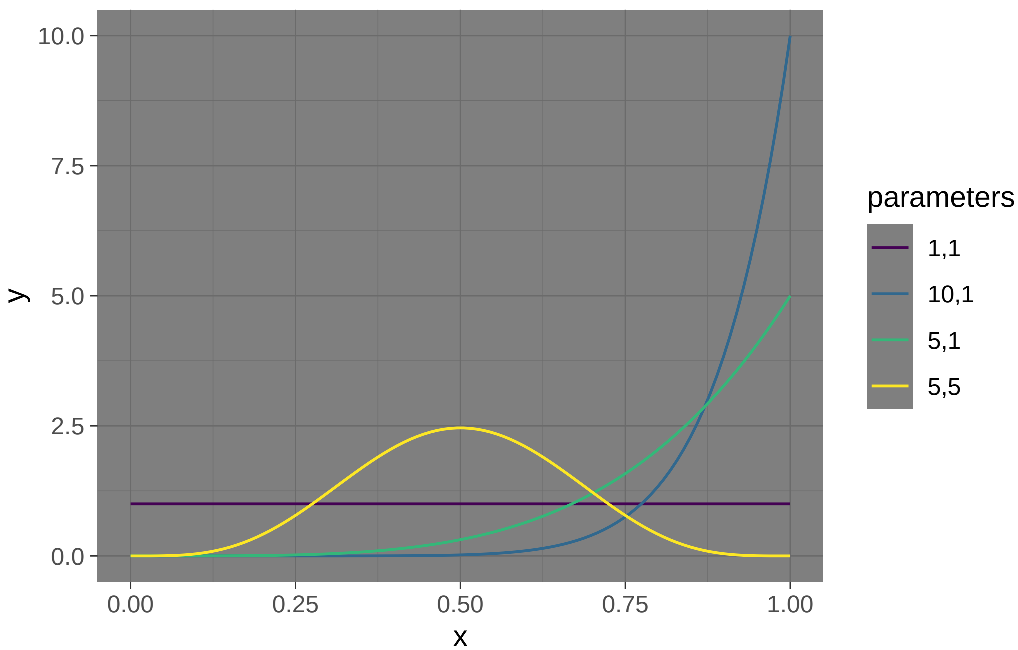
Beta Distribution

The beta distribution,

$p(x) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} x^{\alpha-1}(1-x)^{\beta-1}$, x $\in [0, 1]$

is a common distribution for modeling values between 0 and 1.

```
library(tidyverse)
library(viridis)
num <- 100
x <- seq(0,1, length.out = num)

tibble(y = c(dbeta(x, 10, 1), dbeta(x, 5, 1), dbeta(x, 5, 5), dbeta(x, 1, 1)),
       x = rep(x, 4),
       parameters = c(rep('10,1', num), rep('5,1', num), rep('5,5', num), rep('1,1', num)))
  ggplot(aes(y=y, x=x, color = parameters)) +
  geom_line() +
  theme_dark() +
  scale_color_viridis(discrete=TRUE)
```



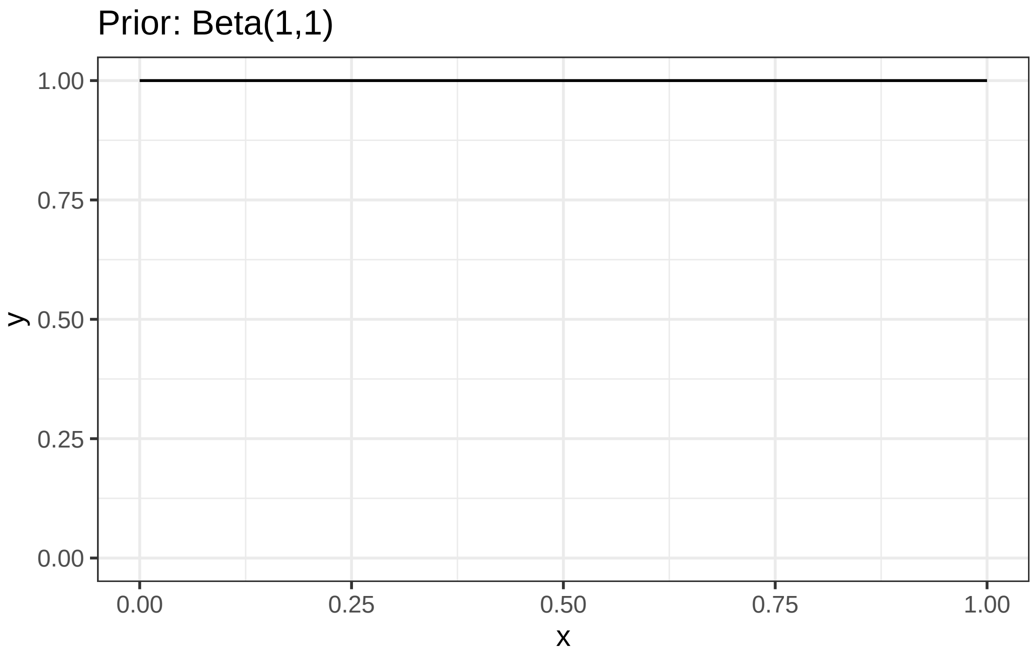
Posterior distribution

A posterior distribution combines prior information with data.

In the case of binary / binomial data, the beta distribution is a convenient prior distribution as the posterior (with beta prior and binomial data results in a beta posterior).

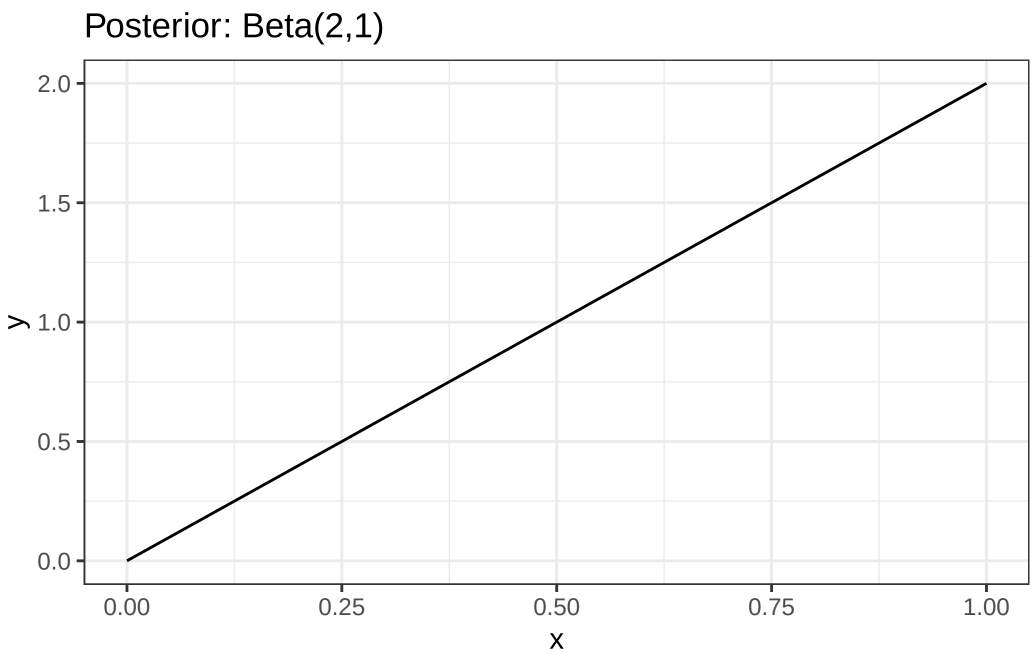
Assume you start with a uniform distribution (beta 1, 1)

```
num <- 100
x <- seq(0,1, length.out = num)

tibble(y = dbeta(x, 1, 1),
       x = x) |>
  ggplot(aes(y=y, x=x)) +
  geom_line() +
  theme_bw() +
  scale_color_viridis(discrete=TRUE) +
  ylim(0,1) +
  ggtitle('Prior: Beta(1,1)')
```



Next, we observe on success (Sweet Peaks), then the new posterior is a beta distribution with parameters (alpha = 1 + 1, beta = 1 + 0)

```r
num <- 100
x <- seq(0,1, length.out = num)

tibble(y = dbeta(x, 2, 1),
       x = x) |>
  ggplot(aes(y=y, x=x)) +
  geom_line() +
  theme_bw() +
  scale_color_viridis(discrete=TRUE) +
  ggtitle('Posterior: Beta(2,1)')
```



Assume, we observe on 7 success (Sweet Peaks) and 9 failures, then the new posterior is a beta distribution with parameters (alpha = 1 + 7, beta = 1 + 9)

```r
num <- 100
x <- seq(0,1, length.out = num)

tibble(y = dbeta(x, 8, 10),
       x = x) |>
  ggplot(aes(y=y, x=x)) +
  geom_line() +
  theme_bw() +
```

```
scale_color_viridis(discrete=TRUE) +
ggtitle('Posterior: Beta(8,10)')
```