Week Three

Last Week

- Inference for Proportions
  - Statistical probability distributions (Binomial & Multinomial)
  - MLE estimation for parameters
  - Estimation vs. Testing
- Bayesian Primer Video
  - Prior Distributions
  - Beta Distribution (for proportions)
  - Posteriors Distributions

This Week: Bayesian Inference for Proportions

Today:

- Activity 2
  - Bayesian Inference for Binomial & Multinomial Distributions
  - Bayesian Inference vs. Maximum Likelihood Estimation
- Thursday: Lab 2

Next Week: Contingency Tables

for Tuesday:

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

Bayesian Inference Overview

The Bayesian statistics paradigm follows three basic steps.

1. Specify prior belief (distribution) about model parameters.

2. Collect data (assumed to follow statistical probability distribution: Likelihood)

3. Posterior distribution defined by Bayes rule: Prior + Likelihood $\rightarrow$ Posterior

Beta Distribution (Priors)

For modeling binary categorical data, we have seen how the binomial distribution is useful. The parameter we hope to estimate is π, which is restricted to be between 0 and 1.

The beta distribution is a good distribution for this case.

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

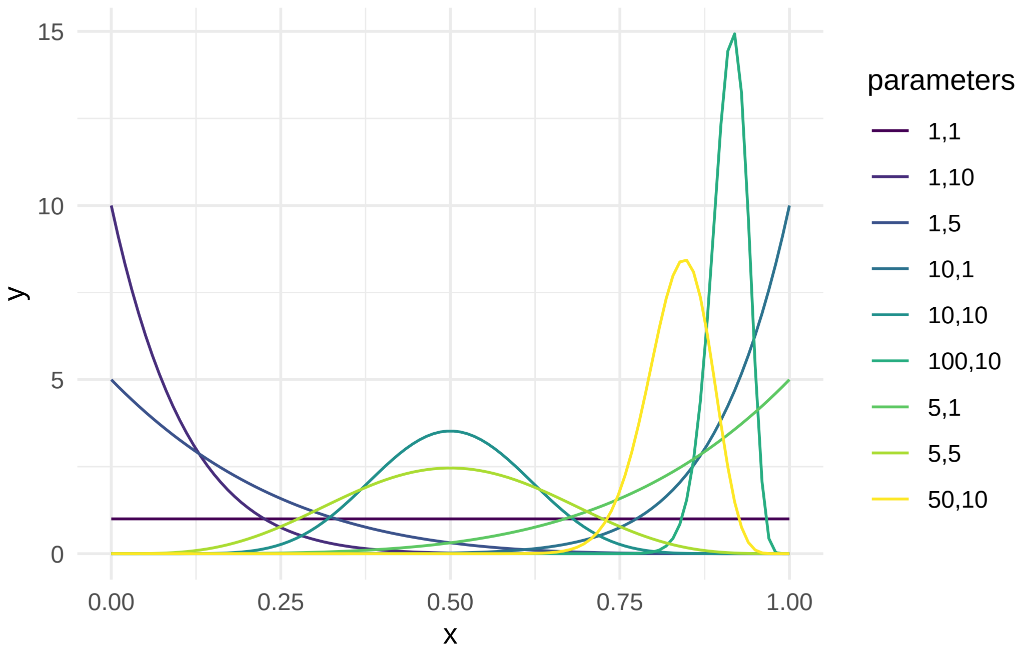
What are the parameters in this distribution (and what are the moments- mean & variance)?

α *and* β

$Mean = \frac{\alpha}{\alpha+\beta}$

$Var = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

How do they impact the shape of the distribution? Overlay curves with a wide range of parameter values.



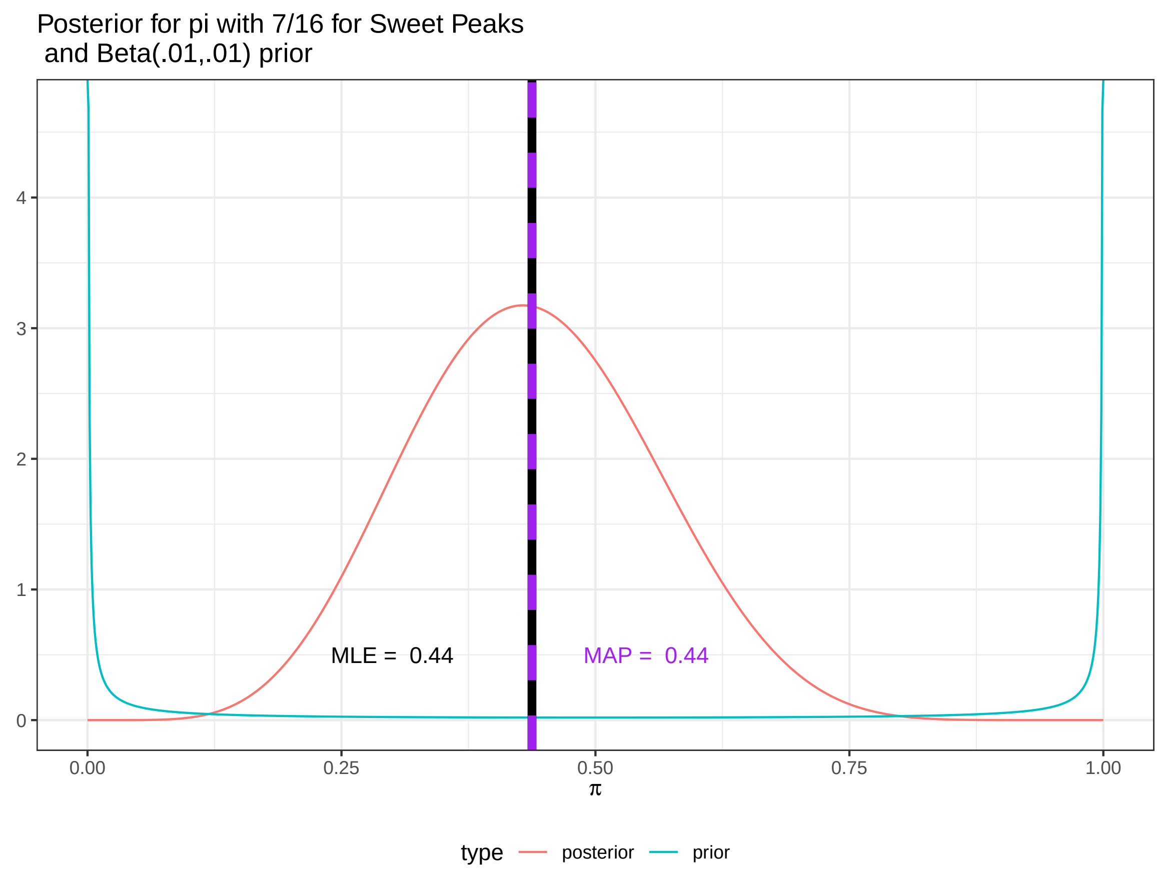
Bayesian Estimation for Binary Data

On the first day of class, we intuitively estimated π for the ice cream problem to be $\frac{7}{16} = 0.44$, which was also the MLE.

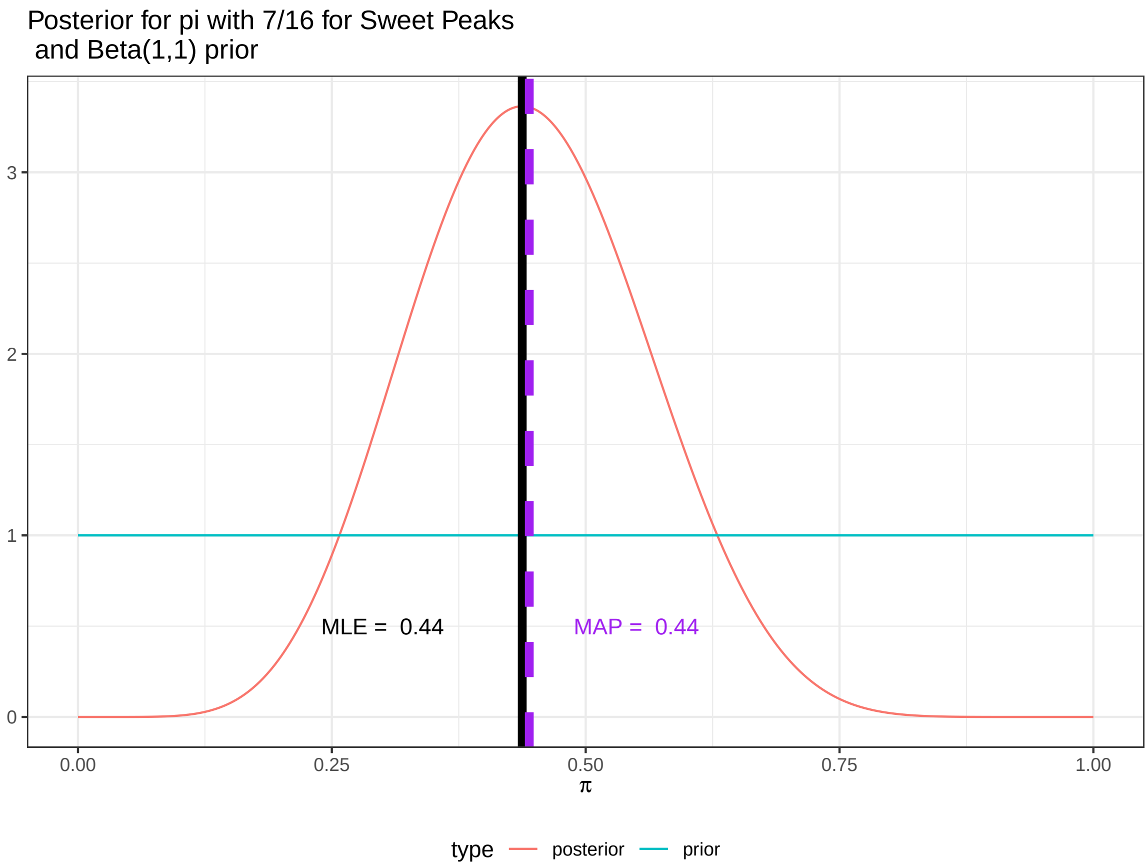
In this setting, the posterior distribution is exactly defined to be a Beta distribution with parameters $y + \alpha$ and $N - y + \beta$.

Let's visualize this setting with a set of different prior distributions and add the posterior mean and MLE.

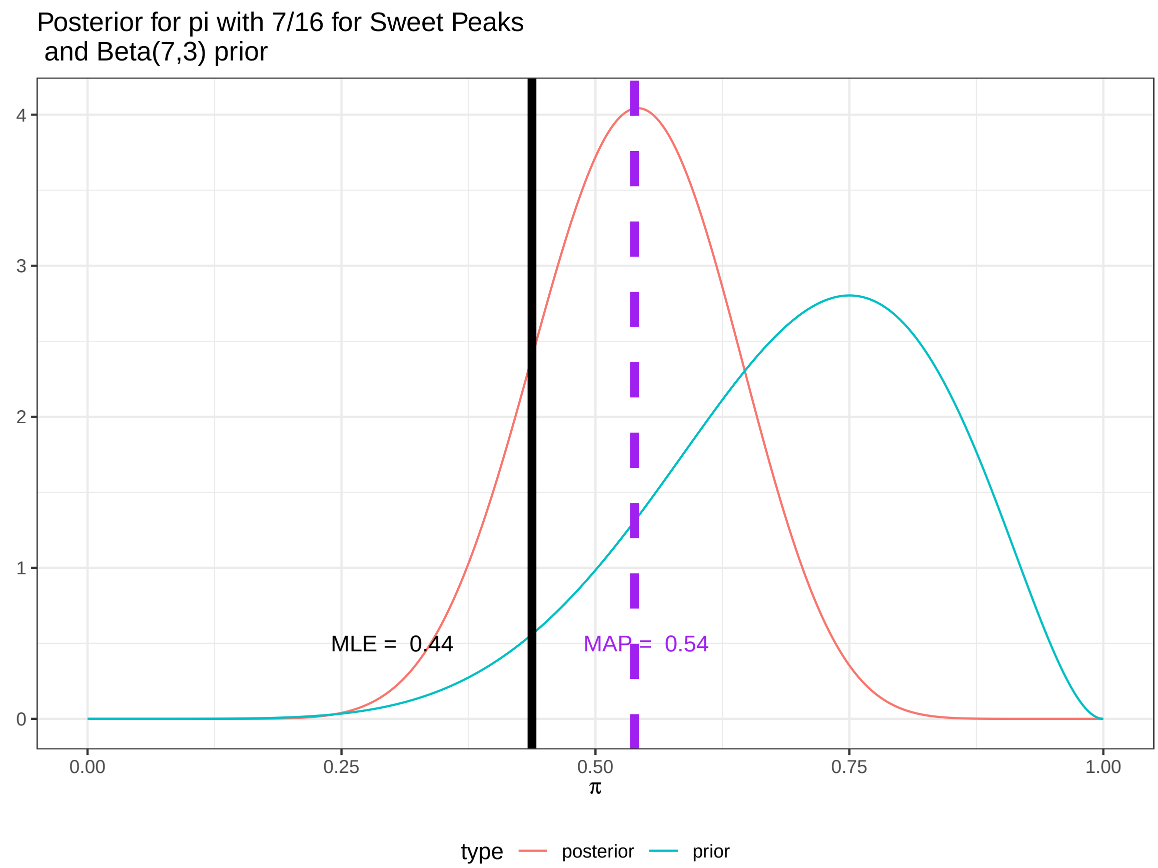
- Beta(.01, .01)



- $\text{Beta}(1,\, 1)$



- $\text{Beta}(7,\, 3)$

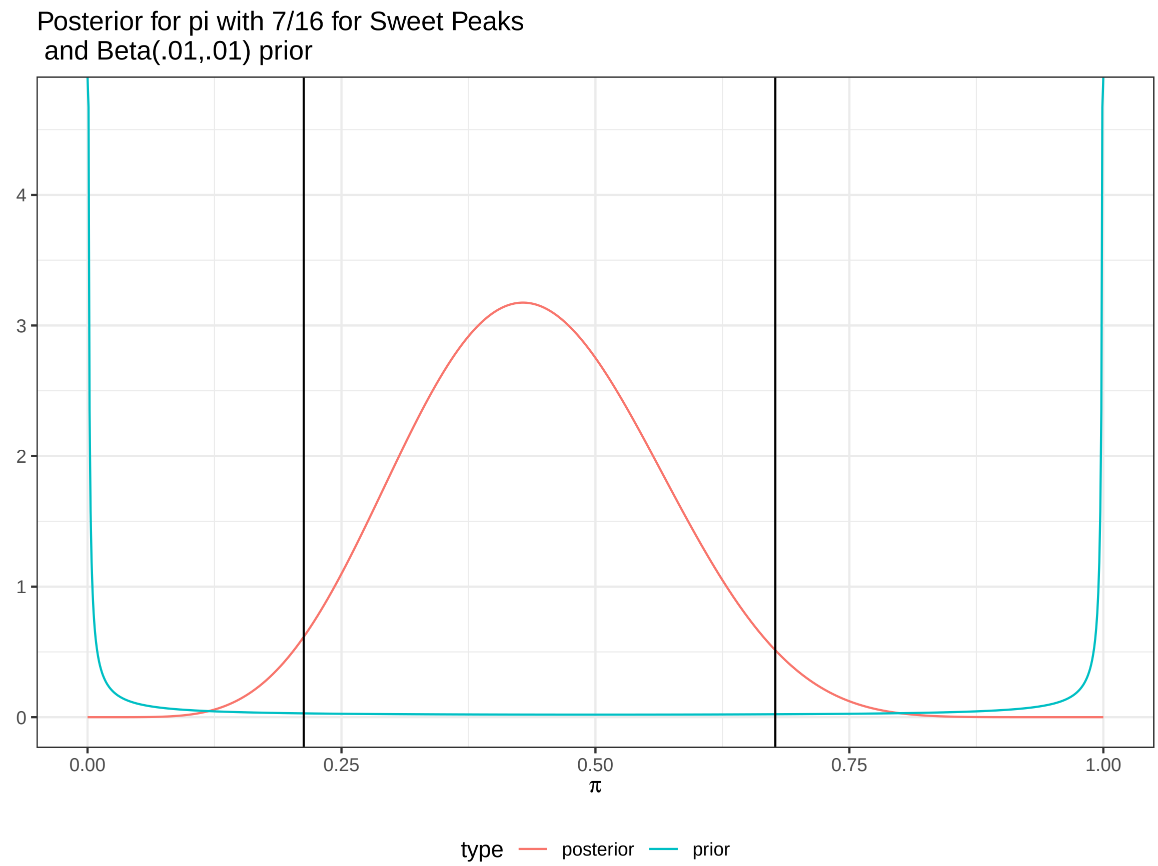


Posterior for pi with 7/16 for Sweet Peaks
and Beta(7,3) prior
MLE = 0.44
MAP = 0.54
0
1
2
3
4
0.00
0.25
0.50
0.75
1.00
π
type
posterior
prior

Estimation: Bayesian Uncertainty intervals

Unlike a likelihood profile, with a Bayesian posterior distribution there is a natural way to obtain a 95% (or other level) uncertainty interval.

Consider the posterior generated using a beta(.01, .01) prior, a Beta(7.01, 9.01) distribution. We can can simply trim the quantiles from the distribution. So the 95% interval boundaries are at 0.21 and 0.68.



Is this the only possible interval?

Multicategory Outcomes

Thus far, we have simplified the ice cream example to a binary situation (Sweet Peaks or not).

However, the data was collected as a multicategory setting with many possible answers.

An extension to the binomial distribution, known as the multinomial distribution is appropriate in this case.

$$P(Y_1 = y_1, Y_2 = y_2, ..., Y_k = y_k) = \frac{n!}{y_1!y_2!...y_k!}\pi_1^{y_1}\pi_2^{y_2}...\pi_k^{y_k}, \quad \text{where } y_i = 0, 1, 2, ...$$

To estimate the parameters in this distribution we can use a Dirichlet distribution as a prior distribution, which is a multivariate extension to the beta distribution.

$p(x) = \frac{\prod_i \Gamma(\alpha_i)}{\Gamma(\sum_i \alpha_i)} \prod_i \pi_i^{\alpha_i - 1}$, _i ∈ [0, 1]

Given the similarities to the previous example, what do you anticipate the form of the posterior distribution would look like in this case with multinomial data and a dirichlet prior.

The posterior would be Dirichlet with parameters $\alpha_i + y_i$.

Recall our data from the first day of class where we had the following data.

- Sweet Peaks 7
- Big Dipper 2
- Ben and Jerry's 2
- DQ 1
- Coldstone 1
- Wendys 1

So that we can visualize the distribution, lets consider just three options: Sweet Peaks, Big Dipper, and Ben and Jerry's.

Here is a prior distribution, given a Dirichlet prior where $\alpha_i = 1$.

```r
# Code modified from Google Gemini
library(ggtern)
```

```
Registered S3 methods overwritten by 'ggtern':
  method           from
  grid.draw.ggplot ggplot2
  plot.ggplot      ggplot2
  print.ggplot     ggplot2
```

```
--
Remember to cite, run citation(package = 'ggtern') for further info.
--
```

```
Attaching package: 'ggtern'
```

```
The following objects are masked from 'package:ggplot2':

    aes, annotate, ggplot, ggplot_build, ggplot_gtable, ggplotGrob,
    ggsave, layer_data, theme_bw, theme_classic, theme_dark,
    theme_gray, theme_light, theme_linedraw, theme_minimal, theme_void
```

```r
library(MCMCpack)
```

```
Loading required package: coda
```

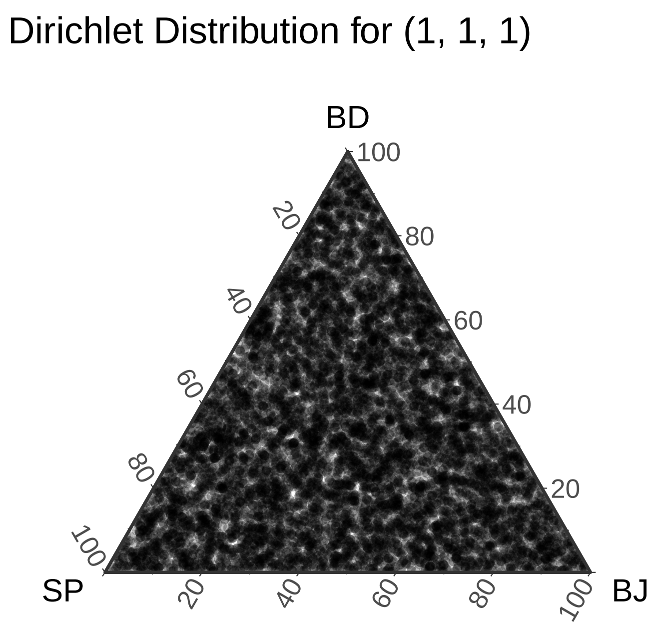
```
Loading required package: MASS
```

```
Attaching package: 'MASS'


The following object is masked from 'package:dplyr':

    select


##
## Markov Chain Monte Carlo Package (MCMCpack)


## Copyright (C) 2003-2025 Andrew D. Martin, Kevin M. Quinn, and Jong Hee Park


##
## Support provided by the U.S. National Science Foundation


## (Grants SES-0350646 and SES-0350613)
##
```

```r
dirichlet_samples <- as_tibble(rdirichlet(10000, c(1, 1, 1)))
```

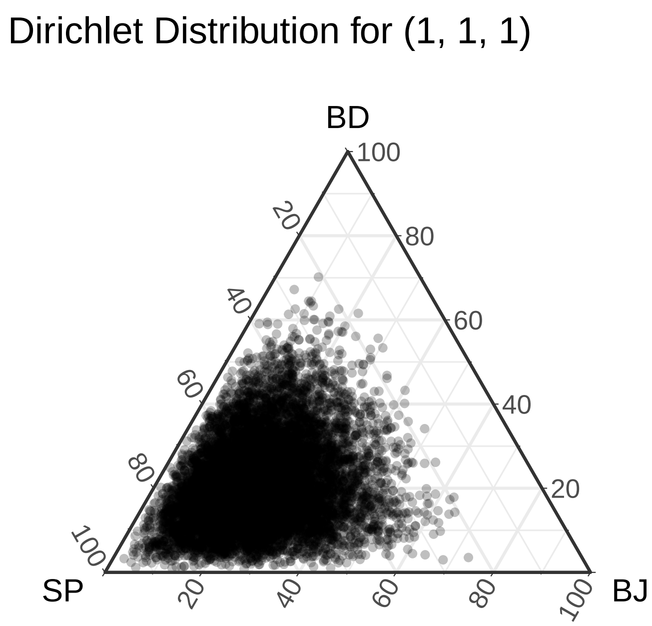
```
Warning: The `x` argument of `as_tibble.matrix()` must have unique column names if
`.name_repair` is omitted as of tibble 2.0.0.
i Using compatibility `.name_repair`.
```

```r
colnames(dirichlet_samples) <- c('SP', 'BD', 'BJ')



# Plot the Dirichlet distribution on a ternary plot
ggtern(data = dirichlet_samples, aes(x = SP, y = BD, z = BJ)) +
  geom_point(alpha = 0.25, size = 1) +
  labs(title = "Dirichlet Distribution for (1, 1, 1)",
       Tlab = "x1", Llab = "x2", Rlab = "x3") +
  theme_bw()
```



Modify this code to construct a posterior distribution and interpret the figure.

```
post_samples <- as_tibble(rdirichlet(10000, c(1 + 7, 1 + 2, 1 + 2)))
colnames(post_samples) <- c('SP', 'BD', 'BJ')


# Plot the Dirichlet distribution on a ternary plot
ggtern(data = post_samples, aes(x = SP, y = BD, z = BJ)) +
  geom_point(alpha = 0.25, size = 1) +
  labs(title = "Dirichlet Distribution for (1, 1, 1)",
       Tlab = "x1", Llab = "x2", Rlab = "x3") +
  theme_bw()
```



From this distribution our mean estimates are:

- SP 0.57
- BD 0.21
- BJ 0.21

We can also contruct uncertainty intervals

```r
rdirichlet(10000, c(8, 3, 3)) |> apply(2, quantile, probs = c(.025, .975))
```

```
            [,1]       [,2]       [,3]
2.5%  0.3126380 0.05012188 0.04950872
97.5% 0.8064262 0.45866513 0.45783998
```