

# Assignment #1

Elliot Smith

8/26/2018

## Problem 1

### Part i

$f(x)$  is the pdf of any of the  $x_i$ , and since they are *iid*, they all follow one pdf.

The joint pdf is:  $f_{X_1, \dots, X_n}(x_1, \dots, x_n) = f(x_1) \times \dots \times f(x_n)$  because they are *iid*.

Since you can multiply the right-hand side in any order, this implies symmetry in regards to the left-hand side.

As a result,  $X_1, \dots, X_n$  are exchangeable.

### Part ii

$$\begin{aligned} p(y_1, \dots, y_n) &= \int p(y_1, \dots, y_n | \theta) p(\theta) d(\theta) \\ &= \int \left( \prod_{i=1}^n p(y_i | \theta) \right) p(\theta) d(\theta) \\ &= \int \left( \prod_{i=1}^n p(y_{\pi_i} | \theta) \right) p(\theta) d(\theta) \\ &= \int p(y_{\pi_1}, \dots, y_{\pi_n} | \theta) p(\theta) d(\theta) \\ &= p(y_{\pi_1}, \dots, y_{\pi_n}) \end{aligned}$$

Where:

- The first line is the definition of marginal probability
- The second line is because the  $Y_i$ 's are conditionally *iid*
- The third line is because the product does not depend on order
- The fourth line we are converting back to the form used in the first line
- Finally, the last line is the definition of marginal probability

## Problem 2

### Part a

Prior Density:  $p(\theta) = 1$

Sampling Distribution:  $p(y|\theta) = \binom{n}{y} \theta^y (1 - \theta)^{n-y}$

Posterior Density:  $p(\theta|y) \propto p(y|\theta)p(\theta) = \binom{n}{y} \theta^y (1 - \theta)^{n-y}$

Posterior Distribution:  $\theta|y \sim \text{Beta}(y + 1, (n - y) + 1)$

$$\begin{aligned}
\binom{n}{y} \frac{\Gamma(y+1)\Gamma(n-y+1)}{\Gamma((y+1)+(n-y+1))} &= \binom{n}{y} \frac{\Gamma(y+1)\Gamma(n-y+1)}{\Gamma(n+2)} \\
&= \frac{n!}{y!(n-y)!} \frac{y!(n-y)!}{(n+1)!} \\
&= \frac{n!}{(n+1)!} \\
&= \frac{1}{n+1}
\end{aligned}$$

## Part b

Prior Distribution:  $\theta \sim \text{Beta}(\alpha, \beta)$

Sampling Distribution:  $y|\theta \sim \text{Binomial}(n, \theta)$

Posterior Density:

$$\begin{aligned}
p(\theta|y) &= \binom{n}{y} \theta^y (1-\theta)^{n-y} \times \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1} \\
&= \binom{n}{y} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{y+\alpha-1} (1-\theta)^{n-y+\beta-1} \\
&\propto \text{Beta}(y+\alpha, (n-y)+\beta)
\end{aligned}$$

Posterior Mean:

$$\begin{aligned}
\frac{y+\alpha}{\alpha+\beta+n} &= \frac{y}{n} + \lambda \left( \frac{\alpha}{\alpha+\beta} - \frac{y}{n} \right) \\
\frac{y+\alpha}{\alpha+\beta+n} - \frac{y}{n} &= \lambda \left( \frac{\alpha}{\alpha+\beta} - \frac{y}{n} \right) \\
\frac{ny+n\alpha-\alpha y-\beta y-ny}{(\alpha+\beta+n)n} &= \lambda \left( \frac{n\alpha-\alpha y-\beta y}{(\alpha+\beta)n} \right) \\
\lambda &= \frac{\alpha+\beta}{\alpha+\beta+n}
\end{aligned}$$

Since  $\lambda$  will always be between 0 and 1 the Posterior Mean will act as a weighted average between our Prior Mean,  $\frac{y}{n}$ , and the data.

## Part c

Posterior Density:  $\theta|y \sim \text{Beta}(y+\alpha, (n-y)+\beta)$

Prior Variance ( $\alpha=1, \beta=1$ ):  $\frac{1}{12}$

Posterior Variance ( $\alpha=1, \beta=1$ ):

$$\begin{aligned}
\frac{(y+1)(n-y+1)}{(n+2)^2(n+3)} &= \frac{ny-y^2+y+n-y+1}{(n^2+4n+4)(n+3)} \\
&= \frac{ny-y^2+y+n-y+1}{n^3+3n^2+4n^2+12n+4n+12} \\
&= \frac{-y^2+ny+n+1}{n^3+7n^2+16n+12}
\end{aligned}$$

Now, we can deduce that smaller values of  $n$  will maximize this quantity; since  $n \geq y$ , we will set  $n = y = 1$ .

Posterior Variance ( $n = y = 1$ ):  $\frac{-y^2+ny+n+1}{n^3+7n^2+16n+12} = \frac{2}{36} = \frac{1}{18}$

Thus, the Posterior Variance, which we just maximized, is always less than the Prior Variance of  $\frac{1}{12}$ .

## Part d

- $n = y = 1$
- $\alpha = 1$
- $\beta = 10$

Prior Variance:  $\frac{(1)(10)}{(11)^2(12)} = \frac{10}{1452} = 0.0069$

Posterior Variance:  $\frac{(2)(10)}{(12)^2(13)} = \frac{20}{1872} = 0.0107$

## Problem 3

### Part a

Prior Distribution:  $\theta \sim \text{Beta}(\alpha, \beta)$

Prior Mean: 0.6

Prior Variance: 0.09

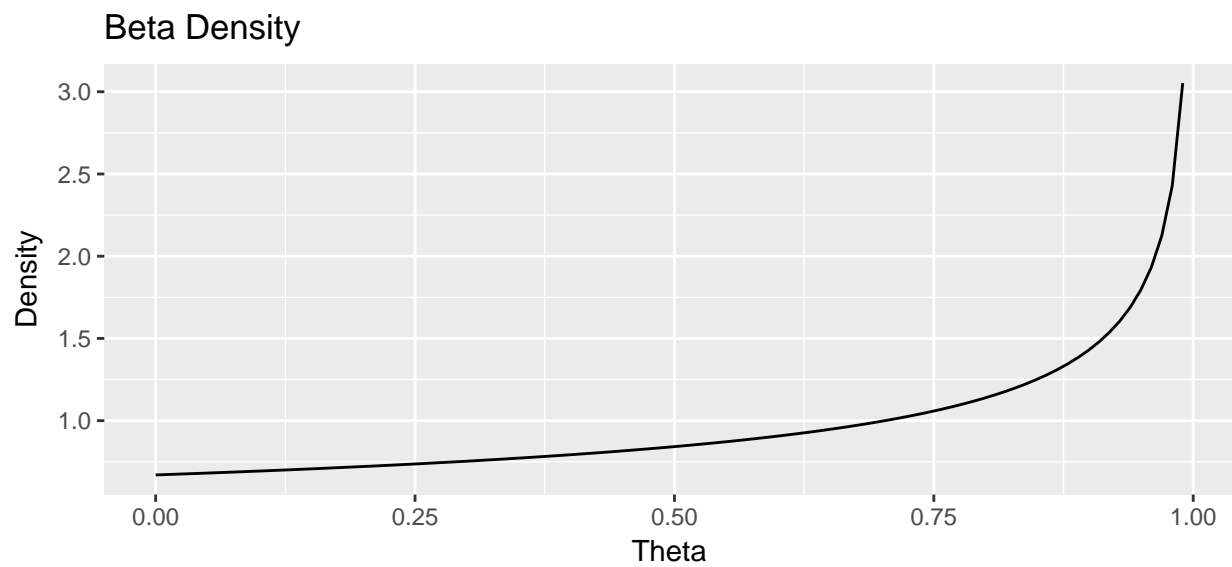
Sampling Distribution:  $y|\theta \sim \text{Binomial}(n, \theta)$

$$\begin{aligned}\mu &= \frac{\alpha}{\alpha + \beta} \\ (\alpha + \beta)\mu &= \alpha \\ \alpha\mu + \beta\mu &= \alpha \\ \beta\mu &= \alpha - \alpha\mu \\ \beta &= \alpha\left(\frac{1}{\mu} - 1\right)\end{aligned}$$

$$\begin{aligned}\sigma^2 &= \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \\ \alpha &= \left(\frac{1 - \mu}{\sigma^2} - \frac{1}{\mu}\right)\mu^2\end{aligned}$$

$\alpha = 1$

$\beta = 0.67$



### Part b

$$n = 1000$$

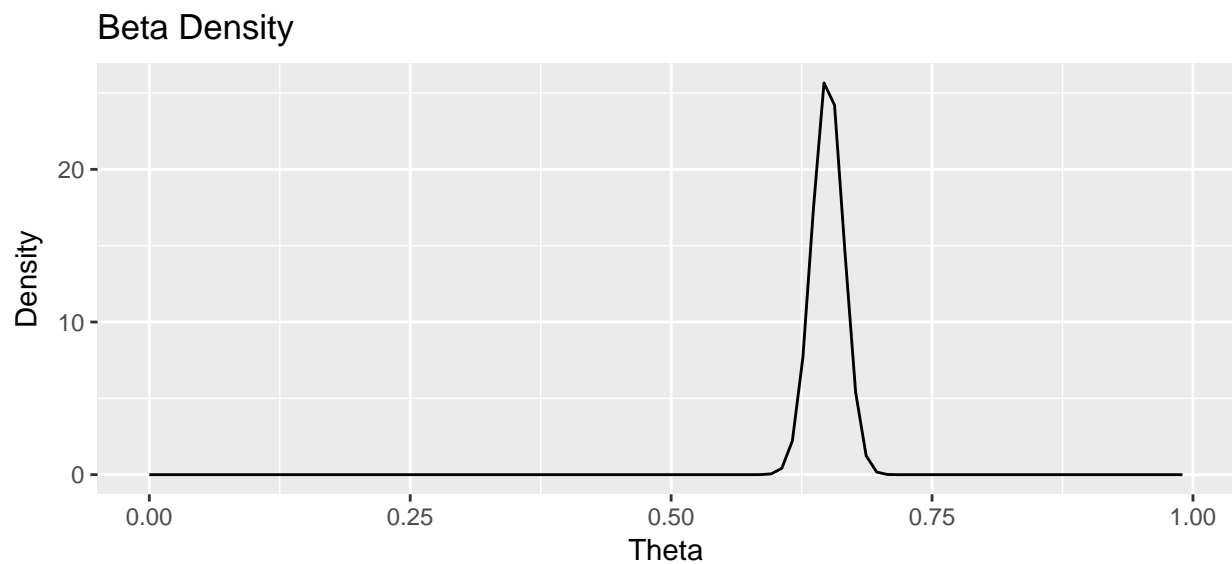
$$y = 650$$

$$\alpha_{post} = y + \alpha$$

$$\beta_{post} = (n - y) + \beta$$

$$\text{Posterior Mean: } \mu_{post} = \frac{y + \alpha}{(y + \alpha) + ((n - y) + \beta)} = 0.6499$$

$$\text{Posterior Variance: } \sigma_{post}^2 = \frac{(y + \alpha)((n - y) + \beta)}{((y + \alpha) + ((n - y) + \beta))^2 ((y + \alpha) + ((n - y) + \beta) + 1)} = 0.0151$$

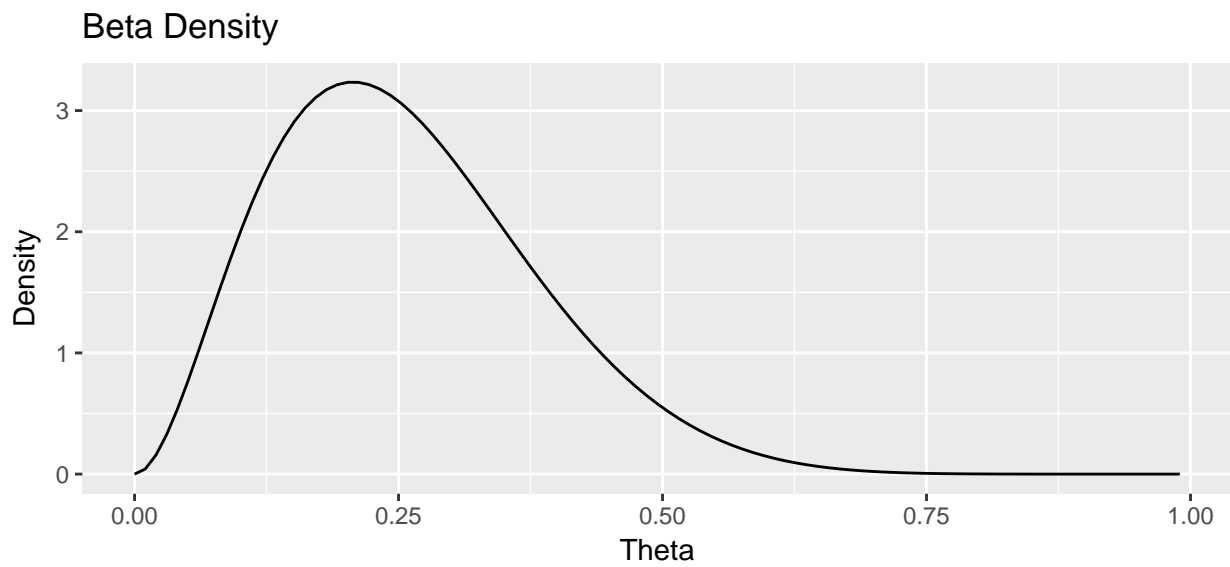


## Part c

### First Sensitivity Check

$$n = 10$$

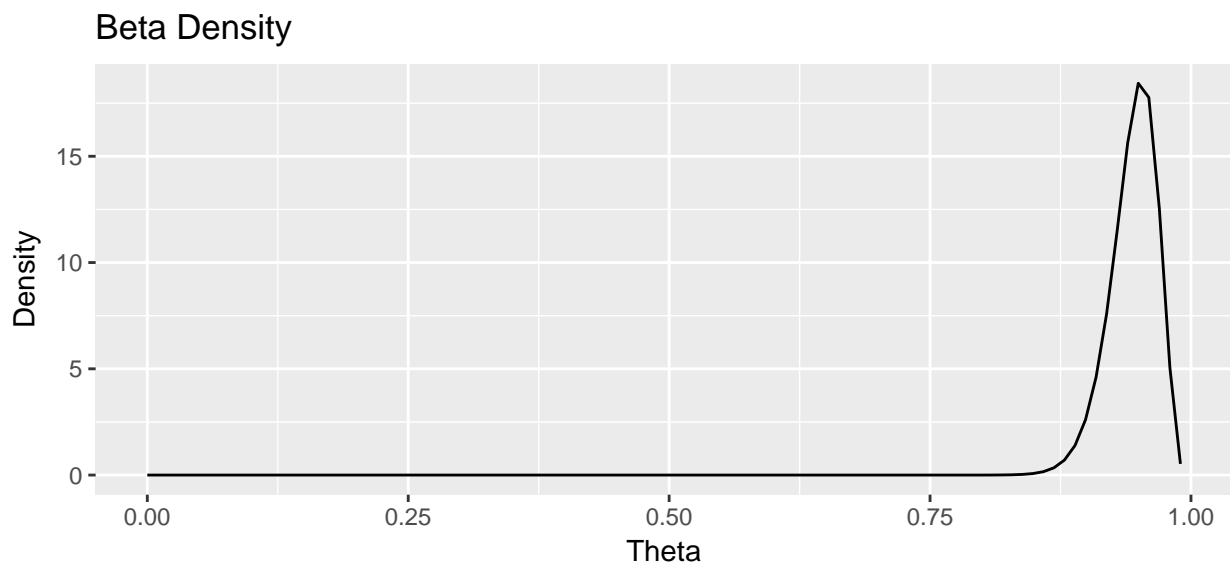
$$y = 2$$



### Second Sensitivity Check

$$n = 100$$

$$y = 95$$



### Third Sensitivity Check

$n = 500$

$y = 250$

