Assignment #2

Elliot Smith

8/30/2018

Problem 1

Part a

Here our goal will be to minimize a to show that $a = E[\theta|y]$ is the unique Bayes estimate of θ:

$$\begin{aligned}
\frac{d}{da}E[L(a|y)] &= \frac{d}{da}\int L(\theta,a)p(\theta|y)d\theta \\
&= \frac{d}{da}\int (\theta-a)^2 p(\theta|y)d\theta \\
&= -2\int (\theta-a)p(\theta|y)d\theta \\
&= -2\Big[\int \theta p(\theta|y)d\theta - a\int p(\theta|y)d\theta\Big] \\
&= -2\big[E[\theta|y]-a\big]
\end{aligned}$$

$-2\big[E[\theta|y]-a\big] = 0$ when $a = E[\theta|y]$

To prove that it is a unique minimizing statistic, we must look at the second derivative:

$\frac{d}{da}(-2\big[E[\theta|y]-a\big]) = 2$

As $2 > 0$, this shows that it is a unique minimzing statistic.

Part b

Here our goal will be to show that for any median value of a, the derivative of $L(\theta, a)$ will evaluate to 0.

$$\begin{aligned}
\frac{d}{da}\big[E[L(a|y)]\big] &= \frac{d}{da}\Big[\int_{-\infty}^{a}(a-\theta)p(\theta|y)d\theta + \int_{a}^{\infty}(\theta-a)p(\theta|y)d\theta\Big] \\
&= \int_{-\infty}^{a}\frac{d}{da}(a-\theta)p(\theta|y)d\theta + \int_{a}^{\infty}\frac{d}{da}(\theta-a)p(\theta|y)d\theta \\
&= \int_{-\infty}^{a}p(\theta|y)d\theta + \int_{a}^{\infty}(-1)p(\theta|y)d\theta \\
&= \int_{-\infty}^{a}p(\theta|y)d\theta - \int_{a}^{\infty}p(\theta|y)d\theta \\
&= \frac{1}{2} - \frac{1}{2} \\
&= 0
\end{aligned}$$

As a result, it has been shown that any posterior median of θ is a Bayes estimate of θ.

Taking the second derivative we again get a positive number, thus again indicating that it is a minimizing statistic.

Part c

Here our goal will be to show that for any value of a, the derivative of $L(\theta, a)$ will evaluate to 0 where k_0 and k_1 are nonnegative numbers.

$$\begin{aligned}
\frac{d}{da}\big[E[L(a|y)]\big] &= \frac{d}{da}\left[\int_{-\infty}^{a} k_1(a-\theta)p(\theta|y)d\theta + \int_{a}^{\infty} k_0(\theta-a)p(\theta|y)d\theta\right] \\
&= \int_{-\infty}^{a} \frac{d}{da}k_1(a-\theta)p(\theta|y)d\theta + \int_{a}^{\infty} \frac{d}{da}k_0(\theta-a)p(\theta|y)d\theta \\
&= \int_{-\infty}^{a} k_1 p(\theta|y)d\theta + \int_{a}^{\infty} (-k_0)p(\theta|y)d\theta \\
&= \int_{-\infty}^{a} k_1 p(\theta|y)d\theta - \int_{a}^{\infty} k_0 p(\theta|y)d\theta \\
&= k_1\int_{-\infty}^{a} p(\theta|y)d\theta - k_0\int_{a}^{\infty} p(\theta|y)d\theta
\end{aligned}$$

Noting that: $k_0\int_a^\infty p(\theta|y)d\theta = k_0 - k_0\int_{-\infty}^{a} p(\theta|y)d\theta$

$$\begin{aligned}
k_1\int_{-\infty}^{a} p(\theta|y)d\theta - k_0\int_{a}^{\infty} p(\theta|y)d\theta &= k_1\int_{-\infty}^{a} p(\theta|y)d\theta - \left[k_0 - k_0\int_{-\infty}^{a} p(\theta|y)d\theta\right] \\
&= k_1\int_{-\infty}^{a} p(\theta|y)d\theta + k_0\int_{-\infty}^{a} p(\theta|y)d\theta - k_0 \\
&= (k_1+k_0)\int_{-\infty}^{a} p(\theta|y)d\theta - k_0
\end{aligned}$$

Now setting $\int_{-\infty}^{a} p(\theta|y)d\theta = \frac{k_0}{k_0+k_1}$ we get our result that any quantile is a Bayes estimate of θ.

Taking the second derivative we again get a positive number, thus again indicating that it is a minimizing statistic.

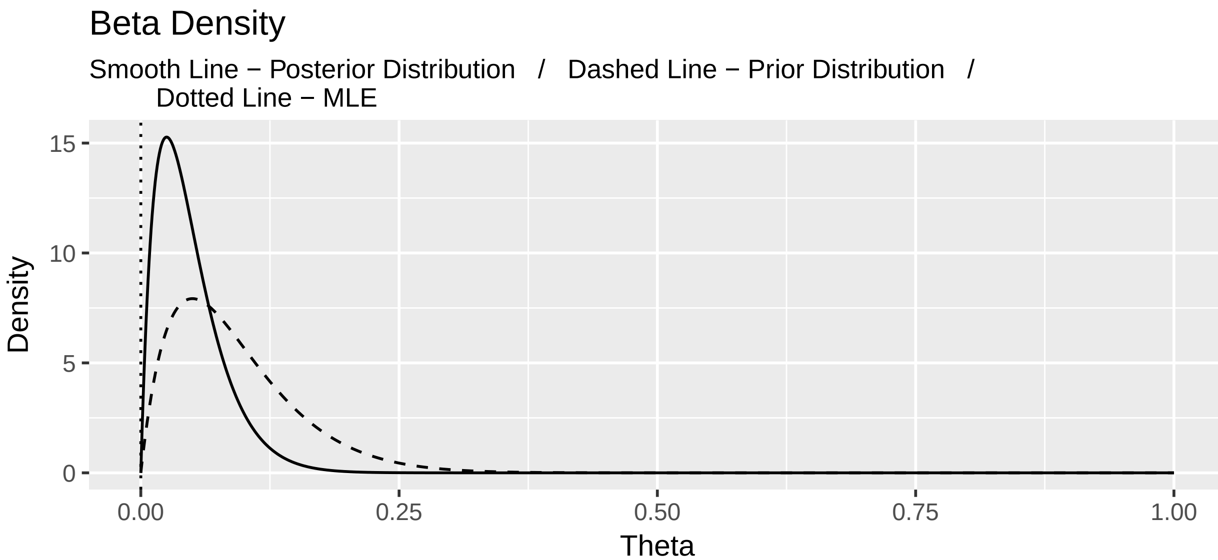
Problem 2

n = 20
Sampling Distribution: $y|\theta \sim Binomial(n = 20, \theta)$
Prior Distribution: $\theta \sim Beta(\alpha = 2, \beta = 20)$
Posterior Distribution:

$$\begin{aligned}
p(\theta|y) &= \binom{n}{y}\theta^y(1-\theta)^{n-y} \times \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1} \\
&= \binom{n}{y}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{y+\alpha-1}(1-\theta)^{n-y+\beta-1} \\
&\propto Beta(y+\alpha, (n-y)+\beta) \\
&= Beta(y+2, (20-y)+20)
\end{aligned}$$

Part i

$y = 0$



From this graph we can see that while the MLE, $\hat{y} = \frac{y}{n}$ is not properly represented the true distribution due to the fact that there were no positive cases in this hospital, the posterior distribution, via the affect of the prior distribution, is impacting our results enough to render a more accurate result.

Part ii

```
## HPD Interval:  ( 0.0006780629 , 0.1087375 )
```

```
## Frequentist Interval:  ( 0 , 0 )
```

From this result, we can see that the Frequestist result compares poorly ($(0, 0)$ is not really an interval at all) against the Bayesian posterior credible interval of θ. Again, the frequentist approach for this hospital is being adversely affected by the fact that there are no positive cases, and as such, we do not get a good representation of the true distribution of positive cases. Whereas the Bayesian approach factors in our prior beliefs on the number of positive cases, thus allowing us to interpret the results for this case.

Part iii

```
## HPD Interval:  ( -4.897083 , -1.763996 )
```

Problem 3

Part i

My methodology for selecting my α and β for my $\theta_1, \theta_2 \sim Gamma(\alpha, \beta)$ distribution is as follows. First, using a Pew Research Center on Social & Demographic Trends article on average number of children per parent (http://www.pewsocialtrends.org/2015/05/07/family-size-among-mothers/), I deduced that the average number of children per parent for women in their 20s during the 1970s was 2.28, while the variance is 0.8905.

These values were obtained by the results from the poll. Now that I had values for the prior mean and variance, I was able to algebraically solve for α and β via the following formulae: $EX = \frac{\alpha}{\beta}$ and $VarX = \frac{\alpha}{\beta^2}$.

```
## Gamma Parameters

## Alpha Parameter:  5.8376

## Beta Parameter:  2.5604
```

Part ii

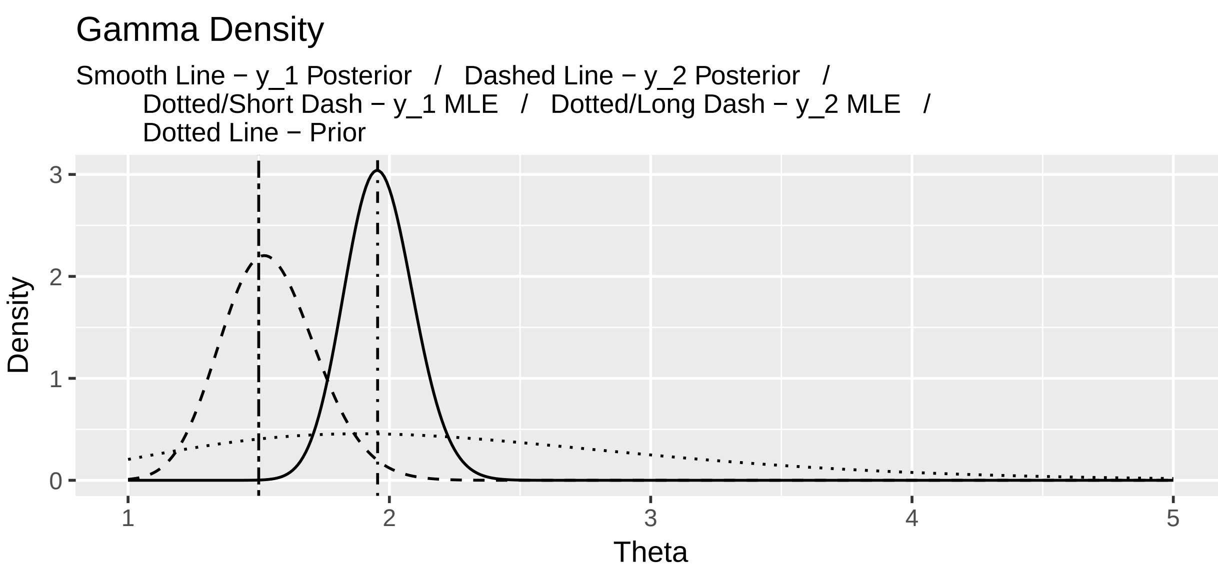
Posterior Distributions:

$$
\begin{aligned}
p(\theta|y) &= p(\theta_1, \theta_2|y_1, y_2) \\
&\propto p(y_1|\theta_1)p(y_2|\theta_2)p(\theta_1)p(\theta_2) \\
&= \frac{{\theta_1}^{y_1} e^{-\theta_1}}{y_1!} \frac{{\theta_2}^{y_2} e^{-\theta_2}}{y_2!} \frac{\beta^\alpha}{\Gamma(\alpha)} {\theta_1}^{\alpha-1} e^{-\beta\theta_1} \frac{\beta^\alpha}{\Gamma(\alpha)} {\theta_2}^{\alpha-1} e^{-\beta\theta_2} \\
&\propto {\theta_1}^{y_1} e^{-\theta_1} {\theta_2}^{y_2} e^{-\theta_2} {\theta_1}^{\alpha-1} e^{-\beta\theta_1} {\theta_2}^{\alpha-1} e^{-\beta\theta_2} \\
&= \theta_1^{y_1+\alpha-1} \theta_2^{y_2+\alpha-1} e^{-[\theta_1(1+\beta)+\theta_2(1+\beta)]}
\end{aligned}
$$

Since we may assume *iid* priors, we get:

$p(\theta_1|y_1) \sim Gamma(y_1 + \alpha, n_1 + \beta)$

$p(\theta_2|y_2) \sim Gamma(y_2 + \alpha, n_2 + \beta)$



As we can see from the plot, we have a fairly non-informative prior; its very hard to tell whether or not our prior distribution is influencing either of our posterior distributions in a meaningful way. Our MLE mean results appear to align very well with our posterior distributions.

Part iii

```
## Posterior Credible Interval, theta_1:  ( 1.714745 , 2.215768 )

## Posterior Credible Interval, theta_2:  ( 1.200902 , 1.918022 )
```

As we can see from our results, the two intervals do overlap!

Part iv

```
## Posterior Credible Interval, theta_1 - theta_2:  ( -0.008211997 , 0.8591994 )
```

As we can see from our results, the interval of $\theta_1 - \theta_2$ does include zero!

Part v

```
## Posterior Probability:  0.9686
```

This provided value dictates the likelihood that a value falls within a particular range in the distribution. In this case, the provided value states how likely that the true answer lies in the range, which in this case, is greater than zero.

The difference between the classical Frequensist approach of a hypothesis test and our Bayesian approach is that the Frequentists believe that our statistic in question is an unknown, but fixed, number, whereas Bayesians believes that it is an unknown distribution. As a result, the Bayesian approach allows us to say whether or not a true answer lies in a given range, while the Frequentist approach relies heavily on weakly defined metrics ("data more extreme", p-value commonly misinterpreted) to determine the probability that we are making an error when accepting our Null Hypothesis as true.