EN.553.732 Homework 1

Joseph High

September 19, 2017

1 Problem 1

Let $p(y_1, ..., y_6|\theta)$ denote the sampling model and $\pi(\theta)$ denote the prior distribution. It is given that $y_1, ..., y_6|\theta \sim U[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$ and that $\theta \sim U[10, 20]$. We then have that
$p(y_1, ..., y_6|\theta) = \dfrac{1}{(\theta + \frac{1}{2}) - (\theta - \frac{1}{2})} = 1, \;\; y_i \in (\theta - \frac{1}{2}, \theta + \frac{1}{2}), \; i = 1, ..., 6$

and $\pi(\theta) = \frac{1}{10}, \;\; \theta \in [10, 20]$

Note that $\theta - \frac{1}{2} < y_i < \theta + \frac{1}{2} \implies |y - \theta| < \frac{1}{2} \implies |\theta - y_i| < \frac{1}{2} \implies y_i - \frac{1}{2} < \theta < y_i + \frac{1}{2}, \; i = 1, ..., 6$

Using the available data $y_i \in \{11.0, 11.5, 11.7, 11.1, 11.4, 10.9\}$, we can find updated bounds for θ.

Indeed, since the above inequality holds for all $i = 1, ..., 6$, then $\max_i\{y_i\} - \frac{1}{2} < \theta < \min_i\{y_i\} + \frac{1}{2}$

$\implies 11.7 - \frac{1}{2} < \theta < 10.9 + \frac{1}{2} \implies 11.2 < \theta < 11.4$

The posterior distribution of θ is $p(\theta|y_1, ..., y_6) \propto p(y_1, ..., y_6|\theta)\pi(\theta)$, implying that the posterior distribution is also uniformly distributed. More precisely, by Bayes' rule

$$p(\theta|y_1, ..., y_6) = \frac{p(y_1, ..., y_6|\theta)\pi(\theta)}{p(y_1, ..., y_6)} = \frac{(1/10)}{p(y_1, ..., y_6)}, \quad 11.2 < \theta < 11.4$$

Now,
$$1 = \int_{11.2}^{11.4} p(\theta|y_1, ..., y_6)d\theta = \int_{11.2}^{11.4} \frac{(1/10)}{p(y_1, ..., y_6)}d\theta = \frac{1}{p(y_1, ..., y_6)} \int_{11.2}^{11.4} \frac{1}{10}d\theta = \frac{0.2/10}{p(y_1, ..., y_6)}$$

$$\implies \frac{1}{p(y_1, ..., y_6)} = 50$$

Thus,
$$p(\theta|y_1, ..., y_6) = \frac{p(y_1, ..., y_6|\theta)\pi(\theta)}{p(y_1, ..., y_6)} = (\frac{1}{10})(50) = 5$$

Therefore, the posterior distribution of θ is $p(\theta|y_1, ..., y_6) = 5, \; 11.2 < \theta < 11.4$.

2 Problem 2

Let $p(y_1, ..., y_{20}|\theta)$ denote the sampling model and $\pi(\theta)$ denote the prior distribution. It is given that $y_1, ..., y_{20}|\theta \sim$ Exponential(θ) and that $\theta \sim$ Gamma(α, β) where E[θ] = 0.2 and $\sigma[\theta] = 1$.

The liklihood function is then $L(\theta) = \theta^{20} e^{-(\sum_{i=1}^{20} y_i)\theta}$ and prior distribution is of the form $\pi(\theta) \propto \theta^{\alpha-1}e^{-\beta\theta}, \; \theta > 0$

The posterior of θ is of then of the form

$$p(\theta|y_1,...,y_{20}) \propto L(\theta)\pi(\theta) \propto \theta^{\alpha+19}e^{-(\sum_{i=1}^{20} y_i+\beta)\theta}, \quad \theta > 0$$

which is proportional to the pdf of a gamma distribution with parameters $\alpha + 20$ and $\beta + \sum\limits_{i=1}^{20} y_i$

To determine the paramaters α and β we note that since $\theta \sim \text{Gamma}(\alpha,\, \beta)$, $\;\text{E}[\theta] = \dfrac{\alpha}{\beta}$ and $\sigma[\theta] = \sqrt{\dfrac{\alpha}{\beta^2}}$

Then from the given information, we have that $\dfrac{\alpha}{\beta} = 0.2$ and $\sqrt{\dfrac{\alpha}{\beta^2}} = 1 \implies \alpha = 0.2\beta \implies \frac{0.2}{\beta} = 1$ $\implies \beta = 0.2 \implies \alpha = 0.04$.

Moreover, we are given that the average time to serve a customer from a sample of 20 customers is 3.8 minutes. That is, $\bar{y} = \frac{1}{20}\sum\limits_{i=1}^{20} y_i = 3.8 \implies \sum\limits_{i=1}^{20} y_i = 76$.

Thus, we have that the parameters of the posterior distribution are $\alpha + 20 = 20.04$ and $\beta + \sum\limits_{i=1}^{20} y_i = 76.2$
Therefore, $\theta|y_1,...,y_{20} \sim \text{Gamma}(20.04,\, 76.2)$.

3 Problem 3

Proof. It is given that the sampling model, $f(y_1,...,y_n|\, p)$, has the negative binomial distribution with unknown parameter p (and known r) and the prior distribution, $\pi(p)$, is the beta distribution with parameters α and β. That is,

$$L(p) = \prod_i^n f(y_i\;|p) = \prod_i^n \binom{y_i + r - 1}{y_i}(1-p)^r p^{y_i} \implies L(p) \propto (1-p)^{nr}p^{\sum_i^n y_i}$$

$$\pi(p) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1} \implies \pi(p) \propto p^{\alpha-1}(1-p)^{\beta-1}$$

Therefore, the posterior distribution is of the form

$$f(p\;|y_1,...,y_n) \propto\; L(p)\pi(p) \propto\; p^{\alpha+\sum_i^n y_i - 1}(1-p)^{\beta+nr-1} \longrightarrow \text{Beta}(\alpha + \sum_i^n y_i,\;\beta + nr)$$

Thus, the posterior distribution, again, follows a beta distribution with parameters $\alpha + \sum_i^n y_i$ and $\beta + nr$
Therefore, the family of beta distributions is a conjugate family of prior distributions for samples from a negative binomial distribution with one unknown parameter. □

4 Problem 4

Joseph Hugh

Rough Draft (To Be input into LaTeX)

Problem 4:

1.) For each $i \in [1, 100]$, $\Pr(Y_i = y_i \mid \theta) = \theta^{y_i}(1-\theta)^{1-y_i}$ since Y_i binary random variables.
- Then, with the assumption of conditional independence of the Y_i on θ, we have

$$\Pr(Y_1 = y_1, \ldots, Y_{100} = y_{100} \mid \theta) = \prod_{i=1}^{100} \Pr(Y_i = y_i \mid \theta)$$

$$= \prod_{i=1}^{100} \theta^{y_i}(1-\theta)^{1-y_i}$$

$$= \theta^{\sum_{i=1}^{100} y_i}(1-\theta)^{100-\sum_{i=1}^{100} y_i}$$

Then $\Pr\left(\sum_{i=1}^{100} Y_i = y \mid \theta\right)$ is the probability that the sum of the binary random variables is equal to y, where the sum of y of 100 binary random variables can be achieved in $\binom{100}{y}$ distinct ways.

Thus, $\Pr\left(\sum_{i=1}^{100} Y_i = y \mid \theta\right) = \binom{100}{y}\theta^{y}(1-\theta)^{100-y}$.

$\square$

2.) (Problem 4)

* Compute $Pr(\sum_{i=1}^{100} y_i = 57 | \theta)$ for each θ.

* Refer to R Code for computation of $Pr(\sum_{i=1}^{100} y_i = 57 | \theta)$ for each θ.

- From part 1 we know that $Pr(\sum_{i=1}^{100} y_i = 57 | \theta) = \binom{100}{57} \theta^{57} (1-\theta)^{43}$

θ	$Pr(\sum_{i=1}^{100} y_i = 57 \| \theta)$
0.0	0.00
0.1	4.107157×10^{-31}
0.2	3.738459×10^{-16}
0.3	1.306895×10^{-8}
0.4	2.285792×10^{-4}
0.5	3.006864×10^{-2}
0.6	6.67289×10^{-2}
0.7	1.853172×10^{-3}
0.8	1.003535×10^{-7}
0.9	9.395858×10^{-18}
1.0	0.00

Plot is attached at end of document

(3) Since $Pr(\theta=0.0) = Pr(\theta=0.1) = \cdots = Pr(\theta=1.0)$

Then $Pr(\theta=0.0) = \cdots = Pr(\theta=1.0) = 1/11$

Since $\theta \sim$ Discrete

Then $$p(\theta \mid \sum_{i=1}^{100} y_i = 57) = \frac{Pr(\sum_{i=1}^{100} y_i = 57 \mid \theta)\, Pr(\theta)}{Pr(\sum_{i=1}^{100} y_i = 57)}$$

$$= \frac{\binom{100}{57} \theta^{57}(1-\theta)^{43} \left(\frac{1}{11}\right)}{Pr(\sum_{i=1}^{100} y_i = 57)}$$

* Where $1/Pr(\sum y_i = 57)$ is just a normalization constant

Then $$Pr(\sum y_i = 57) = \sum_{\theta} \binom{100}{57} \theta^{57}(1-\theta)^{43}\left(\frac{1}{11}\right)$$

From the R code, the values of $p(\theta \mid \sum_{i=1}^{100} y_i = 57)$ were found to be.

θ	0.0	0.1	0.2	0.3
$p(\theta \mid \sum y_i = 57)$	0.00	4.153701×10^{-30}	3.780824×10^{-15}	1.321705×10^{-7}

0.4	0.5	0.6	0.7
2.311695×10^{-3}	3.040939×10^{-1}	6.748515×10^{-01}	1.874172×10^{-2}

0.8	0.9	1.0
1.014907×10^{-6}	9.502335×10^{-17}	0.00

Plot attached.

(4) See Plot and R code.

Here,

$$p(\theta \mid \Sigma y_i = 57) = \frac{\binom{100}{57}\theta^{57}(1-\theta)^{43} \times (1)}{Pr(\Sigma y_i = 57)}$$

for $0 \le \theta \le 1$.

Plot attached to document.

---

(5) Posterior Distribution $\propto$ Beta(58, 44)

Here, there is no normalizing constant since it Beta(58, 44) is the proportional distribution to the posterior.

See plot and R code attached.

Relationship: The plot for the sampling model in Part (2) the discrete type; indeed, it is a discrete distribution.

In Part (3) our prior was discrete, and our posterior was also discrete and was proportional to the distribution of the sampling model.

In parts (4) and (5) our prior was continuous resulting in a continuous posterior, but remained proportional to the sampling model.

Problem 5: See R Code.

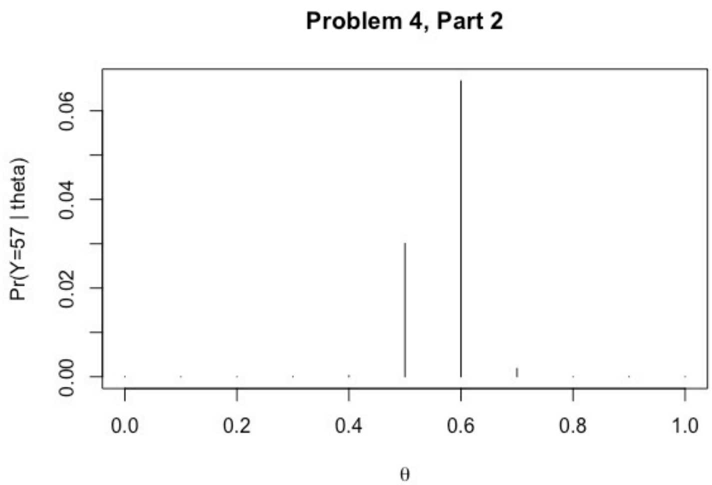
According to Hoff page 6, Chapter 1,

The contour plot suggests that "lower values of n_0 are generally 90% or more certain that" $\theta > 0.5$. — Reference: A First Course in Bayesian statistical methods. Hoff, Chapter 1.

Joseph High
Hopkins ID: 9E1FDC

Problem 4, Part 2 R Code and Plot

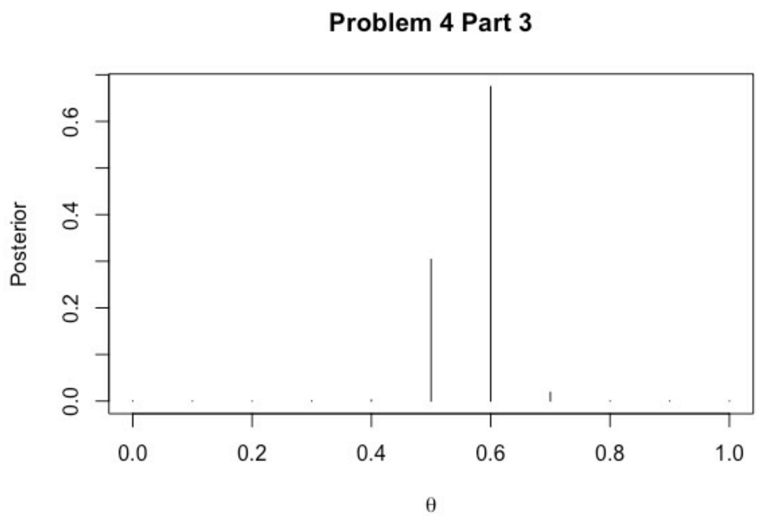
```
> theta<-c(0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0)
> Y<-rep(1,11)
> for(i in 1:11)
+   Y[i]<-choose(100,57)*(theta[i]^57)*(1-theta[i])^43
> print(Y)
 [1] 0.000000e+00 4.107157e-31 3.738459e-16 1.306895e-08 2.285792e-04
3.006864e-02
 [7] 6.672895e-02 1.853172e-03 1.003535e-07 9.395858e-18 0.000000e+00
>
> plot(theta, Y, type = "h", main = "Problem 4, Part 2", xlab =
expression(paste(theta)),
+      ylab="Pr(Y=57 | theta)")
```



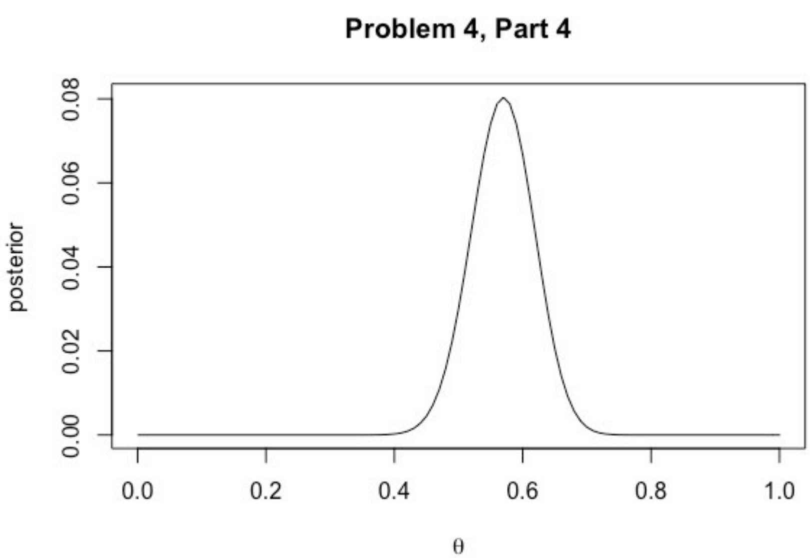
Problem 4, Part 3 R Code and Plot

```
> x<-rep(1,11)
> x1<-rep(1,11)
> for(i in 1:11)
+   x[i]<-(choose(100,57)*(theta[i]^57)*(1-theta[i])^43)*(1/11)
> NormConstant<-1/(sum(x))
> x1<-x*NormConstant
> print(x1)
 [1] 0.000000e+00 4.153701e-30 3.780824e-15 1.321705e-07 2.311695e-03
3.040939e-01
 [7] 6.748515e-01 1.874172e-02 1.014907e-06 9.502335e-17 0.000000e+00
>
> plot(theta, x1, type = "h", main = "Problem 4 Part 3", xlab =
expression(paste(theta)),
+      ylab = "Posterior")
```

Joseph High
Hopkins ID: 9E1FDC

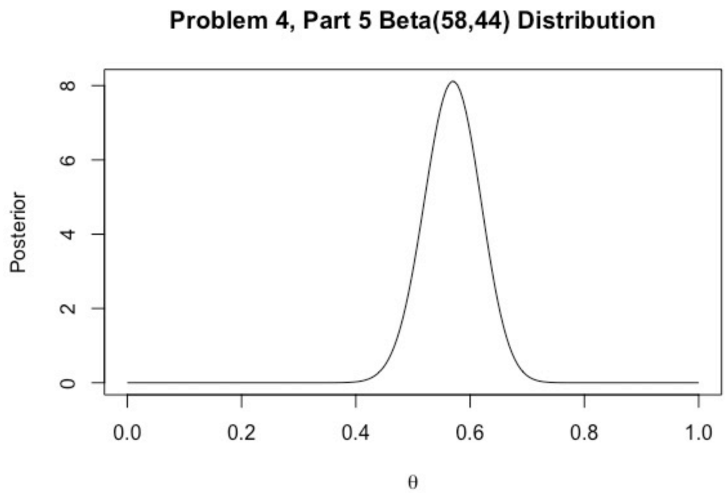


Problem 4, Part 4 R Code and Plot
```
f<-curve(choose(100,57)*x^57*((1-x)^43), from = 0, to = 1, main = "Problem 4, 
Part 4",
         xlab = expression(paste(theta)), ylab = "posterior")
```



Problem 4, Part 5 R Code and Plot
```
v<-seq(0, 1, length = 200)
z<-dbeta(v, 58, 44)
plot(v, z, type = "l", main = "Problem 4, Part 5 Beta(58,44) Distribution",
     xlab = expression(paste(theta)), ylab = "Posterior")
```

Joseph High
Hopkins ID: 9E1FDC



Joseph High
Hopkins ID: 9E1FDC

Problem 5 R Code and Plot

```
theta_0<-c(0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9)
n_0<-c(1,2,8,16,32)

a<-matrix(0L, nrow =length(theta_0), ncol =length(n_0))
b<-matrix(0L, nrow =length(theta_0), ncol =length(n_0))
for (i in 1:length(theta_0))
{for (j in 1:length(n_0))
{a[i,j]=theta_0[i]*n_0[j]
b[i,j]=(1-theta_0[i])*n_0[j]
}
}
Pr<-matrix(0L, nrow =length(theta_0), ncol =length(n_0))
for (i in 1:length(theta_0))
{for (j in 1:length(n_0))
{
  f <- function(x)
    {choose(100,57)*(x^57)*((1-
x)^43)*(gamma(a[i,j]+b[i,j])/(gamma(a[i,j])*gamma(b[i,j])))*(x^(a[ i,j]-
1))*(1-x)^(b[i,j]-1)}
  bot<-integrate(f,0, 1, rel.tol=1e-10)$value
  top<-integrate(f,0.5, 1, rel.tol=1e-10)$value
  Pr[i,j]<-top/bot
}
}
contour(theta_0, n_0, Pr,main = "Problem 5 Countour Plot",
xlab=expression(paste(theta)),
        ylab='n_0')
```