Homework 1

computational

Michael Pena

Question 2

(a).

X_1	$P(X_1)$
0	$.10 + .15 = .25$
1	$.2 + .1 = .3$
2	$.15 + .3 = .45$

(b).

X_2	$P(X_2)$
0	$.1 + .2 + .15 = .45$
1	$.15 + .1 + .3 = .55$

(c).

X_1	X_2	$P(X_1 \| X_2)$
0	0	2/9
1	0	4/9
2	0	1/3
0	1	0.27272
1	1	0.1818
2	1	0.5454

(d).

X_1	X_2	$P(X_2\|X_1)$
0	0	0.4
1	0	2/3
2	0	1/3
0	1	0.6
1	1	1/3
2	1	2/3

(e)

It's not independent because $P(X_2 = 0|X_1 = 0) = 0.4$ and $P(X_2 = 0) = 0.45$ which are not the same.

Question 3

```
yi = c(1.8,1.7,1.4,1.6,1.9,1.5,1.2)
```

(a).

$$\frac{\partial}{\partial\beta}\sum_{i=1}^{7} ln(\frac{\beta^4}{6}y_i^3 e^{-\beta y_i}) = \sum_{i=1}^{7}\left(\frac{\beta^4}{6}y_i^3 e^{-\beta y_i}\right)^{-1}\cdot y_i^3\left(\frac{2\beta^3}{3}e^{-\beta y_i} - \frac{\beta^4}{6}y_i e^{-\beta y_i}\right) = \sum_{i=1}^{7}\frac{6}{\beta}(\frac{2}{3} - \frac{\beta y_i}{6}) = \frac{4(7)}{\beta} - \sum_{i=1}^{7} y_i$$

this implies that $\tilde{\beta} = \frac{28}{\sum_{i=1}^{7} y_i} = 2.522523$

(b).

solving for $\tilde{\alpha}$

$$\frac{\partial}{\partial\alpha}\left[\sum_{i=1}^{7}\Big(\alpha ln2 - ln(\Gamma(\alpha)) + (\alpha-1)ln(y_i) - 2y_i\Big)\right] = \sum_{i=1}^{7}\Big(ln2 - \frac{\Gamma(\alpha)\psi(\alpha)}{\Gamma(\alpha)} + ln(y_i)\Big) = 7ln2 - \psi(\alpha) + \sum_{i=1}^{7} ln(y_i)$$

so if we set the derivative to zero, it implies

$$\psi(\tilde{\alpha}) = 7ln2 + \sum_{i=1}^{7} ln(y_i)$$

furthermore

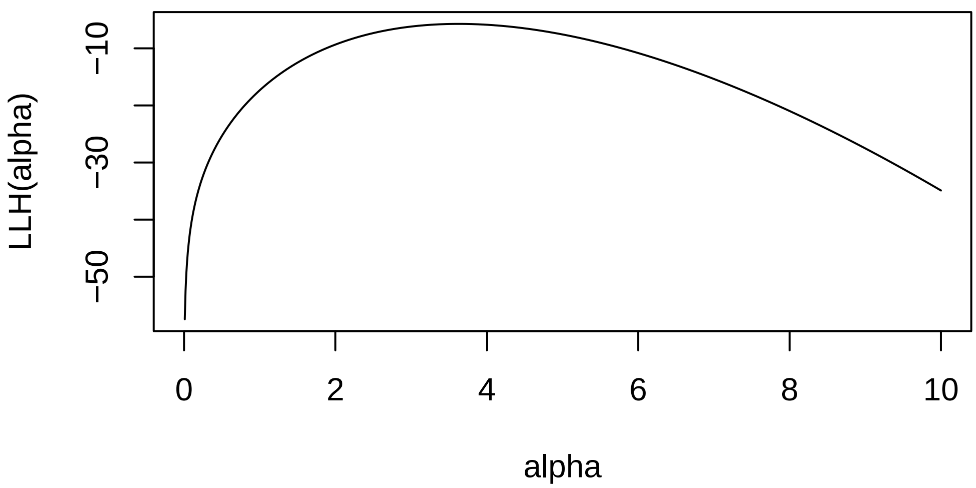
$$\tilde{\alpha} = \psi^{-1}\Big(7ln2 + \sum_{i=1}^{7} ln(y_i)\Big)$$

I am assuming this is this non closed ended issue that was mentioned in the handout. Let's graph the log-likelihood instead.

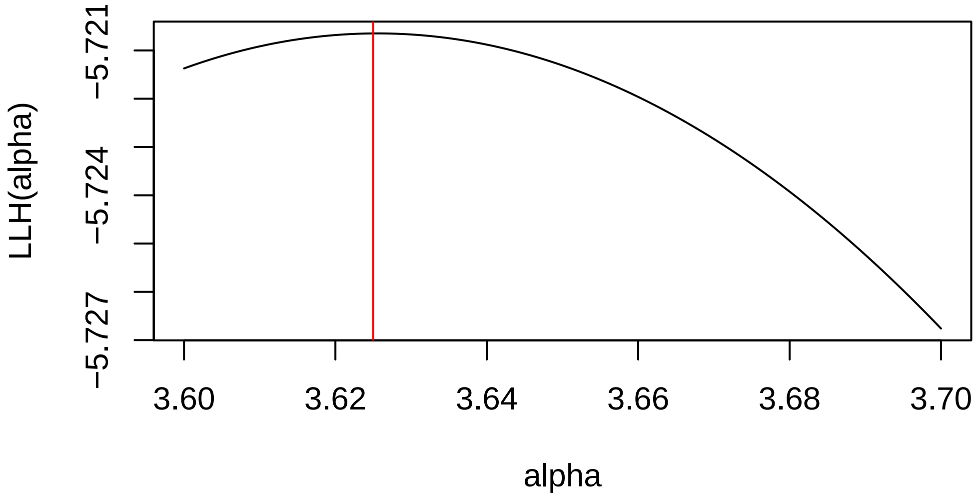
```
LLH <- function(a, y = yi){
  n = length(y)
  f = n*log(2^a/gamma(a)) + (a - 1)*log(prod(y)) - 2*sum(y)
  return(f)
}

alpha <- seq(0,10,length = 1000)
plot(alpha,LLH(alpha),type = 'l')
```

```
Warning in gamma(a): NaNs produced
```




```
alpha <- seq(3.60,3.70,length = 1000)
plot(alpha,LLH(alpha),type = "l")
abline(v=3.625,col = "red")
```



Eyeballing this to be about $\tilde{\alpha} = 3.625$

Question 4

(a).

$$P(\beta|y_i) \propto \Big(\prod_{i=1}^{7} \frac{\beta^4}{6} y_i^3 e^{-\beta y_i}\Big)\Big(0.8e^{-0.8\beta}\Big)$$

$$= \frac{0.8}{6^7}(\prod_i y_i^3)\beta^{28} e^{-\beta(0.8+\sum_i^7 y_i)} = 0.03681441\beta^{28}e^{-11.9\beta}$$

```
0.8/6^7*prod(yi^3) -> c1
0.8+sum(yi) -> c2
c1;c2
```

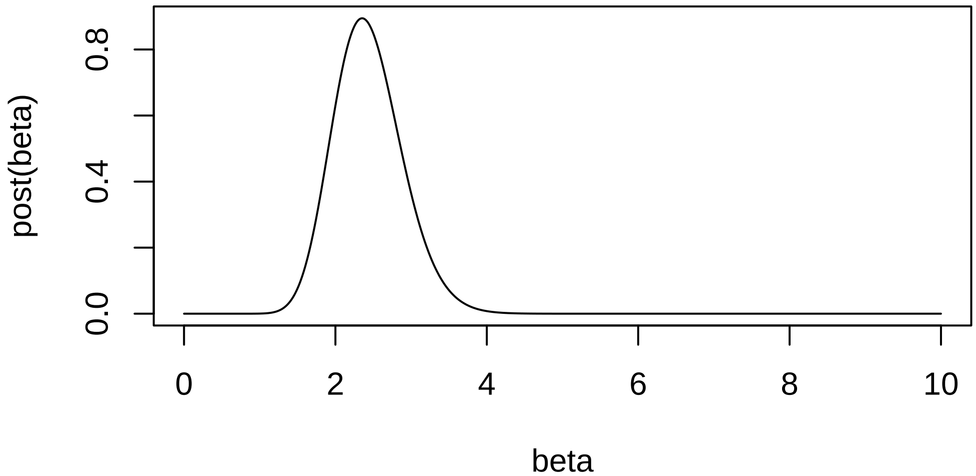
```
[1] 0.03681441
```

```
[1] 11.9
```

$$P(\beta|y_i) = \frac{0.03681441\beta^{28}e^{-11.9\beta}}{\int_0^\infty 0.03681441\beta^{28}e^{-11.9\beta}d\beta}$$

(b).

plotting the posterior
```
# plotting the posterior
post <- function(beta){
 # numerator
  numer <- function(x){
    0.03681441 * x^28 * exp(-c2*x)
  }
 # denominator
  denom <- integrate(numer, lower = 0, upper = Inf)$value
 # output
  return(numer(beta)/denom)
}
# render
beta <- seq(0,10,length = 1000)
plot(beta,post(beta),type = "l")
```



(c).

let's try an easy numerical evaluation approach

```
# make vector between 2 and 3
beta1 <- seq(2,3,length = 500000)
beta1[which(post(beta1) == max(post(beta1)))] -> MLE_beta
MLE_beta
```

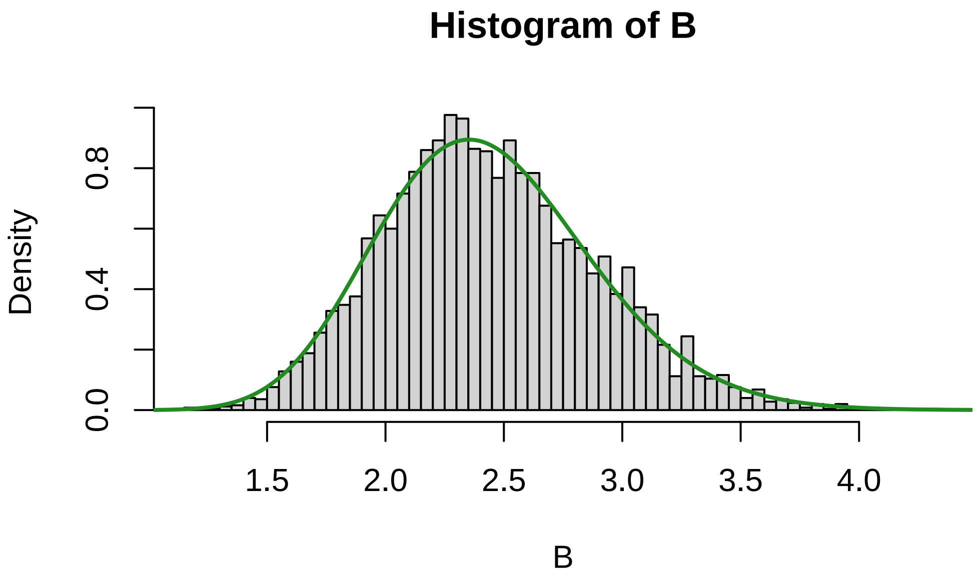
```
[1] 2.352941
```

(d).

let's generate some beta values using sample()

```r
  # render some values of beta from 0 to 6
  beta_sampl = runif(1000000,0,6)
  # get p_values
  p_sampl = post(beta_sampl)
  B = sample(size = 5000, beta_sampl,replace = T, prob = p_sampl)
```

```r
hist(B,breaks = 80, freq = F)
lines(beta,post(beta),col = "forestgreen", lwd = 2)
```




```r
mean(B)
```

```
[1] 2.444675
```

```r
for(v in c(0.025,.5,.975)){
  print(quantile(B,v))
  }
```

```
    2.5%
1.655775
     50%
2.408789
   97.5%
3.389496
```