贝叶斯习题2
朱强强
应用统计学1701

3.1 For the beta density with parameters $\alpha = 2$ and $\beta = 7$, do the following:

1. Refer to Table A.2, calculate the mean and mode as functions of the parameters.

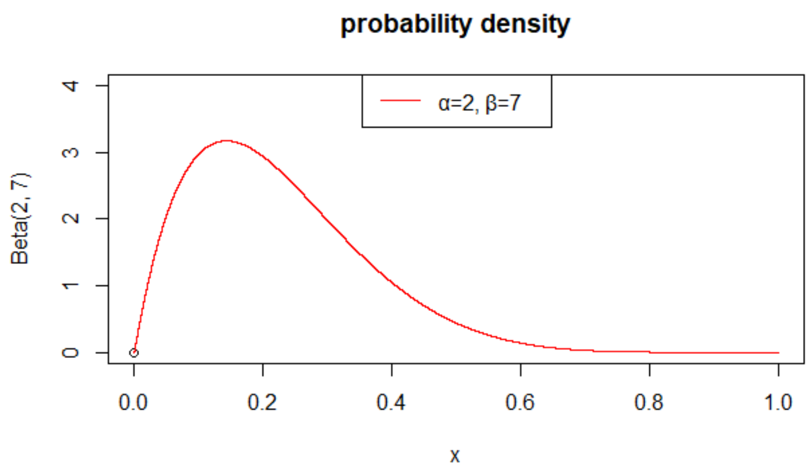
$$mean = \frac{\alpha}{\alpha + \beta} = 0.222$$
$$mode = \frac{\alpha - 1}{\alpha + \beta - 2} = 0.143$$

2. Use an R function to determine the median and a 90% central interval.

   $Median = 0.405$

   a 90% central interval is $[0.04638926, 0.47067941$.

3. Plot the density.

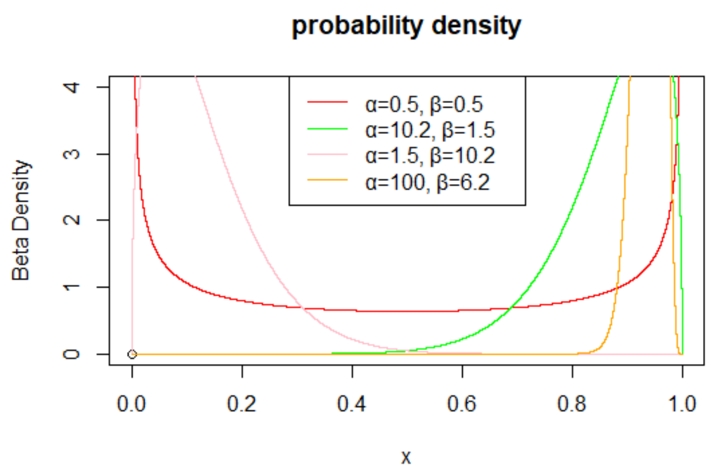



```r
setwd("E:\Bayesian Statistics\Homework 2")
a = 2
b = 7
# calculate the mean and mode
Mean = a / (a + b)
Mode = (a - 1) / (a + b - 2)
print(Mean)
print(Mode)
# calculate the median and a 90% central interval
x = seq(0, 1, length.out=1000000)
y = dbeta(x, a, b)
Median = quantile(y, 0.5)
print(Median)
print(qbeta(c(0.05, 0.95), 2, 7))
# plot the density
```

```
plot(0, 0, main='probability density', xlim=c(0, 1), ylim=c(0, 4),
     ylab='Beta(2, 7)', xlab='x')
lines(x, dbeta(x, a, b), col='red')
legend('top', legend='α=2, β=7', col='red', lwd=1)
```

3.2 Plot different beta densities.

1. $Beta(0.5, 0.5)$
2. $Beta(10.2, 1.5)$
3. $Beta(1.5, 10.2)$
4. $Beta(100, 6.2)$




```
# 3.2
plot(0, 0, main='probability density', xlim=c(0, 1), ylim=c(0, 4),
     ylab='Beta Density', xlab='x')
lines(x, dbeta(x, 0.5, 0.5), col='red')
lines(x, dbeta(x, 10.2, 1.5), col='green')
lines(x, dbeta(x, 1.5, 10.2),col='pink')
lines(x, dbeta(x, 100, 6.2), col='orange')
legend('top', legend=c('α=0.5, β=0.5','α=10.2, β=1.5', 'α=1.5, β=10.2', 'α=100, β=6.2'),
       col=c('red', 'green', 'pink', 'orange'), lwd=1)
```

3.3 The uniform distribution is a special case of the beta distribution.

1.

$$\begin{aligned} U(0,1) &= p^k(1-p)^{1-k}, k = 0, 1 \\ &= Beta(1,1) \end{aligned}$$

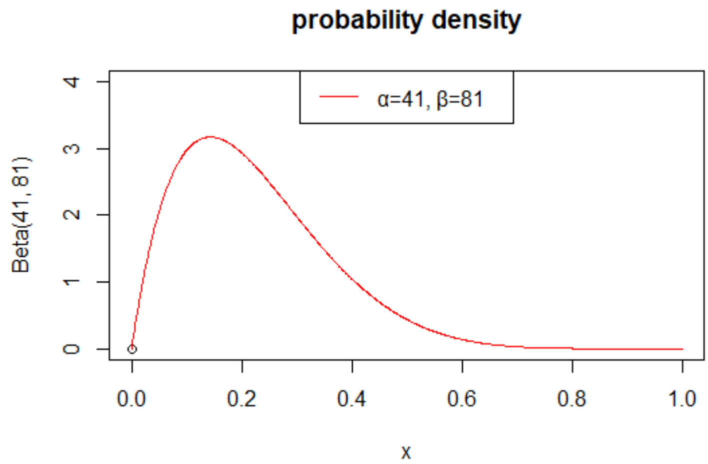
2. The information of posterior density for $Beta(1,1)$ is the equivalent prior sample size for a $U(0,1)$ prior.

$$\begin{aligned} \pi &\sim Beta(1,1) \\ p(\pi) &\propto \pi^{\alpha-1}(1-\pi)^{\beta-1} \\ &= U(0,1) \end{aligned}$$

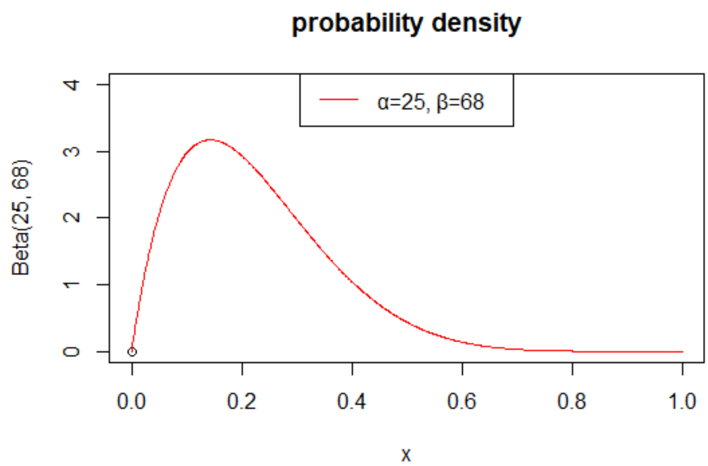
3.4

1. θ is her underlying true probability of getting a hit in any bat.

$$p(\theta|Data) \propto \theta^{40}(1-\theta)^{80}$$
$$\therefore \theta \sim Beta(41,81)$$



2. Assume that $p(y|\theta) \propto \theta^{24}(1-\theta)^{67} \propto Beta(25,68)$.

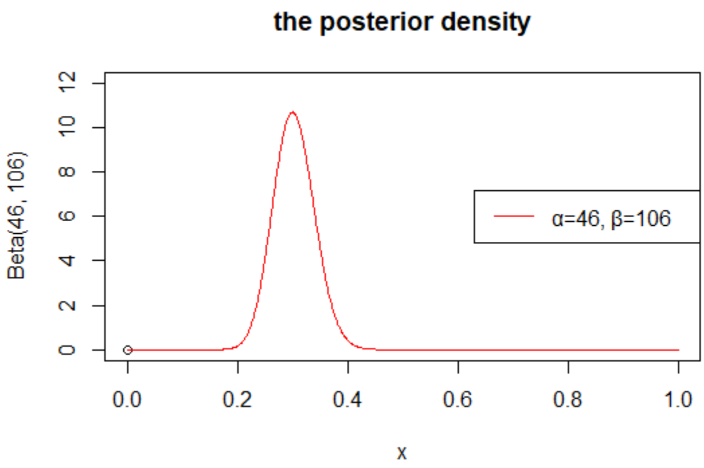


3. **Reason:** If the ball in front is hit, the player's mood will be better, and the ball behind is more likely to be hit. On the other hand, if the player keeps missing the ball, the probability of the ball behind him being hit will be affected and may decrease.

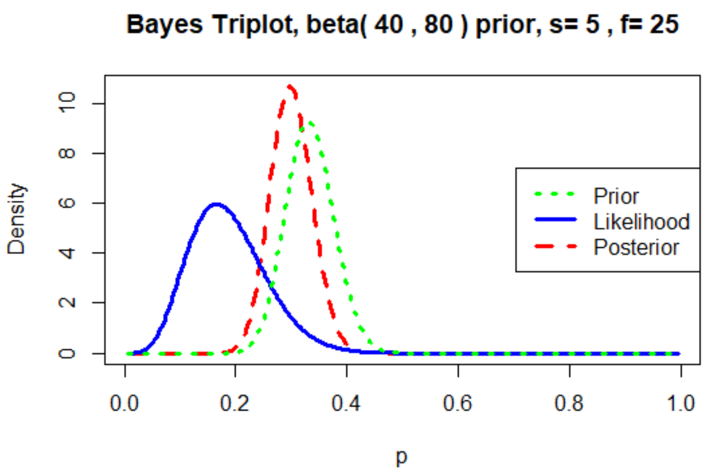
4. $y = 5, n = 3$

$$
\begin{aligned}
p(\theta|y) &\propto p(y|\theta) \cdot \pi(\theta) \\
&= \theta^5 (1-\theta)^{25} \cdot \theta^{40} (1-\theta)^{80} \\
&= \theta^{45} (1-\theta)^{105} \\
\theta|y &= Beta(46, 106) \\
Mean &= 0.3026316 \\
Mean &= 0.3
\end{aligned}
$$

Posterior density plot.



A plot showing the prior density, the likelihood, and the posterior density.




```
a = 46
b = 106
Mean = a / (a + b)
Mode = (a - 1) / (a + b - 2)
print(Mean)
print(Mode)
plot(0, 0, main='the posterior density', xlim=c(0, 1), ylim=c(0, 12),
     ylab='Beta(46, 106)', xlab='x')
lines(x, dbeta(x, a, b), col='red')
legend('right', legend='α=46, β=106', col='red', lwd=1)
triplot(c(40, 80), c(5, 25), "right")
```

4.1 $p(y|\theta) \propto Beta(46, 106)$

a. Calculate the mean and mode of posterior distribution.

$\alpha = 41, \beta = 81, y = 5, n = 30$

$$mean = \frac{\alpha + y}{\alpha + \beta + n} = 0.3026316$$
$$mode = \frac{\alpha + y - 1}{\alpha + \beta + n - 2} = 0.3$$

b. 95% posterior interval for θ.

The result computed by R is $[0.2324309, 0.3777516]$.

c. Posterior probability that $\theta > 0.25$.

The result computed by R is 0.3272937.

```r
# 4.1
a = 41
b = 81
y = 5
n = 30
# calculate the mean and mode
Mean = (a + y) / (a + b + n)
Mode = (a + y - 1) / (a + b + n - 2)
print(Mean)
print(Mode)
# 95% posterior interval for theta
print(qbeta(c(0.025, 0.975), a + y, b + n - y))
# Posterior probability that theta > 0.25
print(qbeta(0.25, a + y, b + n - y, lower.tail=FALSE))
```

4.2

1. $p\_value = 0.0001928$
2. $\pi \sim Beta(132, 197)$

$$H_0 : \pi \geq 0.5$$
$$H_1 : \pi < 0.5$$
$$p(H_0|\pi) = 0.0001590998$$
$$p(H_1|\pi) = 0.9998409$$

```r
# 4.2
binom.test(131, 327, p=0.5, alternative="less")
print(pbeta(0.5,132,197))
print(1-pbeta(0.5,132,197))
```

4.2

1. $y^* = 8$

   $Pr(y^* = 8|y) = 0.1744675$

2. set = [8 7 9 6 10 5 11 12 4]

```r
# 4.3
library(LearnBayes)
pbetap(c(132, 197), 20, 8)
prob = pbetap(c(132, 197), 20, 0:20)
prob1 = data.frame(0:20, prob)
prob1 = prob1[order(prob1$prob, decreasing = T),]
i = 0
while (i < nrow(prob1)) {
  if (sum(prob1[1:i, 2]) < 0.95) {
    i = i + 1
  }
  else {
    print(i)
    break
  }
}
prob1[1:i, 1]
```

5.1

We know that in case of the binomial likelihood, the density is as $Beta(\frac{1}{2}, \frac{1}{2})$.

1. **Jeffreys prior**

   a. Calculate the mean and mode of posterior distribution.

   $\alpha = \frac{1}{2}, \beta = \frac{1}{2}, y = 5, n = 30$

$$mean = \frac{\alpha + y}{\alpha + \beta + n} = 0.1774194$$
$$mode = \frac{\alpha + y - 1}{\alpha + \beta + n - 2} = 0.1551724$$

b. 95% posterior interval for θ.

The result computed by R is $[0.06657395, 0.32742775]$.

c. Posterior probability that $\theta > 0.25$.

The result computed by R is 0.2193795.

2. **Beta(0,0)**

a. Calculate the mean and mode of posterior distribution.

$\alpha = 0, \beta = 0, y = 5, n = 30$

$$mean = \frac{\alpha + y}{\alpha + \beta + n} = 0.1666667$$
$$mode = \frac{\alpha + y - 1}{\alpha + \beta + n - 2} = 0.1428571$$

b. 95% posterior interval for θ.

The result computed by R is $[0.05845608, 0.31664061]$.

c. Posterior probability that $\theta > 0.25$.

The result computed by R is 0.2078793.

5.2

Prior density	$Pr(\pi > 0.25\|y)$	$E(\pi\|y)$	**95% equal tail credible set**
$Beta(41, 81)$	0.3272937	0.3026316	[0.2324309, 0.3777516]
$Beta(25, 86)$	0.2692383	0.2439024	[0.1724670, 0.3232227]
$U(0, 0)$	0.2300388	0.1875	[0.0745199, 0.3372716]
$Beta(0.5, 0.5)$	0.2193795	0.1774194	[0.06657395, 0.32742775]
$Beta(0, 0)$	0.2078793	0.1666667	[0.05845608, 0.31664061]

From the table above we can know that the prior we assume is not relatively robust.

5.3

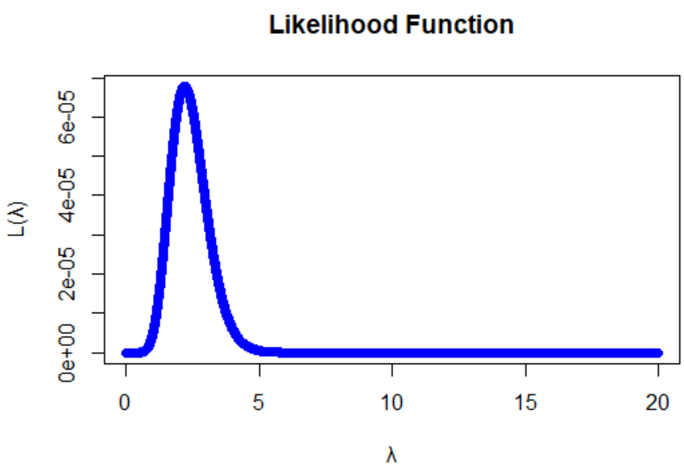
$$\phi = g(\pi) = log(\frac{\pi}{1-\pi})$$
$$\pi = g^{-1}(\phi) = \frac{exp(\phi)}{1+exp(\phi)}$$
$$p_\phi(\phi) = p_\pi(f(\phi))\left|\frac{d\pi}{d\phi}\right|$$
$$= \frac{exp(\phi)}{[1+exp(\phi)]^2}$$

5.4

1. The joint probability distribution of $y_1, y_2, \ldots, y_n$ is

$$p(y_1, y_2, \ldots, y_n|\lambda) = \prod_{i=1}^{n} \frac{\lambda^{y_i}}{y_i!}e^{-\lambda}$$
$$= \frac{\lambda^{\sum_{i=1}^{n} y_i}}{y_1!y_2!\ldots y_n!}e^{-n\lambda}$$
$$y_i = 0, 1, 2, \ldots$$

2. $$L(\lambda) = \frac{\lambda^{11}}{1440}e^{-5\lambda}$$



3. Find the mle of λ.

$$\begin{aligned} L(\lambda) &= log(p(y_1, y_2, \ldots, y_n|\lambda)) \\ &= log(\lambda) \cdot \sum_{i=1}^{n} y_i - log(\sum_{i=1}^{n} y_i!) - n\lambda \\ \frac{\partial L(\lambda)}{\partial \lambda} &= \frac{1}{\lambda} \sum_{i=1}^{n} y_i - n = 0 \\ \lambda_{MLE} &= \frac{1}{n} \sum_{i=1}^{n} y_i = \bar{y} \end{aligned}$$

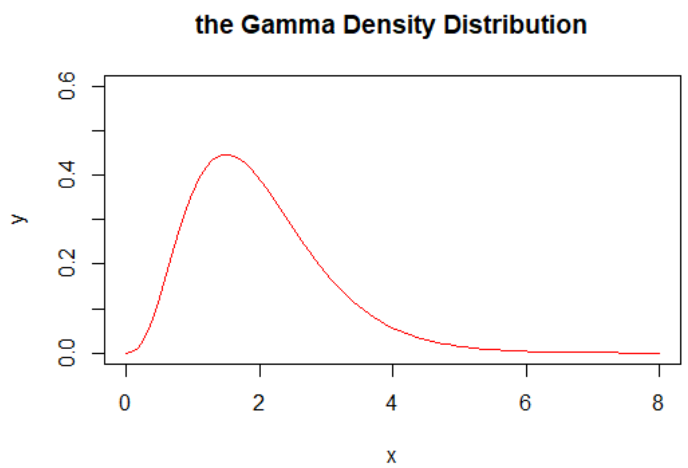
The mle of λ computed by R is also 2.2, and a 95% frequentist confidence interval for λ is [1.098232, 3.936408].

```r
# 5.4
x = seq(0, 20, length.out=100000)
y = x^11/1440*exp(-5*x)
plot(x, y, main='Likelihood Function', xlab='λ', ylab='L(λ)', col='blue')
poisson.test(x=sum(c(2, 5, 1, 0, 3)), T=5, alternative="two.sided",                  conf.level=0.95)
```

5.5

$$p(\lambda) \propto \lambda^3 exp(-2\lambda), \lambda > 0$$

1. $p(\lambda \propto Gamma(4, 2)$
2. From the figure below we can conclude that the most likely value of λ is within [1, 2] , and almost all values of λ are within [0, 6].



```
# 5.5.2
x = seq(0, 6, length.out=100)
y = dgamma(x, 4, 2)
plot(x, y, main="the Gamma Density Distribution", xlim=c(0,8),
ylim=c(0,0.6),       col="red", type="l")
```

3. We know that

$$
\begin{aligned}
p(\lambda|y) &\propto \lambda^3 exp(-2\lambda) \cdot \lambda^{\sum_{i=1}^n y_i} e^{-n\lambda} \\
&= \lambda^{3+\sum_{i=1}^n y_i} exp(-(2+n)\lambda) \\
\therefore \lambda|y &\sim Gamma(\sum_{i=1}^n y_i + 4, n+2)
\end{aligned}
$$

4. $\lambda|y \sim Gamma(15, 7)$

   $Mean = \frac{15}{7} = 2.142857$

   The 95% central credible set for λ is [1.199341, 3.355660].

```
# 5.3.4
mean = 15 / 7
print(mean)
qgamma(c(0.025, 0.975), 15, 7)
```

5. Yes. The prior, posterior and λ have the same kernel function, which is the form of $Gamma$ distribution, so they are conjugate.

5.6

1. Derive the Jeffreys prior that goes with Poisson likelihood.

$$
\begin{aligned}
&L(\lambda) = log(p(y_1, y_2, \ldots, y_n|\lambda)) = log(\lambda) \cdot \sum_{i=1}^n y_i - log(\sum_{i=1}^n y_i!) - n\lambda \\
&\frac{\partial L(\lambda)}{\partial \lambda} = \frac{1}{\lambda} \sum_{i=1}^n y_i - n \\
&\frac{\partial^2 L(\lambda)}{\partial \lambda^2} = -\frac{1}{\lambda^2} \sum_{i=1}^n y_i \\
&I(\lambda) = -E\left(-\frac{\sum_{i=1}^n y_i}{\lambda^2}\right) = \frac{n\lambda}{\lambda^2} = \frac{n}{\lambda} \\
&p(\lambda) \propto \sqrt{I(\lambda)} = \sqrt{n}\lambda^{-\frac{1}{2}}
\end{aligned}
$$

We recognize this density as $Gamma(\frac{1}{2}, 0)$.

2. $p(\lambda) \propto \lambda^{-0.5}\lambda^{11}exp(-5\lambda) = \lambda^{10.5}exp(-5\lambda)$

   $\therefore \lambda \sim Gamma(11.5, 5)$

3. $Mean = \frac{11.5}{5} = 2.3$

   The 95% central credible set for λ is [1.168855, 3.807563].

```
# 5.6.3
mean = 11.5 / 5
print(mean)
qgamma(c(0.025, 0.975), 11.5, 5)
```

4. From the table below we can see that the difference between Bayesian point estimation and classical point estimation is not large, but the interval estimation of Bayesian method is narrower than that of classical method.

	Point estimation	Interval estimation
Bayesian method	2.3	[1.168855, 3.807563]
Classical method	2.2	[1.098232, 3.936408]

5. Because new employee are of prior infomation, he have to give an interval estimation on the basis of noninformative prior which is Jeffreys prior distribution. So the range of set is wider than that of president. However, the mean difference between the two methods is relatively small.

	Point estimation	Interval estimation
New employee	2.3	[1.168855, 3.807563]
President	2.142857	[1.199341, 3.355660]

6. $p(\lambda > 2|Data1) = 0.6419118$

7. $p(\lambda > 2|Data1) = 0.5704367$

```
# 5.6.6
pgamma(2, 11.5, 5, lower.tail=FALSE)
# 5.6.7
pgamma(2, 15, 7, lower.tail=FALSE)
```

5.7

$$Beta(0,0) \sim \pi^{-1}(1-\pi)^{-1}$$
$$p(\pi|y) = \pi^{-1}(1-\pi)^{-1} \cdot \pi^{y}(1-\pi)^{n-y} = \pi^{y-1}(1-\pi)^{n-y-1}$$
$$\therefore \pi|y \propto Beta(y, n-y)$$
$$E(\pi|y) = \frac{y}{n}$$
$$L(y_i;\pi) = log(\prod_{i=1}^{n} \pi^{y_i}(1-\pi)^{1-y_i}) = log(\pi)\sum_{i=1}^{n} y_i + (n - \sum_{i=1}^{n} y_i) log(1-\pi)$$
$$\frac{\partial L(y_i;\pi)}{\partial \pi} = \frac{\sum_{i=1}^{n} y_i}{\pi} - \frac{n - \sum_{i=1}^{n} y_i}{1-\pi} = 0$$
$$\pi_{MLE} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{y}{n}$$