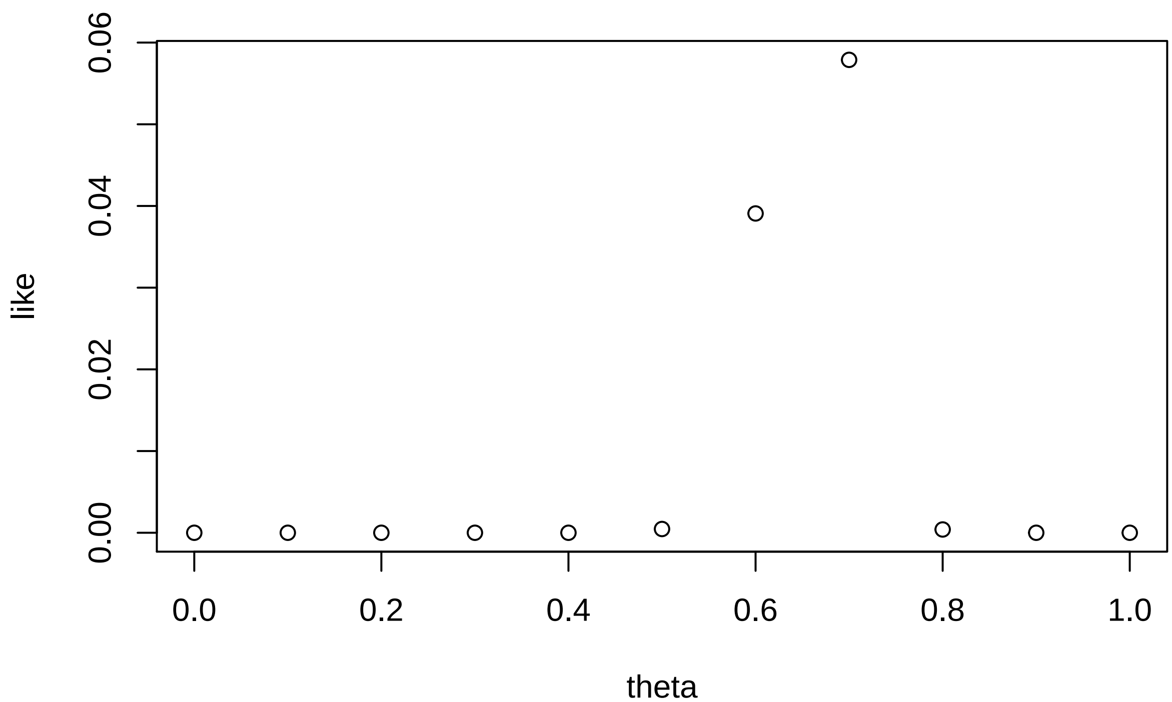
Biostat682 Hw1

David (Daiwei) Zhang

September 26, 2017

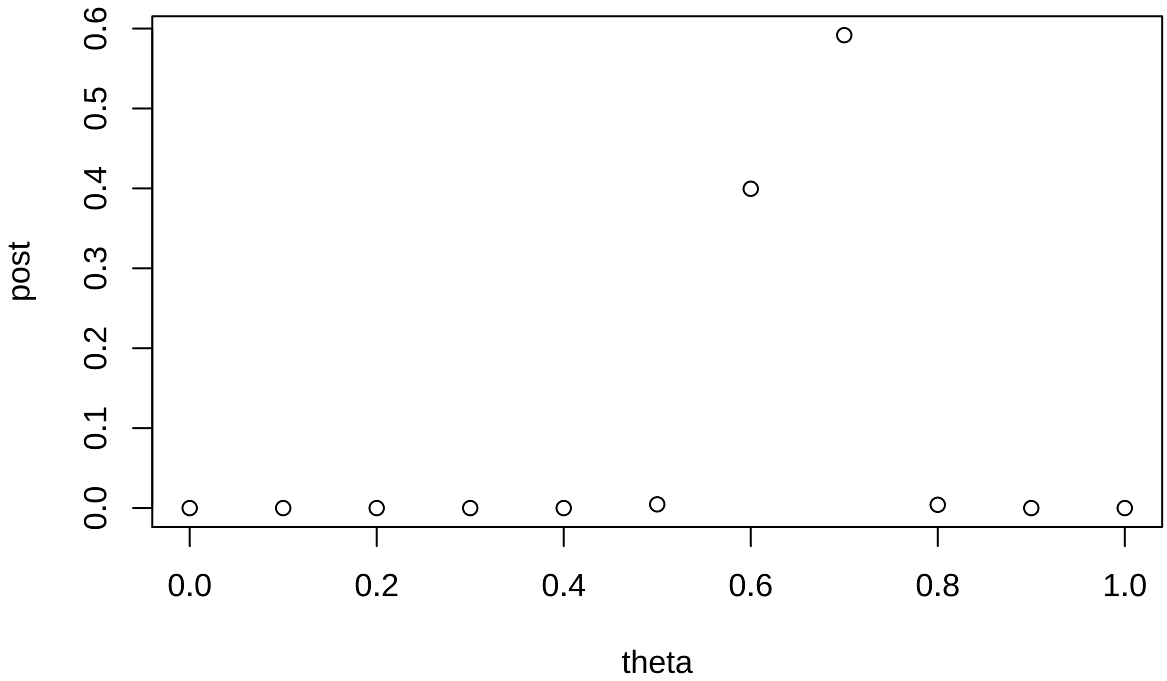
Problem 2(b)

```
theta <- c(0:10) * 0.1
y <- 66
n <- 100
likelihood <- function(theta){
  choose(n, y) * theta^y * (1-theta)^(n-y)
}
like <- likelihood(theta)
plot(theta, like)
```



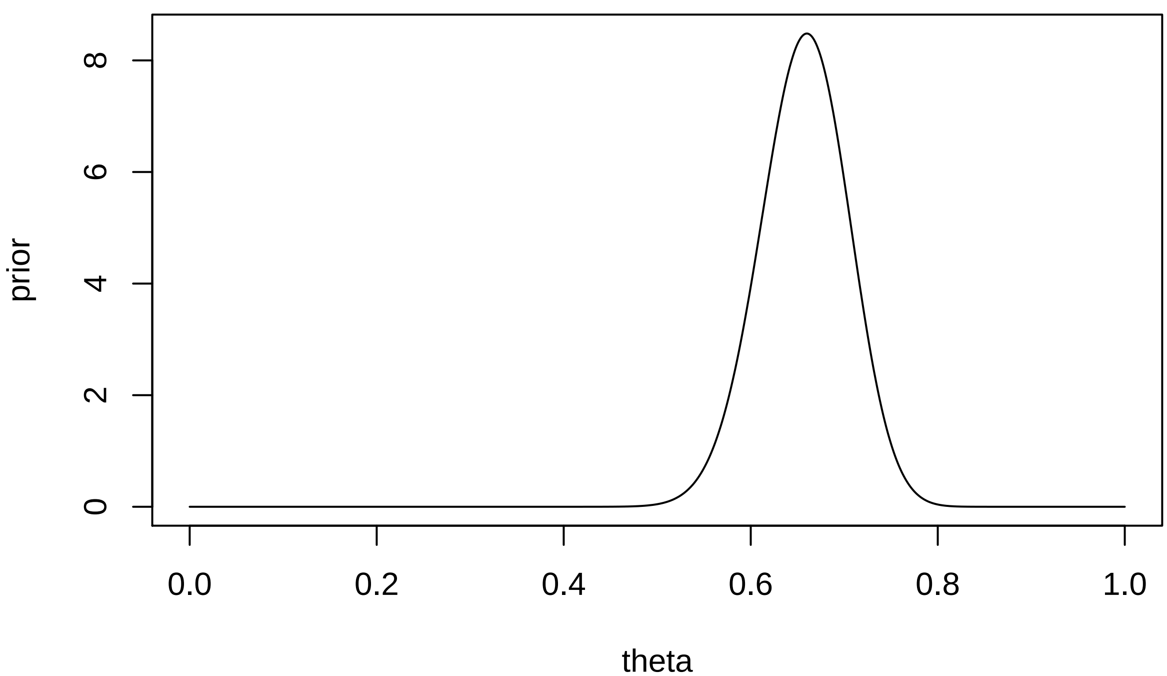
Problem 2(c)

```
m <- length(theta)
prior <- rep(1/m, m)
post <- prior * like
post <- post / sum(post)
plot(theta, post)
```



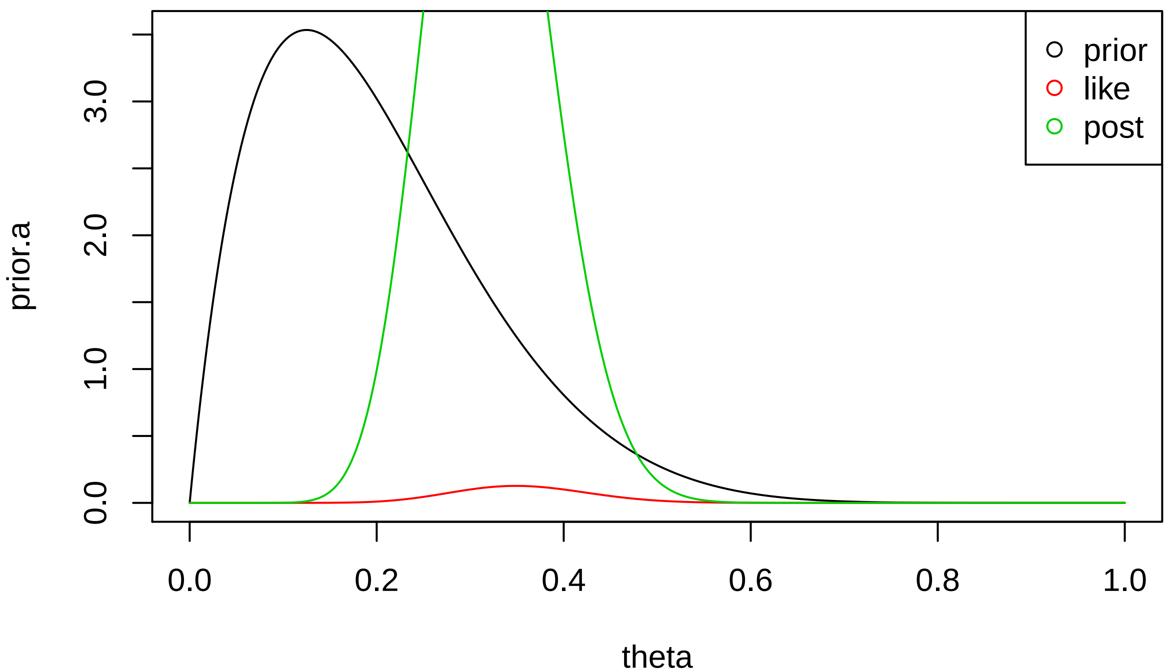
Problem 2(d)

```
m <- 1000
theta <- c(0:m) / m
like <- likelihood(theta)
prior <- like * 1.0
prior.norm <- sum(prior) / m
prior <- prior / prior.norm
plot(theta, prior, type = "l")
```



Problem 3(a)

```r
m <- 1000
theta <- c(0:m) / m
beta.a <- 2
beta.b <- 8
y <- 15
n <- 43
prior.a <- dbeta(theta, beta.a, beta.b)
like <- dbinom(y, n, theta)
post <- prior.a * like
post.norm <- sum(post) / m
post <- post / post.norm
plot(theta, prior.a, type = "l", col = 1)
points(theta, like, type = "l", col = 2)
points(theta, post, type = "l", col = 3)
legend("topright", col = 1:3, legend = c("prior", "like", "post"), pch = 1)
```



We have

$$\pi(\theta) \sim Beta(2, 8),$$

so

$$\pi(\theta|y) \sim Beta(2 + 15, 8 + 28) = Beta(17, 36).$$

Then

$$E[\theta|y] = \frac{17}{17 + 36} = 0.32$$

$$\text{argmax}\, \pi(\theta|y) = \frac{17 - 1}{17 + 36 - 2} = 0.31$$

$$\sigma = \sqrt{\frac{17 * 36}{(17 + 36)^2(17 + 36 + 1)}} = 0.06$$

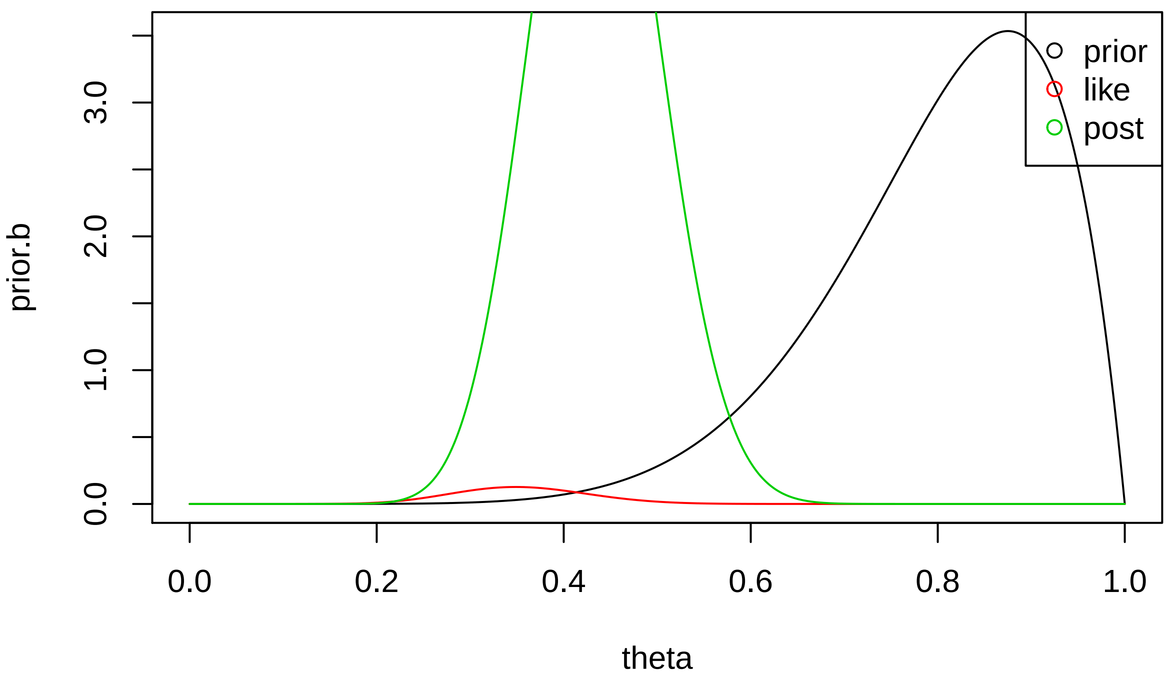
and the 95% crediblity interval is

```
ci <- qbeta(c(0.025, 0.975), 17, 36)
print(ci)
```

```
## [1] 0.2032978 0.4510240
```

Problem 3(b)

```
m <- 1000
theta <- c(0:m) / m
beta.a <- 8
beta.b <- 2
y <- 15
n <- 43
prior.b <- dbeta(theta, beta.a, beta.b)
like <- dbinom(y, n, theta)
post <- prior.b * like
post.norm <- sum(post) / m
post <- post / post.norm
plot(theta, prior.b, type = "l", col = 1)
points(theta, like, type = "l", col = 2)
points(theta, post, type = "l", col = 3)
legend("topright", col = 1:3, legend = c("prior", "like", "post"), pch = 1)
```



We have

$$\pi(\theta) \sim Beta(8, 2),$$

so

$$\pi(\theta|y) \sim Beta(8 + 15, 2 + 28) = Beta(23, 30).$$

Then

$$E[\theta|y] = \frac{23}{23+30} = 0.43$$
$$\operatorname{argmax} \pi(\theta|y) = \frac{23-1}{23+30-2} = 0.43$$
$$\sigma = \sqrt{\frac{23*30}{(17+29)^2(17+29+1)}} = 0.07$$

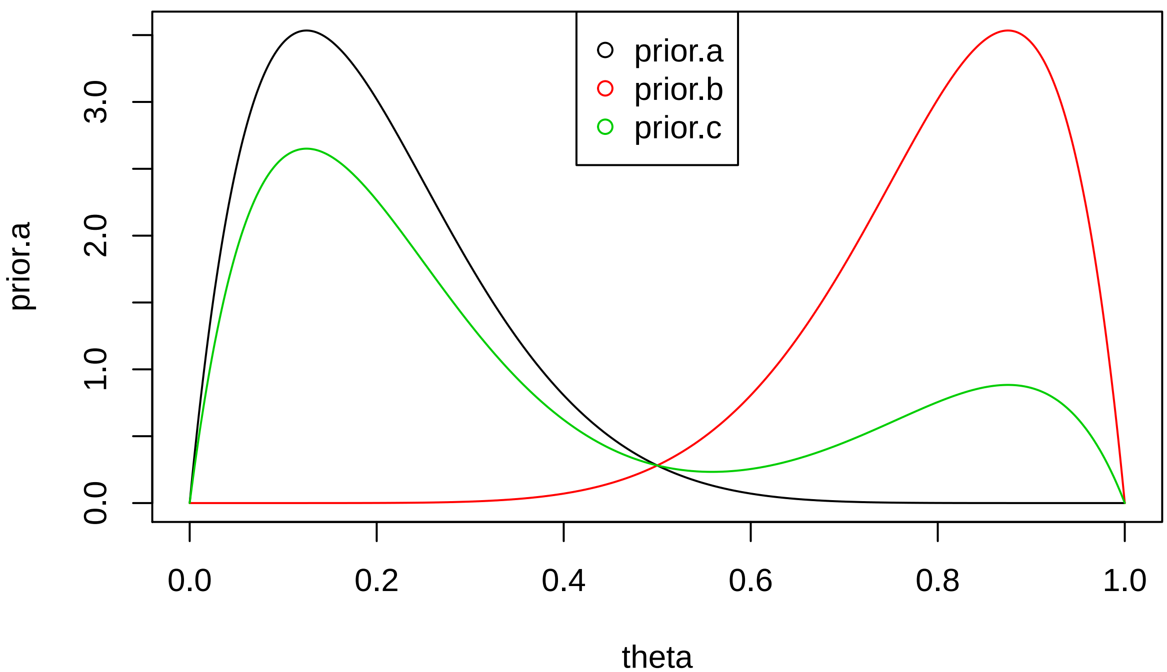
and the 95% crediblity interval is

```
ci <- qbeta(c(0.025, 0.975), 23, 30)
print(ci)
```

```
## [1] 0.3046956 0.5679528
```

Problem 3(c)

```
prior.fun <- function(theta){
  1/4 * gamma(10) / (gamma(2) * gamma(8)) * (3 * theta * (1-theta)^7 + theta^7*(1-theta))
}
prior.c <- prior.fun(theta)
plot(theta, prior.a, col = 1, type = "l")
points(theta, prior.b, col = 2, type = "l")
points(theta, prior.c, col = 3, type = "l")
legend("top", col = 1:3, legend = c("prior.a", "prior.b", "prior.c"), pch = 1)
```



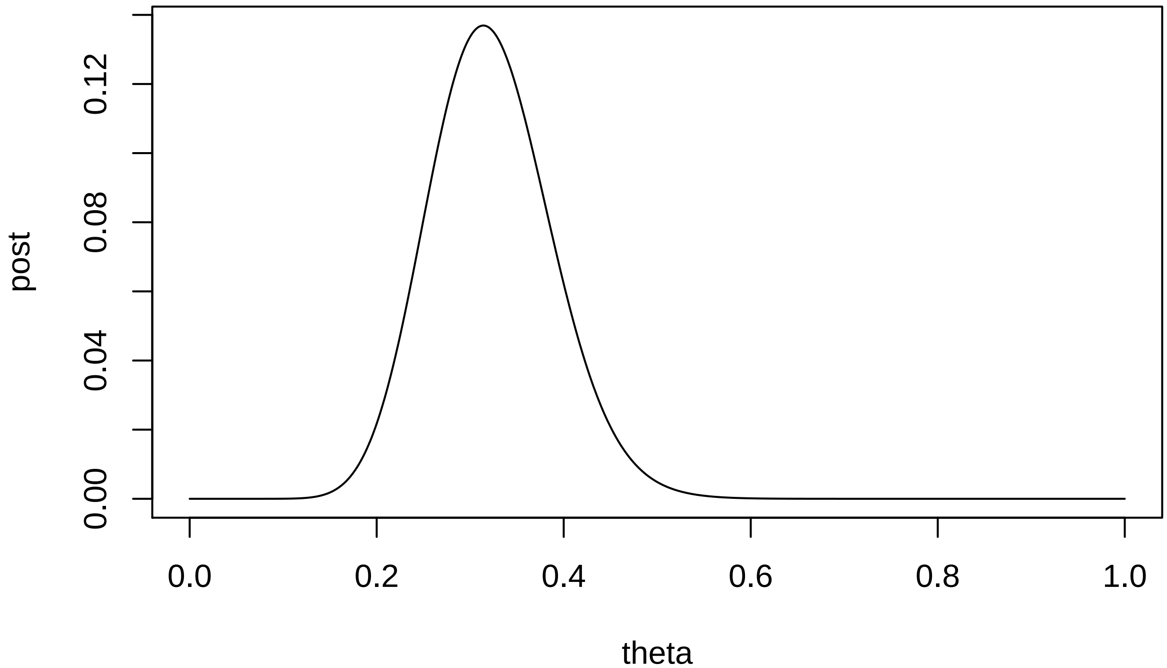
Prior opinion expressed: "I believe that the odds of recidivism is 7 to 1 or 1 to 7, and I believe the former has a 0.75 chance to be true and the latter 0.25."

Problem 3(d)(iii)

```
posterior <- function(theta){
  1/4 * gamma(44) / (gamma(16) * gamma(29)) * gamma(10) / (gamma(2) * gamma(8)) * (3 * theta^16 * (1-the
}
post <- posterior(theta)
post.mode.idx <- which.max(post)
post.mode <- theta[post.mode.idx]
print(post.mode)
```

```
## [1] 0.314
```

```
plot(theta, post, type = "l")
```



Here the mode is between the modes for the previous two posteriors, but it is much closer to the mode of the posterior with the 0.75 weight than that with the 0.25 weight.