Homework 1

Umberto Junior Mele 1388371

13 marzo 2017

a)

Y_i is a random variable with distribution *Bernoulli* and *i.i.d.* so we can write down the joint distribution, like this:

$$Pr(Y_1 = y_1, ..., Y_{100} = y_{100}|\theta) = \prod_{i=1}^{100} \theta^{y_i}(1-\theta)^{1-y_i} = L(\theta)$$

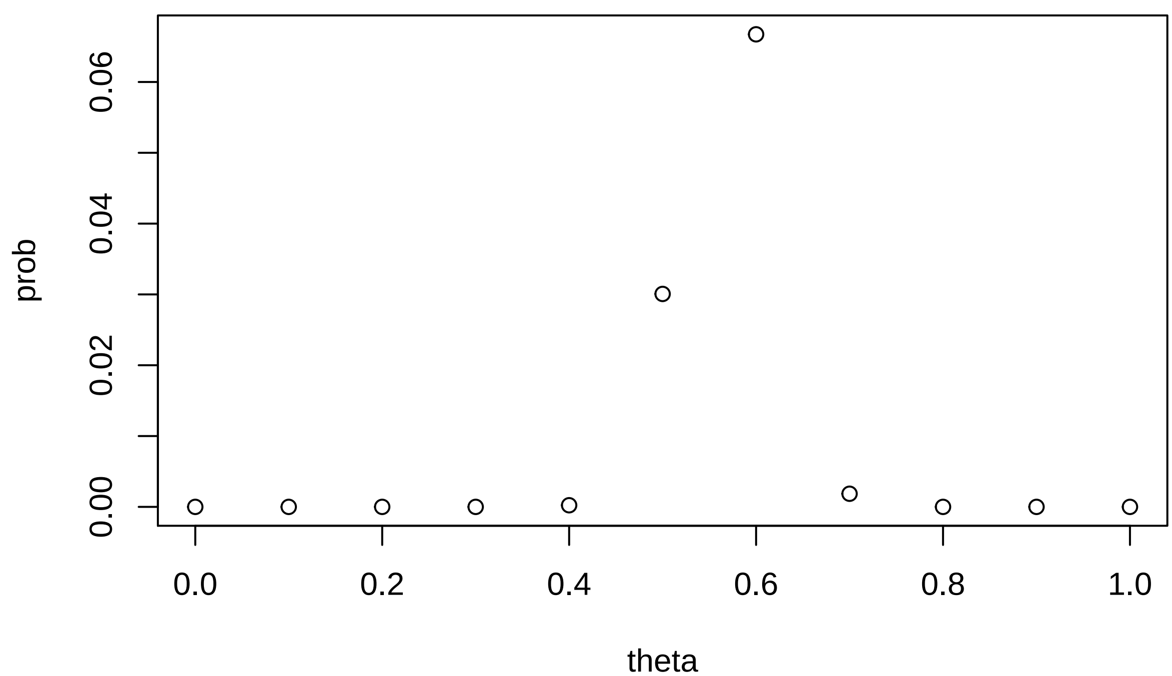
where $L(\theta)$ is the Likelihood function.

While the probability distribution of $Pr(\sum_{i=1}^{100} Y_i = y|\theta)$ is defined by the random variable $Z = Y_1 + ... + Y_{100}$, that is distribuited like a Binomial distribution, so:

$$Pr(\sum_{i=1}^{100} Y_i = y|\theta) = Pr(Z = y) = \binom{100}{y}\theta^y(1-\theta)^{100-y}$$

2)

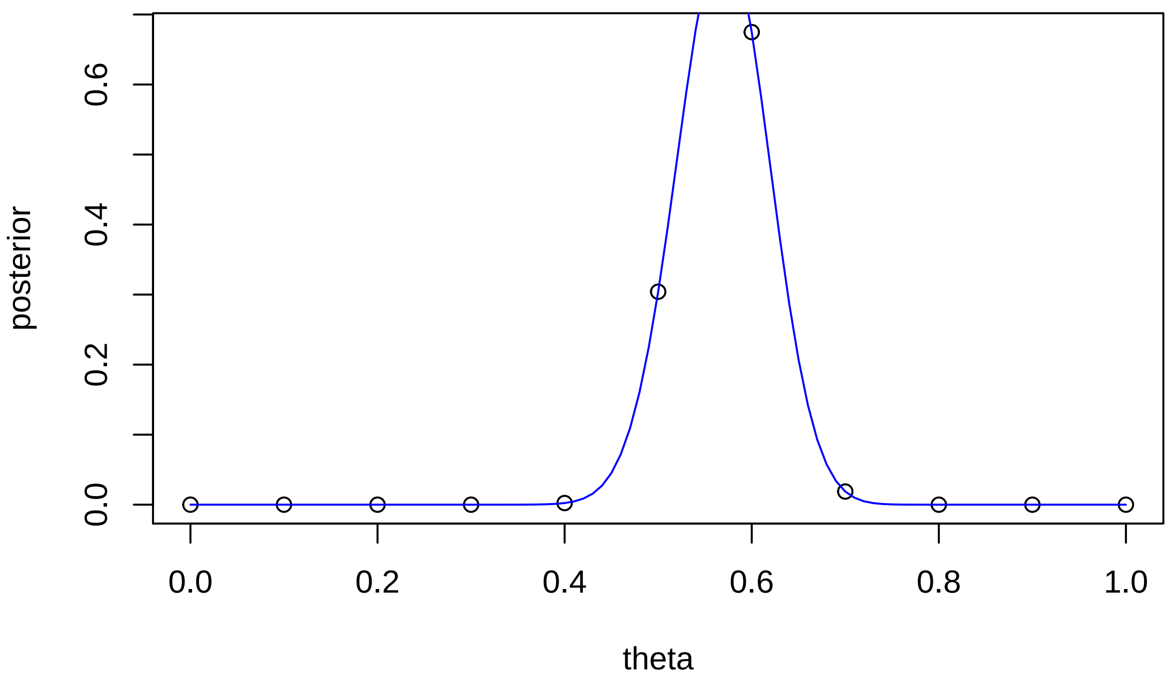
```r
theta <- seq(0,1,0.1)
prob_Z.equal.57 <- dbinom(57 ,100, theta)
plot(theta, prob_Z.equal.57, xlab = 'theta' , ylab = 'prob')
```



3) Remembering that the posterior distribution is:

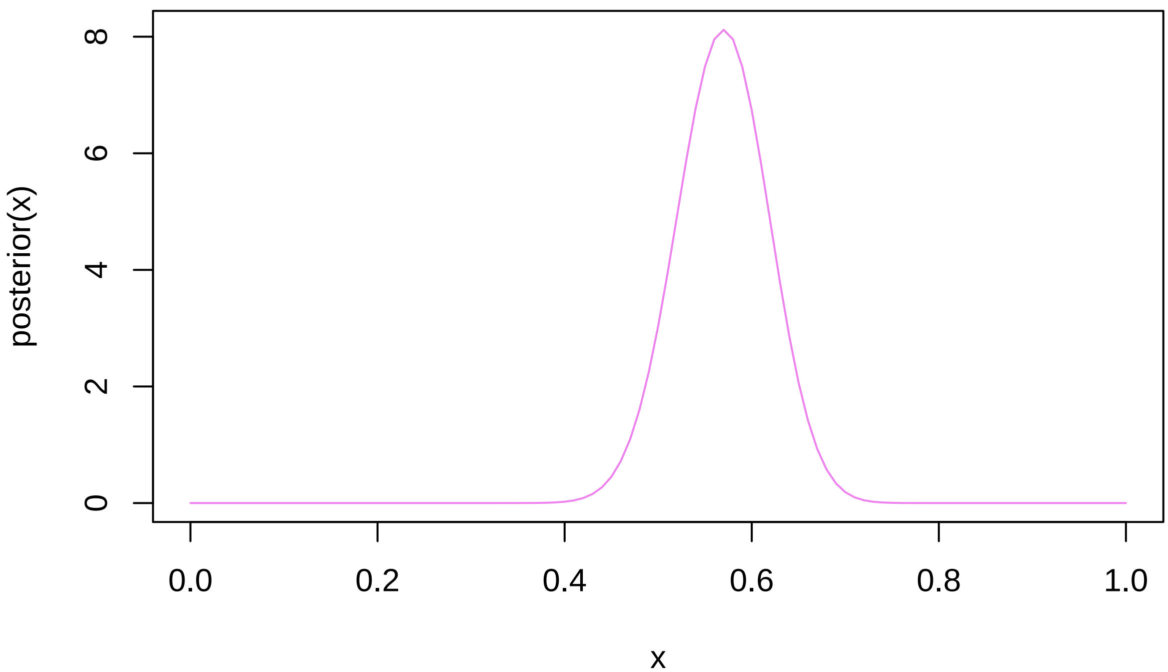
$$\pi(\theta|x) = \frac{L(\theta) \cdot \pi(\theta)}{m(x)}$$

```
prior <- function(t)  1/11 *(t < 1 & t >0)
L <-  function(t) dbinom(57,100,t)
m <- function(all.theta){
  su <- 0
  for(t in all.theta){
    su <- su + prior(t)*L(t)
  }
  return(su)
}

posterior <- function(t) prior(t)*L(t)/m(theta)
posterior <- Vectorize(posterior)

plot(theta ,posterior(theta), ylab = 'posterior')
curve(posterior(x), add = T, col='blue')
```




```
prior <- function(t) dunif(t)*(t>0 & t<1 )
L <-  function(t) dbinom(57,100,t)
joint <- function(t) L(t)*prior(t)
m <- integrate(joint,0,1)

posterior <- function(t) prior(t)*L(t)/m$value
posterior <- Vectorize(posterior)
```

```r
curve(posterior, col = 'violet')
```

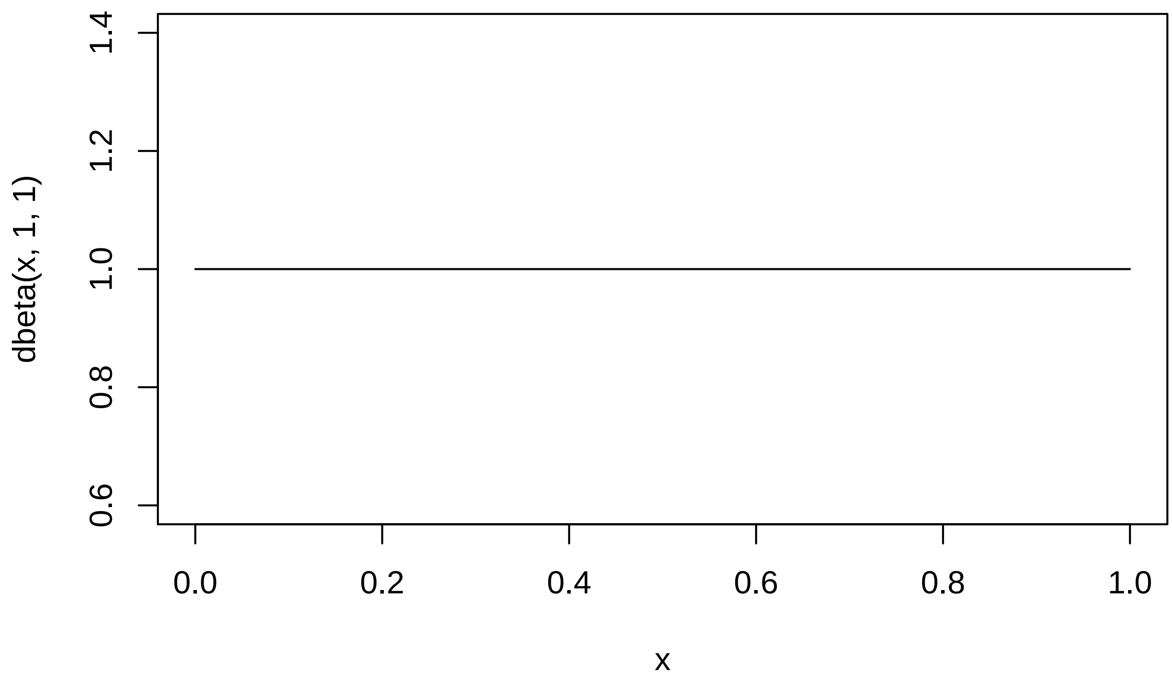


e) We know that the posterior distribution is a $Beta(1+57, 1+43)$ because if we use the Conjugate Analysis for this problem, so we use a $Beta$ distro as Prior:

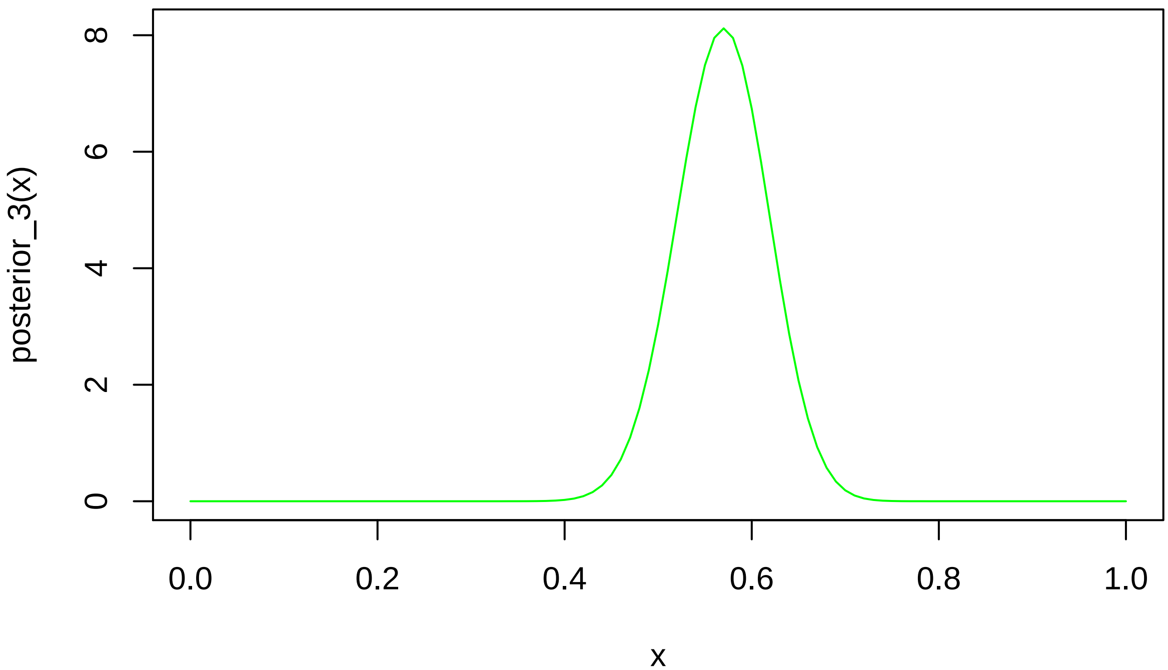
$$\pi(\theta|\bar{Z}) = \binom{100}{57}\theta^{57}(1-\theta)^{100-57} \cdot \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{\beta(\alpha, \beta)} \propto \theta^{57+\alpha-1}(1-\theta)^{43+\beta-1}$$

so the posterior distribution is a $Beta$ distribution with parameters $\dot{\alpha} = 57 + \alpha$ and $\dot{\beta} = 43 + \beta$, so the prior distribution is a $Beta(1, 1)$ that is the same distribution of a $Unif(0, 1)$.

```r
curve(dbeta(x,1,1))
```




```
prior <- function(t) dbeta(t,1,1)*(t>0 & t<1 )
L <-  function(t) dbinom(57,100,t)
joint <- function(t) L(t)*prior(t)
m <- integrate(joint,0,1)

posterior_3 <- function(t) prior(t)*L(t)/m$value
posterior_3 <- Vectorize(posterior_3)

curve(posterior_3, col = 'green')
```



2)

a)

The prior predictive distribution is defined by:

$$m(\cdot) = \int f(\cdot|\theta)\pi(\theta)d\theta = \int \frac{\sqrt{\lambda}}{\sqrt{2\pi}}e^{-\frac{\lambda(x-\theta)^2}{2}} \cdot \frac{\sqrt{\nu}}{\sqrt{2\pi}}e^{-\frac{\nu(\theta-\mu)^2}{2}}d\theta$$

$$\propto \int e^{-\frac{\lambda(x^2-2\theta x+\theta^2)+\nu(\theta^2-2\theta\mu+\mu^2)}{2}}d\theta \quad \propto e^{-\frac{\lambda x^2+\nu\mu^2}{2}} \int e^{-\frac{(\lambda+\nu)\theta^2}{2}+(x+\mu)\theta}d\theta$$

Now remembering that:

$$N(\mu = \frac{b}{a}; \sigma^2 = \frac{1}{a}) \propto exp\{-\frac{ax^2}{2} + bx\}$$

and..

$$\int \frac{\sqrt{a}}{\sqrt{2\pi}} \cdot e^{-\frac{ax^2}{2}+bx-\frac{b^2}{2a}}dx = 1$$

we can compute:

$$\int e^{-\frac{ax^2}{2}+bx} = \frac{\sqrt{2\pi}}{\sqrt{a}} \cdot e^{\frac{b^2}{2a}}$$

so:

$$m(\cdot) \propto \quad e^{-\frac{\lambda x^2+\nu\mu^2}{2}} \int e^{-\frac{(\lambda+\nu)\theta^2}{2}+(x+\mu)\theta}d\theta \quad \propto e^{-\frac{\lambda x^2+\nu\mu^2}{2}} \cdot e^{\frac{(x+\mu)^2}{2(\lambda+\nu)}}$$

$$\propto e^{-\frac{\lambda x^2}{2}} \cdot e^{\frac{x^2+2x\mu}{2(\lambda+\nu)}} \propto e^{-\frac{1}{2}(\lambda-\frac{1}{\lambda+\nu})x^2+\frac{\mu}{\lambda+\nu}x}$$

the last function is a $N(a = (\lambda - \frac{1}{\lambda+\nu}); b = \frac{\mu}{\lambda+\nu})$ and is the prior predictive distribution.

//: <> (but we can use also a trick ...

$$X \sim N(\theta, \lambda = \frac{1}{\sigma^2})$$

$$\theta \sim N(\mu, \nu = \frac{1}{\sigma_\mu^2})$$

$$X = (X - \theta) + \theta$$

$$(X - \theta) \sim N(0, \lambda)$$

$$X \sim N(0, \lambda) + N(\mu, \nu) \sim N(\mu, \frac{\lambda + \nu}{\lambda\nu})$$

)

b) Remembering that:

$$m(x_{new}|\hat{x}) = \int f(x_{new}|\theta)\pi(\theta|\hat{x})d\theta$$

we know that $\pi(\theta|x)$ is the posterior distribution that is a $N(\mu_\theta^{'}, \lambda_\theta^{'})$:

$$\mu_\theta^{'} = w \cdot \mu + (1 - w) \cdot \bar{x}_n$$

$$\nu_\theta^{'} = \nu + N\lambda$$

and:

$$w = \frac{\nu}{\nu + N\lambda}$$

so it's easy now to compute the posterior predictive distribution:

$$m(x_{new}|x) = \int f(x_{new}|\theta)\pi(\theta|x)d\theta = \int \frac{\sqrt{\lambda}}{\sqrt{2\pi}} e^{-\frac{\lambda(x-\theta)^2}{2}} \cdot \frac{\sqrt{\nu_\theta^{'}}}{\sqrt{2\pi}} e^{-\frac{\nu_\theta^{'}(\theta - \mu_\theta^{'})^2}{2}} d\theta$$

$$\propto e^{-\frac{1}{2}(\lambda - \frac{1}{\lambda + \nu_\theta^{'}})x^2 + \frac{\mu_\theta^{'}}{\lambda + \nu_\theta^{'}}x}$$

because is the same integral we compute before.

So the posterior predictive distribution is a $N(a = (\lambda - \frac{1}{\lambda + \nu_\theta^{'}}); b = \frac{\mu_\theta^{'}}{\lambda + \nu_\theta^{'}})$

c)

We want that $P_\pi(-5 \le \theta \le 5) = 0.96$, and $\mu = 0$ so:

$$P_\pi(-5 \le \theta \le 5) = P_\pi(\frac{-5 - \mu}{\sigma} \le \frac{\theta - \mu}{\sigma} \le \frac{5 - \mu}{\sigma}) = P_\pi(\frac{-5}{\sigma} \le \ Z \le \frac{5}{\sigma}) = \Phi(\frac{5}{\sigma}) - \Phi(-\frac{5}{\sigma}) = 1 - 2\Phi(-\frac{5}{\sigma})$$

$$\Phi(-\frac{5}{\sigma}) = 0.02$$

```
qnorm(0.02)
```

```
## [1] -2.053749
```

$$\sigma = \frac{-5}{-2.053749} = 2.4346$$

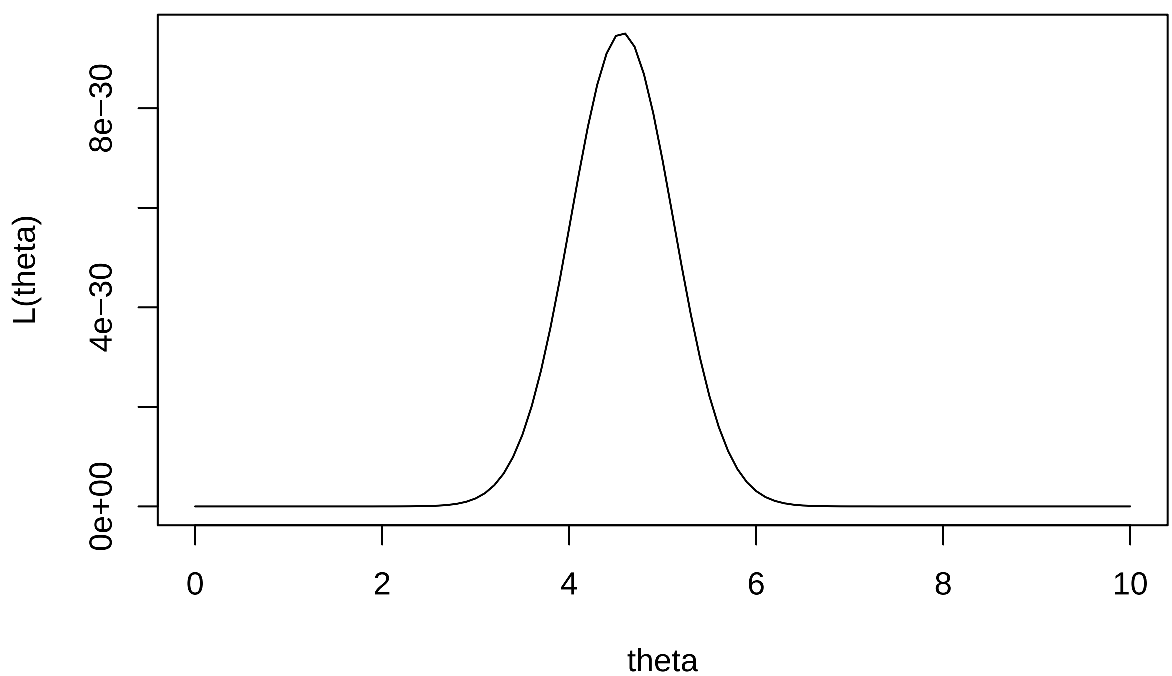
so the prior is a $N(0, 5.93)$.

d)

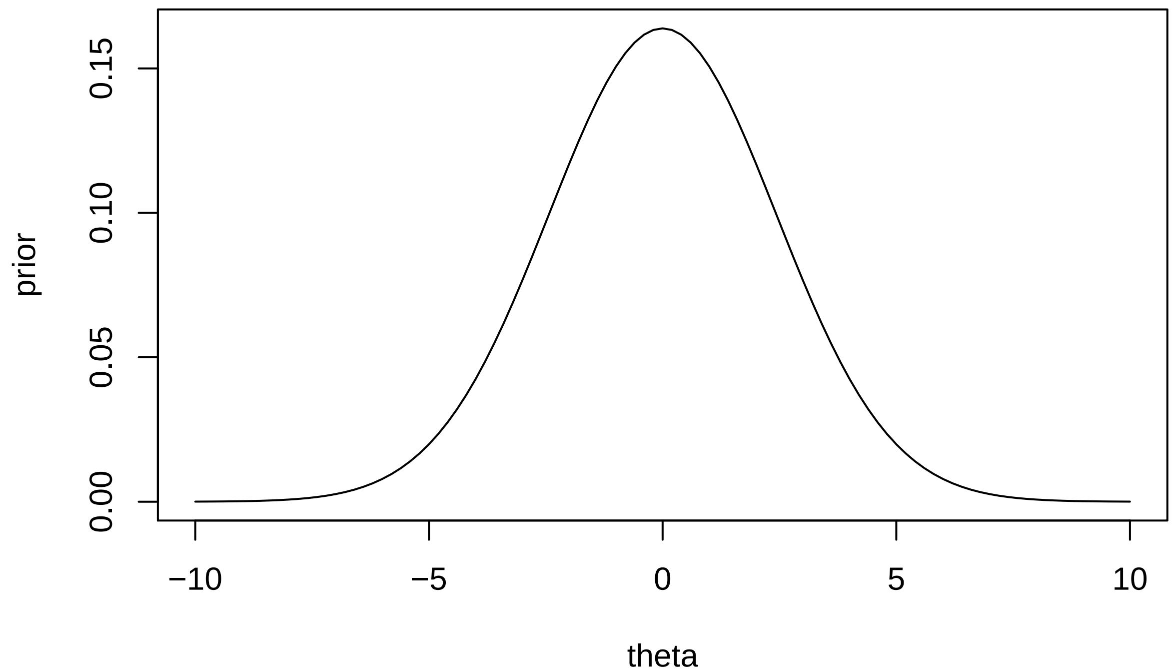
So:

$$\pi(\theta|\bar{X}) \propto L(\theta) \cdot \pi(\theta)$$

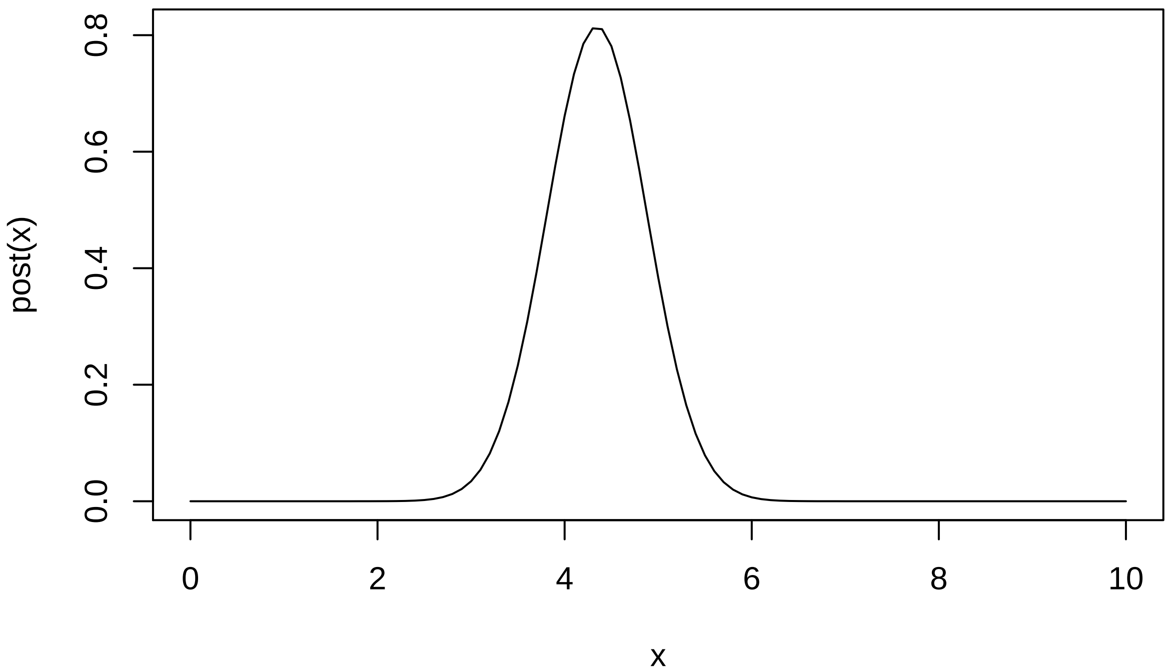
```
X.bar <- c(-1.25, 8.77, 1.18, 10.66, 11.81, -6.09, 3.56, 10.85, 4.03, 2.13)

L <- function(thet){
  prod=1
  for(x in X.bar){
    prod = prod*dnorm(x, mean=thet, sd=sqrt(3))
  }
  return(prod)
}
L <- Vectorize(L)
curve(L, xlab = 'theta', ylab = 'L(theta)', from = 0, to = 10)
```




```
pr <- function(thet) dnorm(thet, 0, (-5/qnorm(0.02)))
pr <- Vectorize(pr)

curve(pr, ylab = 'prior', xlab='theta', from = -10, to = 10)
```




```
jnt <- function(t) L(t)*pr(t)
m.x_bar <- integrate(jnt, -Inf, Inf)
post <- function(t) jnt(t)/m.x_bar$value
curve(post, from = 0, to = 10)
```


0.0
0.2
0.4
0.6
0.8
post(x)
0
2
4
6
8
10
x