Lab 3: Bayes Theorem

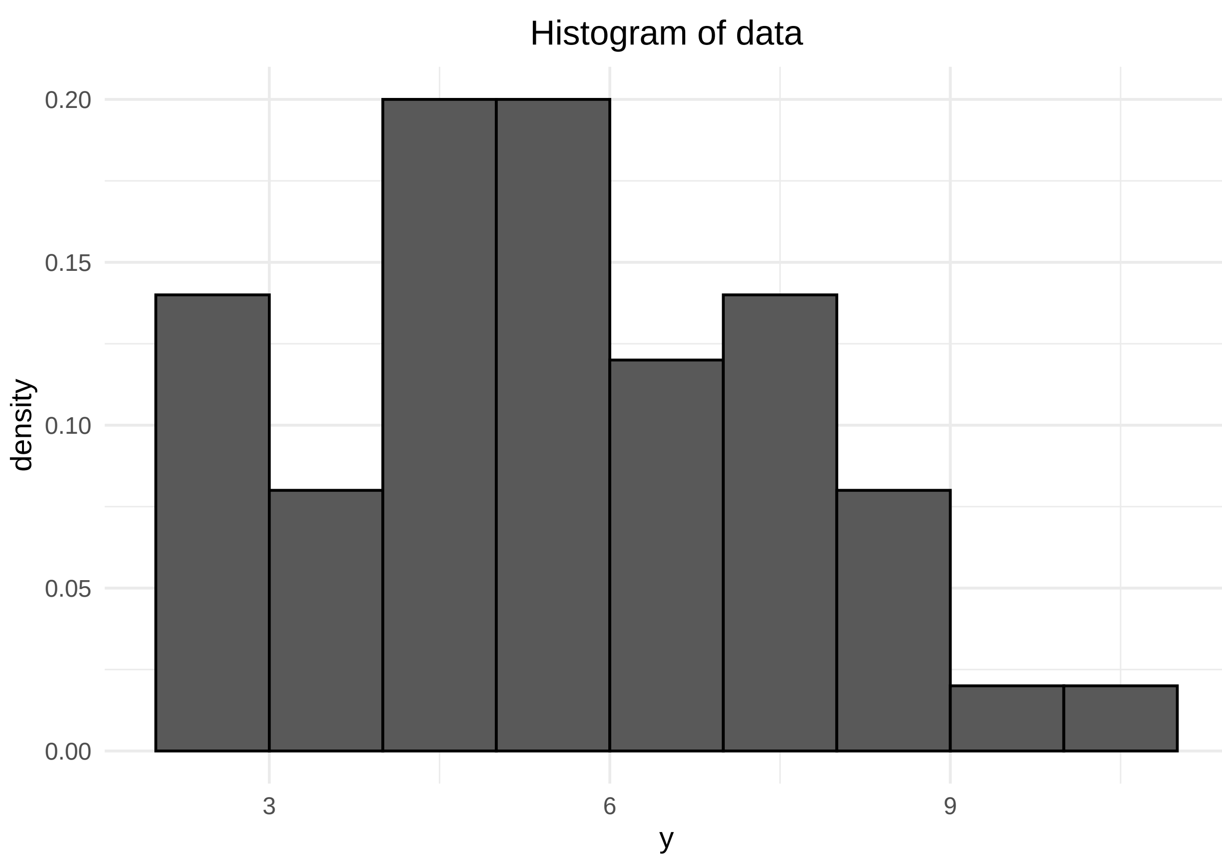
Team USA: Yuting Deng, Kendra Gilbertson, Lily Durkee, Bennett Hardy

2022-09-27

Contact:
Hermione.Deng@colostate.edu
kendra01@colostate.edu
L.Durkee@colostate.edu
Bennett.Hardy@colostate.edu

Preliminaries

1. Simulate 50 data points from a Poisson distribution with mean $\theta = 6.4$ to represent the data set.
2. Plot a histogram of the data with density on the y-axis.



The histogram function in R will graph abundance data, so we need to set `freq=F` to put density on the y-axis.

3. Set values for the prior mean (mu.prior) and standard deviation (sigma.prior).

```
## [1] 10.2
```

```
## [1] 0.5
```

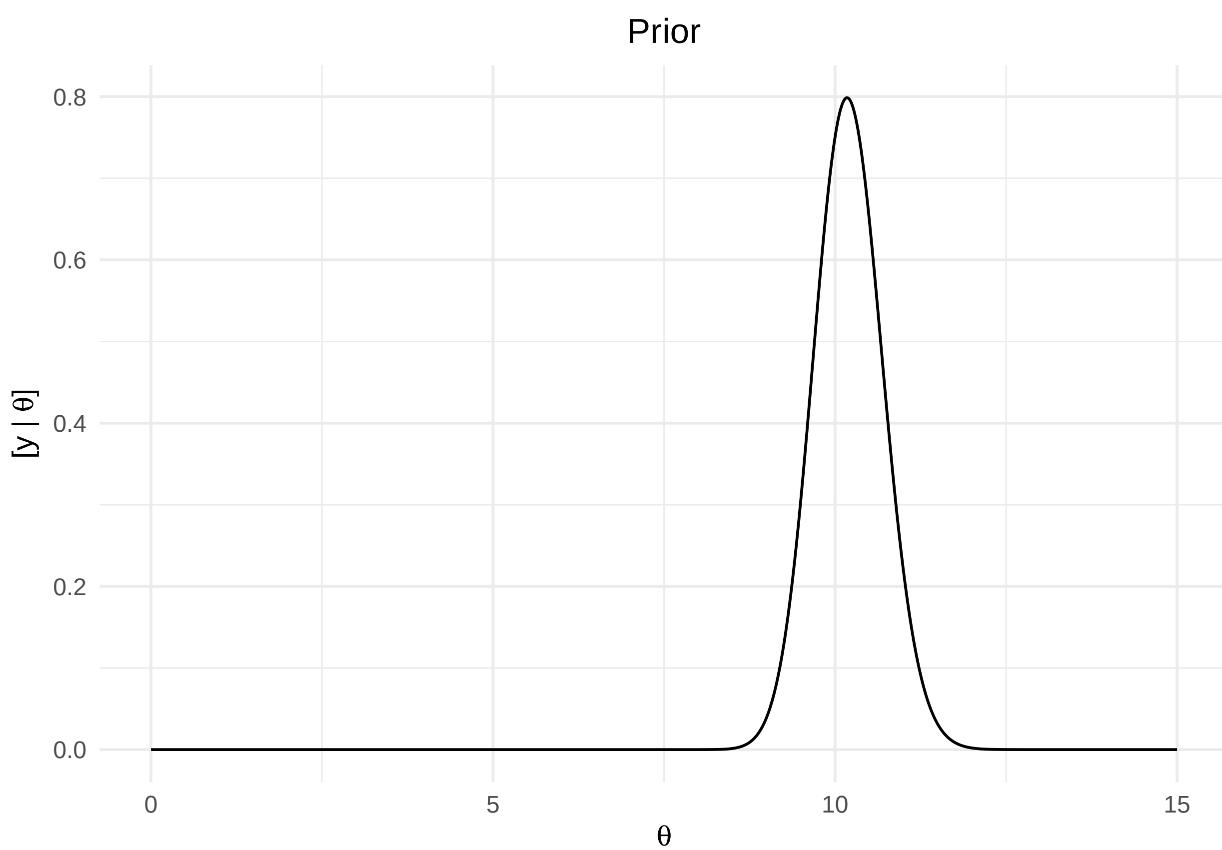
4. Set up a vector containing a sequence of values for θ, the mean number of invasive plants.

Prior distribution of θ

5. Write the mathematical expression for a gamma prior on θ

$$\theta \sim gamma(\alpha, \beta)$$
$$\theta \sim gamma(\mu^2/\sigma^2, \mu/\sigma^2)$$

6. Plot the prior distribution of θ.



7. Check your moment matching by generating 100,000 random variates from a gamma distribution with parameters matched to the prior mean and standard deviation.

```
## [1] 10.19569
```

```
## [1] 0.5000646
```

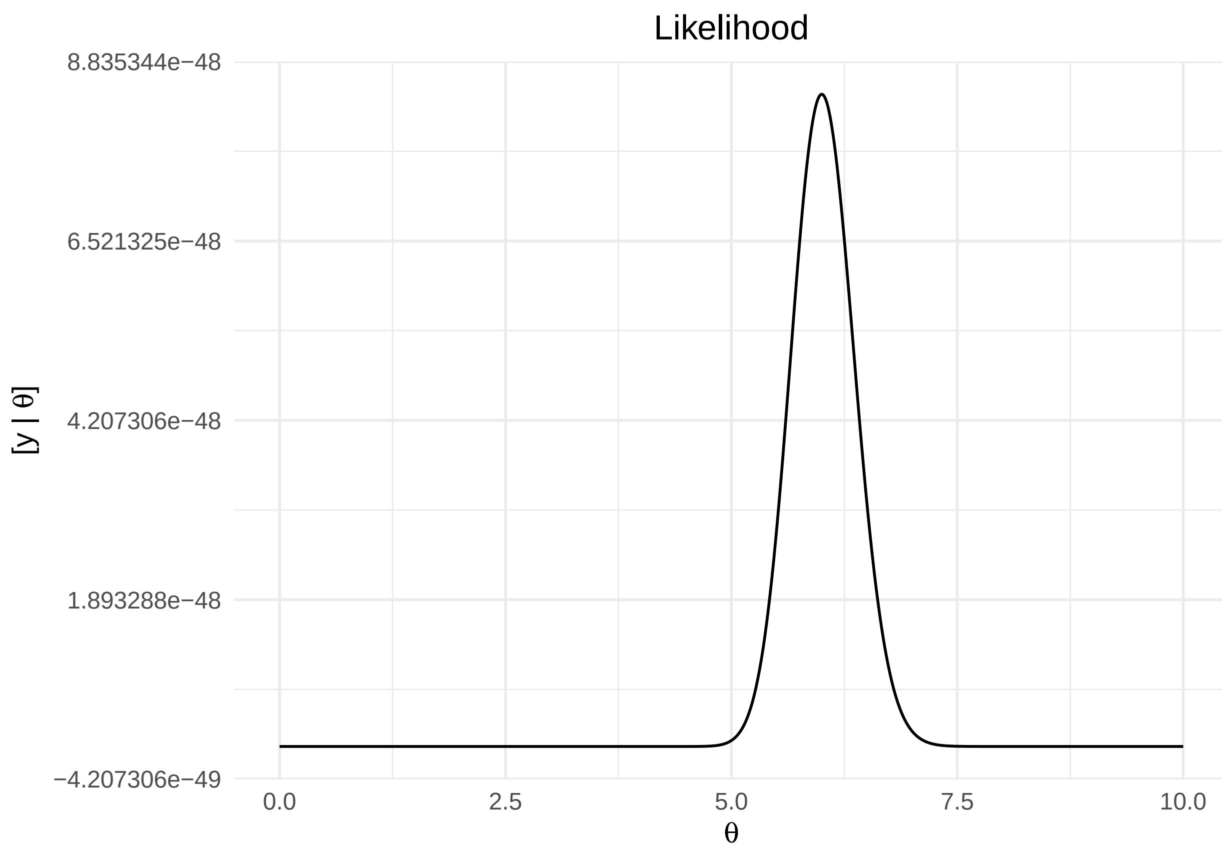
With a large sample size, the mean and variance will approximate the true mean and variance, which verifies the moment matching we performed.

Likelihood

8. What is the mathematical expression for the likelihood assuming that the data are conditionally independent?

$$\prod_{i=1}^{n} \text{Poisson}(y_i|\theta)$$

9. Plot the likelihood holding the data constant and varying θ.

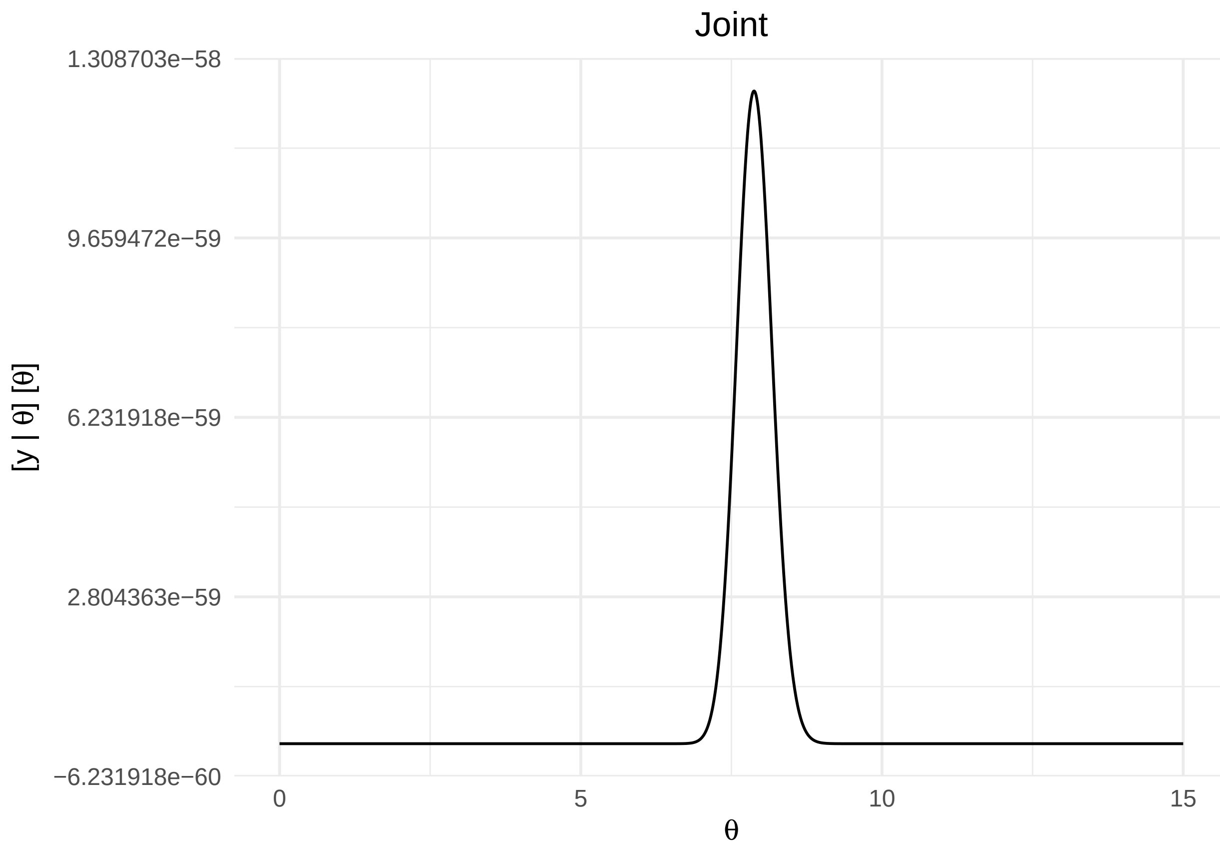


This plot is called likelihood profile, and the area under the curve does not sum to one. The relative inference is not affected whether or not we multiple the curve by a constant.

The joint distribution

10. What is the mathematical expression for the joint distribution?

$$\prod_{i=1}^{n} \text{Poisson}(y_i|\theta)\text{gamma}(\theta, \frac{10.2^2}{0.5^2}, \frac{10.2}{0.5^2})$$



The small number seems reasonable because we have multiplied to densities together, and multiplying numbers less than 1 gives us an even smaller number.

Marginal probability of the data

11. What is the mathematical expression for the marginal probability of the data [y]?

$$\int_{-\infty}^{\infty} \prod_{i=1}^{n} \text{Poisson}(y_i|\theta)\text{gamma}(\theta, \frac{10.2^2}{0.5^2}, \frac{10.2}{0.5^2})$$

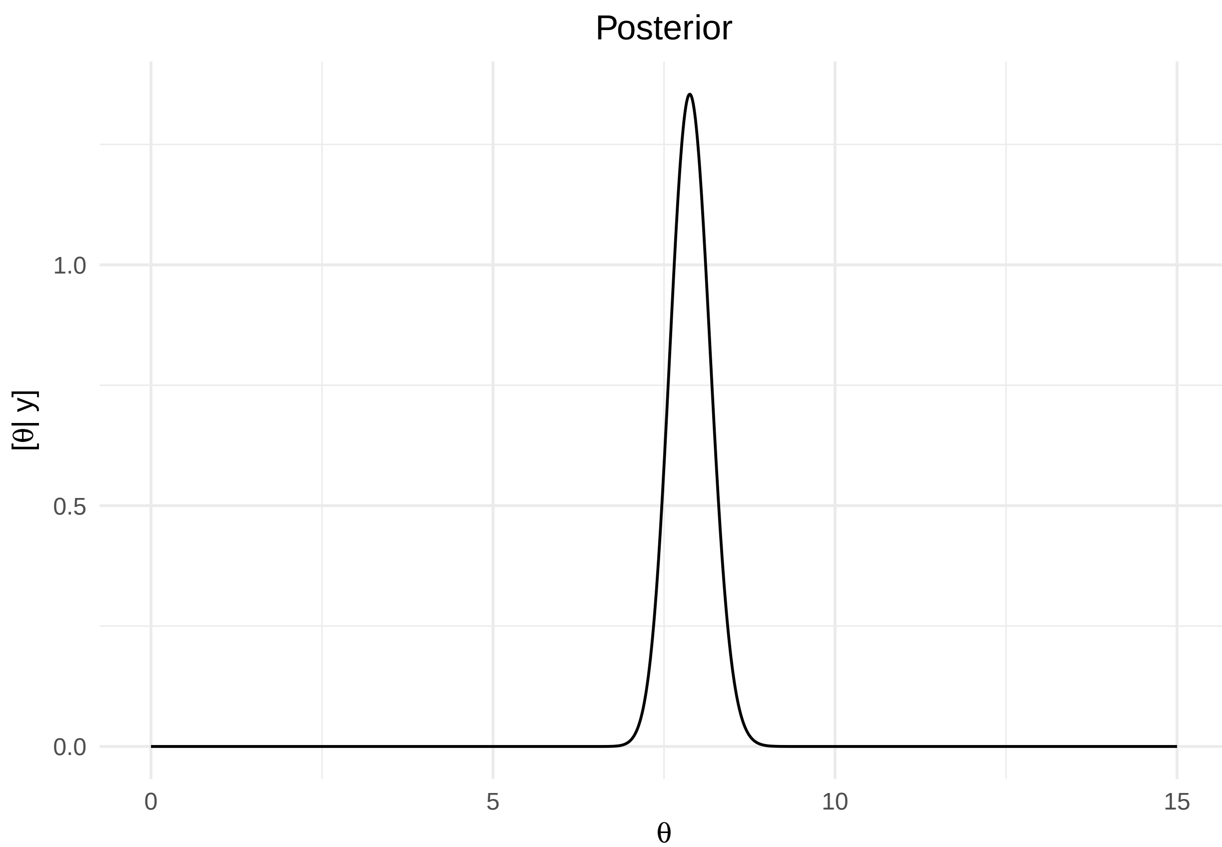
```
## [1] 9.203307e-57
```

To find the area under the curve we take the integral, or approximate the integral by summing increasingly thinner bars that make up the distribution. Here we multiple the height of each bar by our predetermined width (0.01) to find the area. Y is a random variable governed by a distribution until we collect the data, at which point it becomes a vector of values that evaluates to a scalar.

Posterior distribution

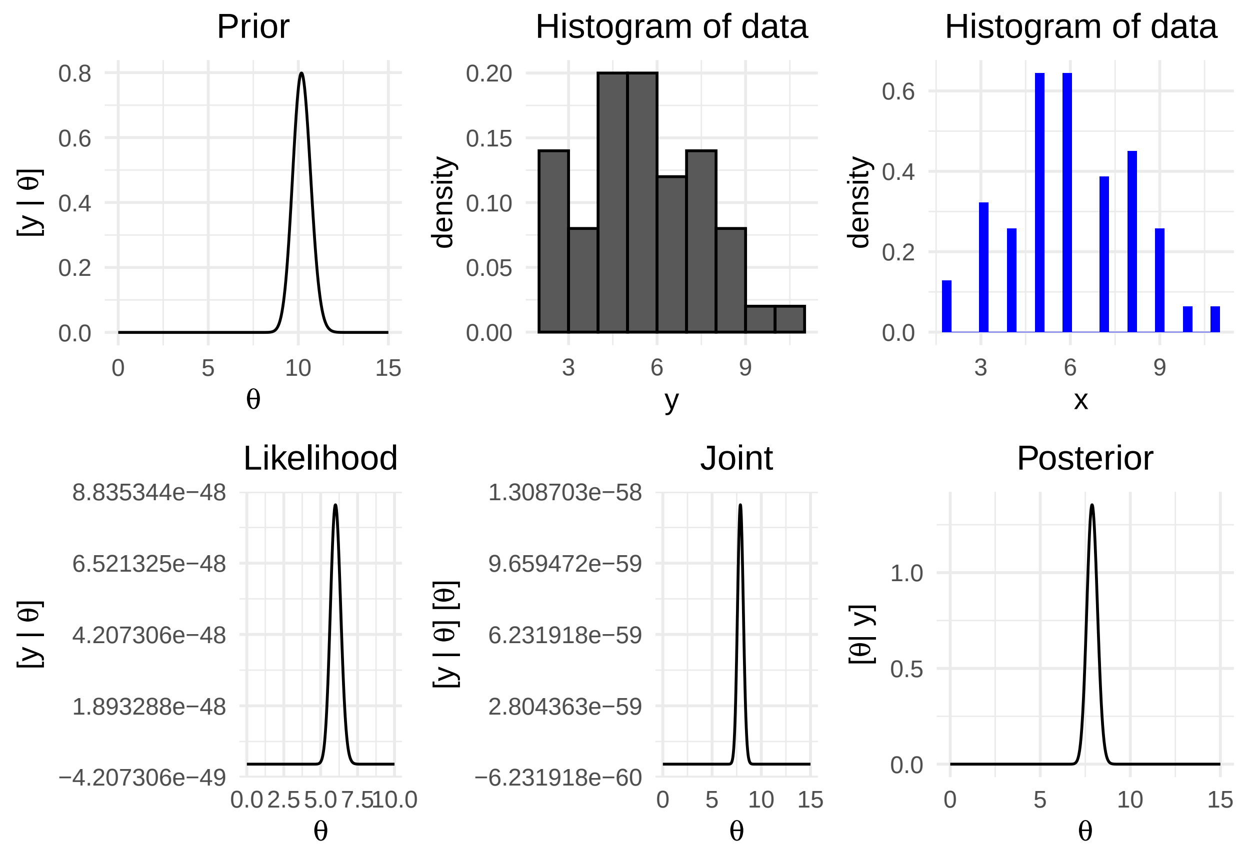
12. What is the mathematical expression for the posterior distribution?

$$[\theta|y] = \frac{\prod_{i=1}^{n} \text{Poisson}(y_i|\theta)\text{gamma}(\theta, \frac{10.2^2}{0.5^2}, \frac{10.2}{0.5^2})}{\int_{-\infty}^{\infty} \prod_{i=1}^{n} \text{Poisson}(y_i|\theta)\text{gamma}(\theta, \frac{10.2^2}{0.5^2}, \frac{10.2}{0.5^2})}$$



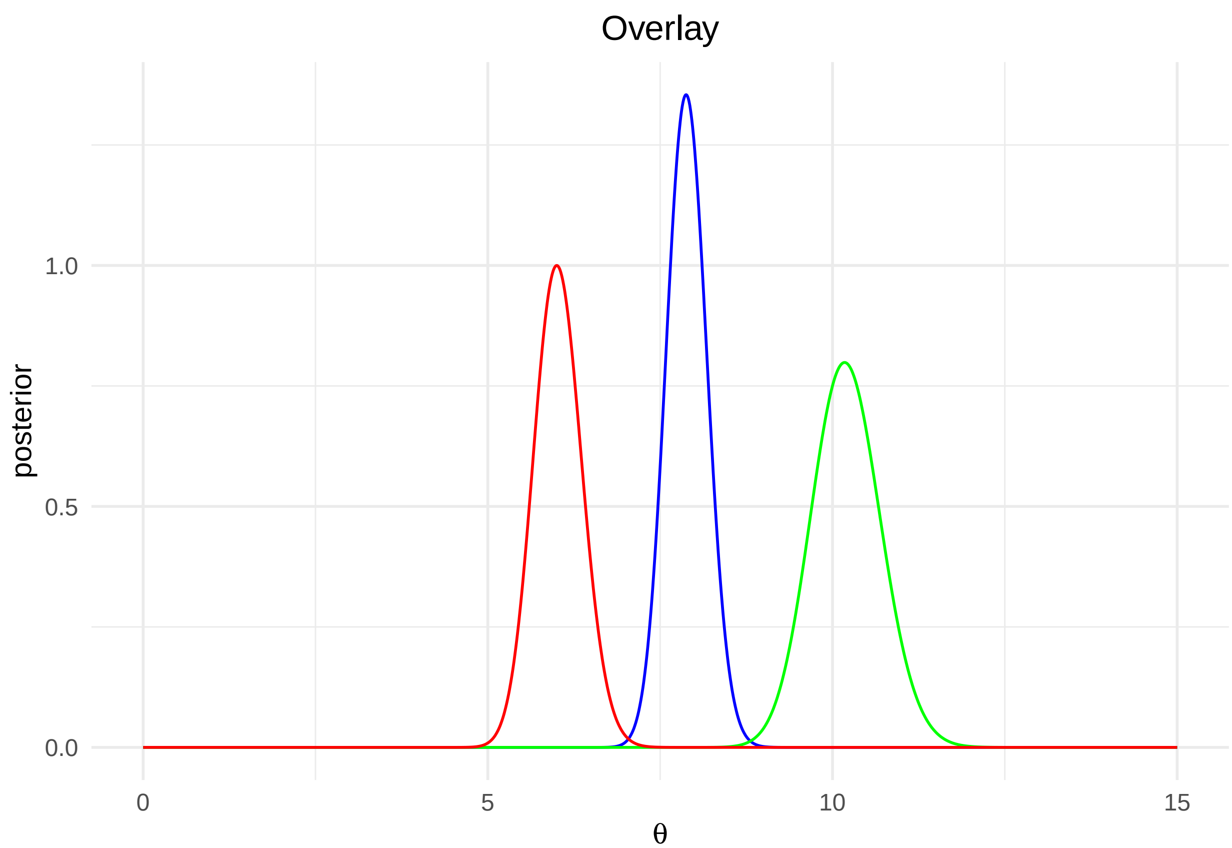
Putting it all together

13. Plot the prior, a histogram of the data, the likelihood, the joint, and the posterior in a six panel layout.

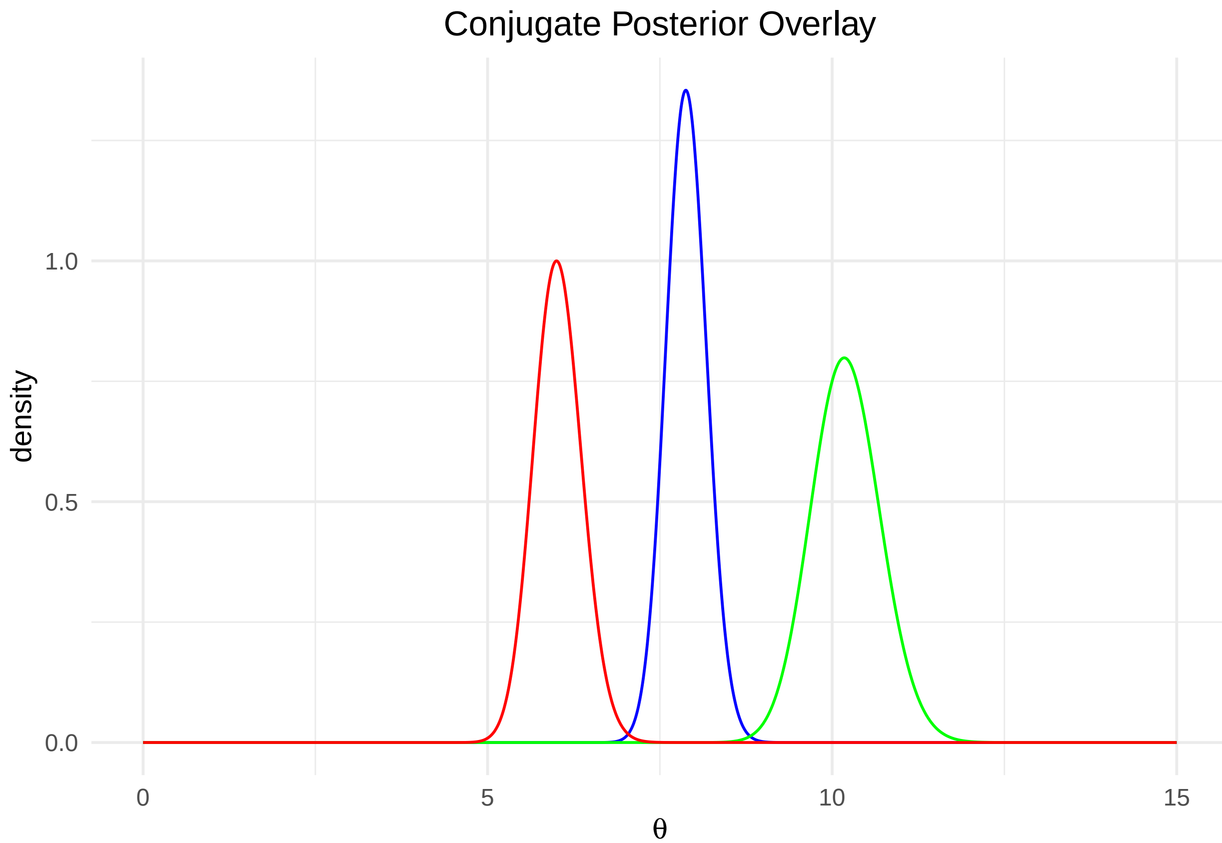


14. Overlay the prior, the likelihood, and the posterior on a single plot.

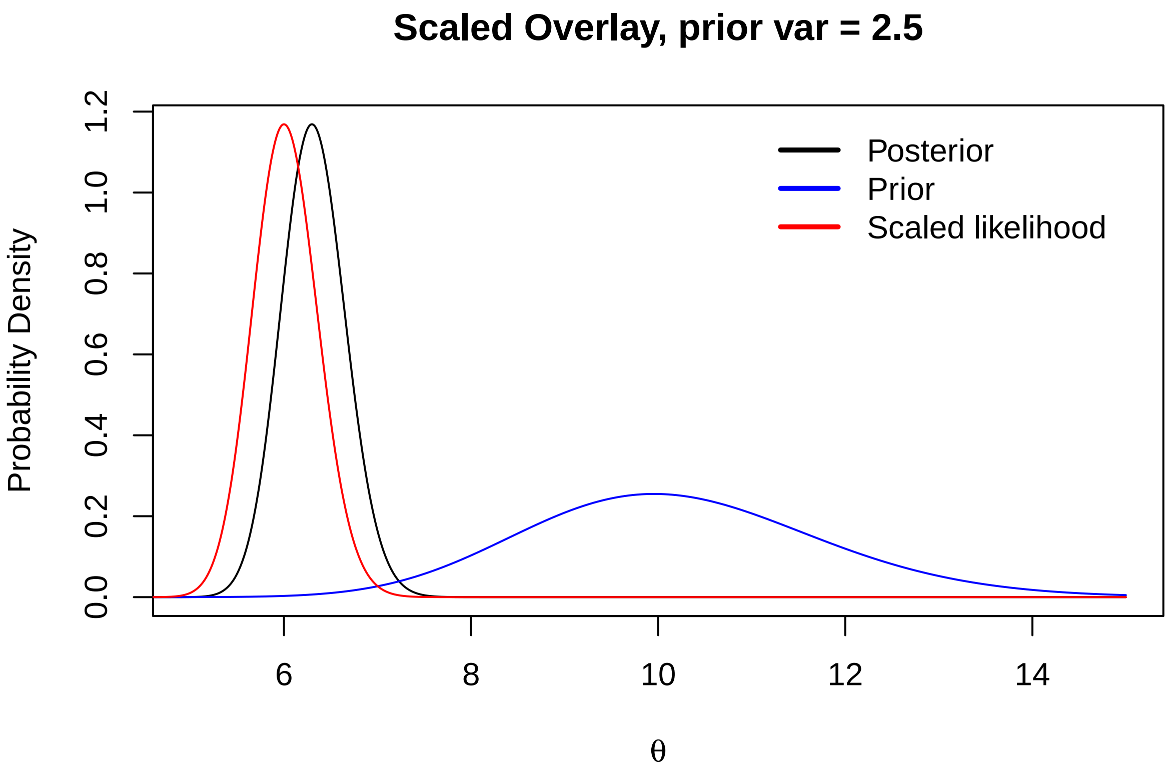
```
## [1] 8.414613e-48
```

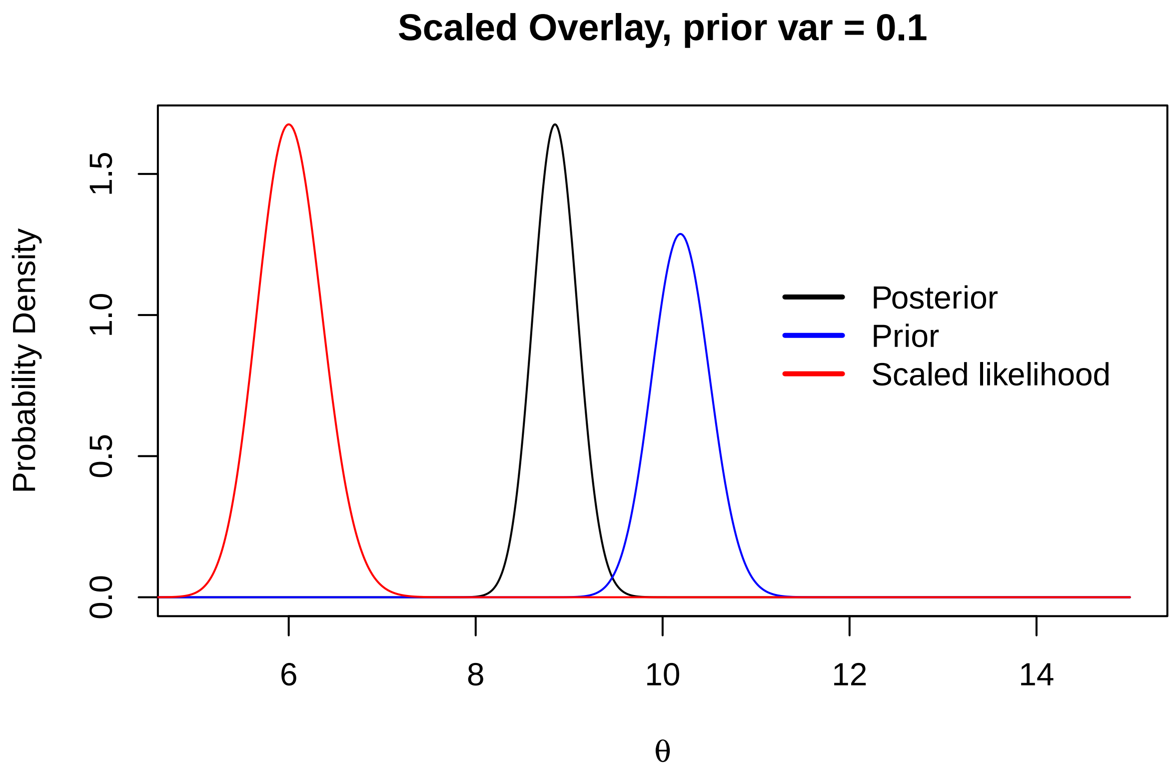


15. Check to be sure that everything is correct using the gamma-Poisson conjugate relationship.

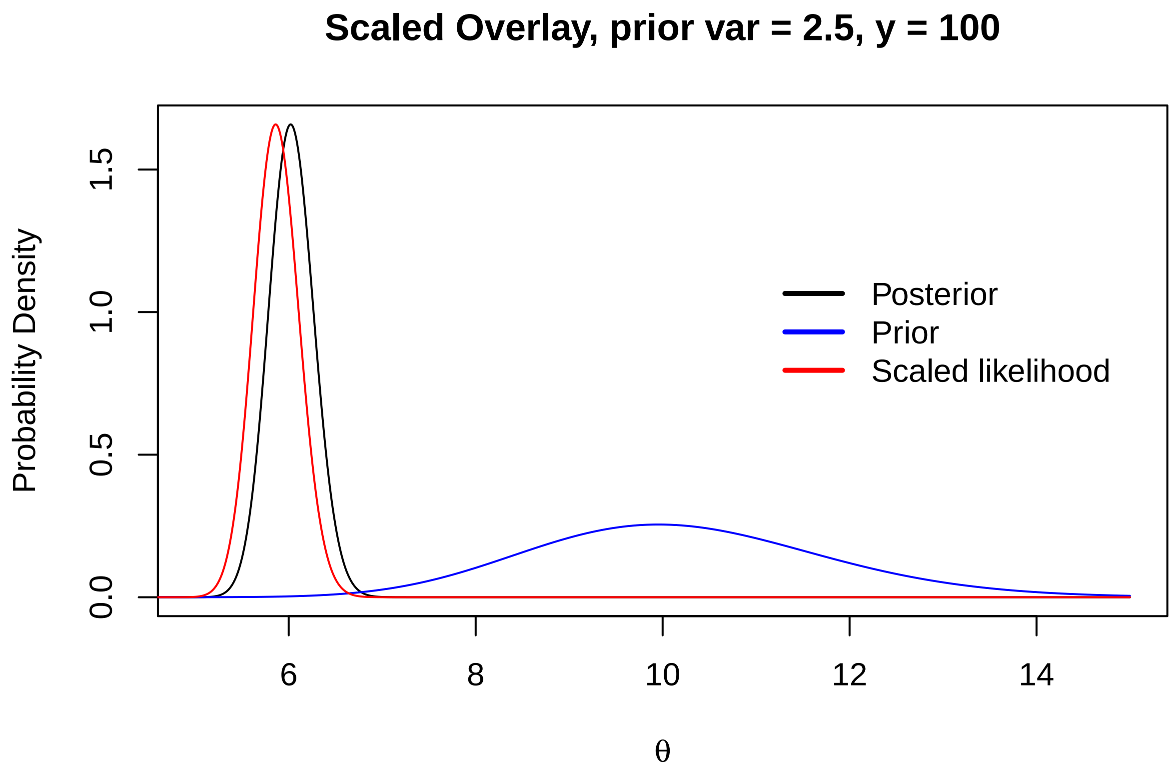


The likelihood profile for θ has much less dispersion than our histogram because we are using 1450 more data points. The more data we have, the less variance within our sample.


Scaled Overlay, prior var = 2.5
Posterior
Prior
Scaled likelihood
Probability Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
6
8
10
12
14
θ



Scaled Overlay, prior var = 0.1
Probability Density
0.0
0.5
1.0
1.5
6
8
10
12
14
θ
Posterior
Prior
Scaled likelihood



17. Gather some classmates and discuss the position of the prior, likelihood and posterior along the x axis and their variances.

Increasing the variance of the prior gave us a flatter distribution and gave the likelihood more weight in the posterior distribution. Decreasing the variance of the prior has the opposite effect, and the prior is more informative. Increasing the sample size decreased the variance in both the likelihood and the posterior.

Code

```r
knitr::opts_chunk$set(
    echo = FALSE,
    message = FALSE,
    warning = FALSE,
    attr.source = ".numberLines"
)
library(ggplot2)
library(gridExtra)
library(dplyr)
library(ggpubr)
library(mathjaxr)
set.seed(10)
data <- rpois(n=50, lambda=6.4)
y_df <- data.frame(data)
f_hist1 <- ggplot(y_df, aes(x=data))+
  geom_histogram(aes(y = ..density..), breaks=c(2,3,4,5,6,7,8,9,10,11), color = "black")+
  theme_minimal()+
  labs(title="Histogram of data", x="y")+
  theme(plot.title = element_text(hjust = 0.5))
f_hist1

f_hist2 <- ggplot(y_df, aes(x=data))+
  geom_histogram(aes(y = ..density..), fill = "blue")+
  theme_minimal()+
  labs(title="Histogram of data", x="x")+
  theme(plot.title = element_text(hjust = 0.5))
mu.prior <- 10.2
sigma.prior <- 0.5

mu.prior
sigma.prior
step <- .01
theta <- seq(0,15,step)
prior_fx <- function(theta, mu = mu.prior, sigma = sigma.prior) {
  prior <- dgamma(theta, rate=mu/sigma^2, shape=mu^2/sigma^2)
  return(prior)
}

prior <- prior_fx(theta, mu.prior, sigma.prior)
prior_df <- data.frame(cbind(theta, prior))

# plot(x=theta, y=prior)
f_prior <- ggplot(prior_df, aes(theta, prior)) +
  geom_line() +
  theme_minimal()+
  labs(x=expression(theta), y = expression(paste("[y | ", theta, "]")), title="Prior")+
  theme(plot.title = element_text(hjust = 0.5))


f_prior
beta <- mu.prior/sigma.prior^2
```

```r
alpha <- mu.prior^2/sigma.prior^2

check <- rgamma(n=100000, alpha, beta)

mean(check)
sd(check)
# function "like" generates the likelihood of the data given theta
# uses arguments: y = a vector of collected data
# theta = a vector of means

like_fx <- function(theta, y){
    y_theta = prod(dpois(y,theta))
  return(y_theta)
}

like_data = c()
for (i in 1:length(theta)){
  like_data[i] = like_fx(theta[i], data)
}

like_df = data.frame(theta, like_data)

f_likelihood <- ggplot(like_df, aes(x=theta, y=like_data))+
  geom_line()+
  theme_minimal()+
  labs(x=expression(theta), y = expression(paste("[y | ", theta, "]")), title="Likelihood")+
  theme(plot.title = element_text(hjust = 0.5))+
  xlim(0,10)

f_likelihood

#10
joint_theta = c()
for (i in 1:length(theta)){
  joint_theta[i] = like_fx(theta[i], data)*prior_fx(theta[i],mu.prior, sigma.prior)
}

joint_df <- data.frame(theta, joint=joint_theta)
f_joint <- ggplot(joint_df, aes(x=theta, y=joint))+
  geom_line()+
  theme_minimal()+
  labs(x=expression(theta), y=expression(paste("[y | ", theta, "] [",theta,"]")), title="Joint")+
  theme(plot.title = element_text(hjust = 0.5))

f_joint
sum(joint_theta)
posterior_df <- data.frame(theta, posterior = joint_df$joint/sum(joint_theta*0.01))
f_posterior <- ggplot(posterior_df, aes(x=theta, y=posterior))+
  geom_line() +
  theme_minimal()+
  labs(x=expression(theta), y=expression(paste("[", theta, "| y]")), title="Posterior")+
  theme(plot.title = element_text(hjust = 0.5))
```

```r
f_posterior
library(ggpubr)
ggarrange(f_prior, f_hist1, f_hist2, f_likelihood, f_joint, f_posterior, ncol = 3, nrow = 2)
max(like_df$like_data)

like_df <- like_df %>%
  mutate(likelihood.2 = like_data/max(like_df$like_data))

ggplot()+
  geom_line(posterior_df, mapping=aes(x=theta, y=posterior), color="blue")+
  geom_line(prior_df, mapping=aes(x=theta, y=prior), color="green")+
  geom_line(like_df, mapping=aes(x=theta, y=likelihood.2), color="red")+
  theme_minimal()+
  labs(x=expression(theta), title="Overlay")+
  theme(plot.title = element_text(hjust = 0.5))
#15
posterior_conj <- function(theta){
  #gamma distribution
  alpha_conj = alpha + sum(data)
  beta_conj = beta + 50
  prob_density = dgamma(theta, alpha_conj, beta_conj)
  return (prob_density)
}

posterior_df_conj = data.frame(theta, density = posterior_conj(theta))

ggplot()+
  geom_line(posterior_df_conj, mapping=aes(x=theta, y=density), color="blue")+
  geom_line(prior_df, mapping=aes(x=theta, y=prior), color="green")+
  geom_line(like_df, mapping=aes(x=theta, y=likelihood.2), color="red")+
  theme_minimal()+
  labs(x=expression(theta), title="Conjugate Posterior Overlay")+
  theme(plot.title = element_text(hjust = 0.5))

Like <- like_data

like <- function(y, theta){
  L <- numeric(length(theta))
  for (i in 1:length(theta)){
    L[i] <- prod(dpois(y,theta[i]))
  }
  return(L)
}
#### INCREASE TO 2.5 ########

#### increasing variance ####
sigma.prior1 <- 1.58 ## var = 2.5
prior1 <- function(theta, mu.prior, sigma.prior1){
  alpha = mu.prior^2/sigma.prior1^2
  beta = mu.prior/sigma.prior1^2
  priordist = dgamma(theta, alpha, beta)
  return(priordist)
}
```

```r
prior.prob1 <- prior1(theta,mu.prior,sigma.prior1)

joint1 <- function(theta, mu.prior, sigma.prior1, y){
  prior.prob <- prior_fx(theta,mu.prior,sigma.prior1)
  Like <- like(y,theta)
  j <- prior.prob * Like
  return(j)
}
Jdist1 <- joint1(theta, mu.prior, sigma.prior1, data)

marg1 <- sum(step*Jdist1)

post1 <- Jdist1/marg1

Like_sort1 <- sort(Like, decreasing = TRUE)
like.rescale1 <- Like / Like_sort1[1]
post_sort1 <- sort(post1, decreasing = TRUE)
rescale_like1 <- (like.rescale1 * post_sort1[1])


plot(theta, post1, main = "Scaled Overlay, prior var = 2.5", xlab = expression(theta), ylab = "Probabil
lines(theta, prior.prob1, col = "blue")
lines(theta, rescale_like1, col = "red")
legend(11,1.2, c("Posterior", "Prior", "Scaled likelihood"),
       lwd=c(2.5, 2.5, 2.5),col=c("black", "blue", "red"), bty = "n")


### DECREASE TO 0.1 ##############


sigma.prior2 <- 0.31 ## var = 0.1
prior2 <- function(theta, mu.prior, sigma.prior2){
  alpha = mu.prior^2/sigma.prior2^2
  beta = mu.prior/sigma.prior2^2
  priordist = dgamma(theta, alpha, beta)
  return(priordist)
}
prior.prob2 <- prior2(theta,mu.prior,sigma.prior2)

joint2 <- function(theta, mu.prior, sigma.prior2, y){
  prior.prob <- prior_fx(theta,mu.prior,sigma.prior2)
  Like <- like(y, theta)
  j <- prior.prob * Like
  return(j)
}
Jdist2 <- joint2(theta, mu.prior, sigma.prior2, data)

marg2 <- sum(step*Jdist2)

post2 <- Jdist2/marg2

Like_sort2 <- sort(Like, decreasing = TRUE)
like.rescale2 <- Like / Like_sort2[1]
```

```r
post_sort2 <- sort(post2, decreasing = TRUE)
rescale_like2 <- (like.rescale2 * post_sort2[1])



plot(theta, post2, main = "Scaled Overlay, prior var = 0.1", xlim = c(5,15), xlab = expression(theta), y
lines(theta, prior.prob2, col = "blue")
lines(theta, rescale_like2, col = "red")
legend(11,1.2, c("Posterior", "Prior", "Scaled likelihood"),
       lwd=c(2.5, 2.5, 2.5),col=c("black", "blue", "red"), bty = "n")


########## INCREASE TO 100 #############

set.seed(10)
y1 <- rpois(100, 6.4) # increase number of obs to 100

# set values for prior mean and sd
mu.prior <- 10.2
sigma.prior1 <- 1.58 ## var = 2.5

########## Prior Distribution of theta ###########
prior1 <- function(theta, mu.prior, sigma.prior1){
  alpha = mu.prior^2/sigma.prior1^2
  beta = mu.prior/sigma.prior1^2
  priordist = dgamma(theta, alpha, beta)
  return(priordist)
}
prior.prob1 <- prior1(theta,mu.prior,sigma.prior1)

######## The likelihood #########
# function "like" generates the likelihood of the data given theta.
# uses arguments: y = a vector of collected data
#                 theta = a vector of means of number of invasive plants
like1 <- function(y1, theta){
  L <- numeric(length(theta))
  for (i in 1:length(theta)){
    L[i] <- prod(dpois(y1,theta[i]))
  }
  return(L)
}
Like1 <- like1(y1, theta)

######### The Joint Distribution ###########
joint1 <- function(theta, mu.prior, sigma.prior1, y1){
  prior.prob <- prior_fx(theta,mu.prior,sigma.prior1)
  Like1 <- like(y1, theta)
  j <- prior.prob * Like1
  return(j)
}
Jdist11 <- joint1(theta, mu.prior, sigma.prior1, y1)

########### The Marginal Probability of the data #############
```

```
marg11 <- sum(step*Jdist11)

########## The Posterior Distribution ####################
post11 <- Jdist11/marg11

Like_sort1 <- sort(Like1, decreasing = TRUE)
like.rescale1 <- Like1 / Like_sort1[1]
post_sort1 <- sort(post11, decreasing = TRUE)
rescale_like1 <- (like.rescale1 * post_sort1[1])


# make overlay plot
plot(theta, post11, main = "Scaled Overlay, prior var = 2.5, y = 100", xlim = c(5,15), xlab = expression
lines(theta, prior.prob1, col = "blue")
lines(theta, rescale_like1, col = "red")
legend(11,1.2, c("Posterior", "Prior", "Scaled likelihood"),
       lwd=c(2.5, 2.5, 2.5),col=c("black", "blue", "red"), bty = "n")
# this R markdown chunk generates a code appendix
```