Worksheet 3 Group Task Ex 4

Andrea Staub

2024-04-07

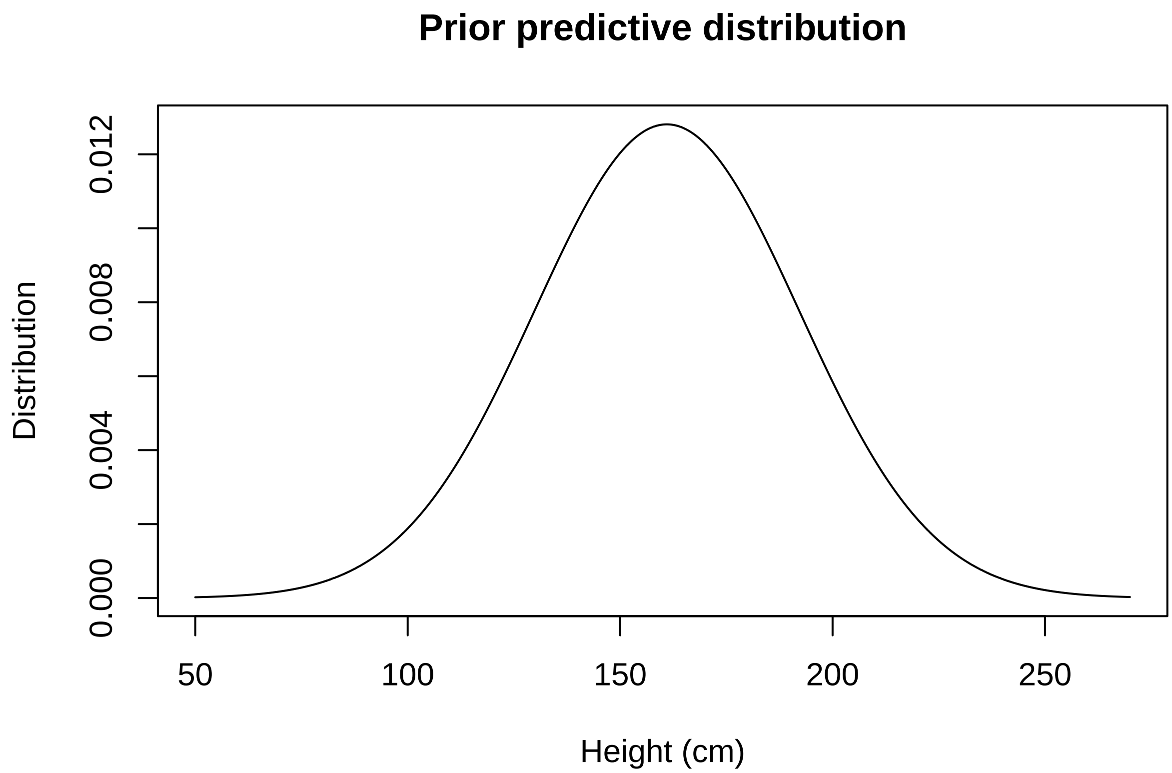
Exercise 4

Apply analytical formulas derived in Exercise 3 above to the vector of height (cm) measurements 166, 168, 168, 177, 160, 170, 172, 159, 175, 164, 175, 167, 164 of 13 Swiss females. Assume that $y_1, ..., y_n$ are observations generated by N(m, κ^{-1}) distribution with $\kappa = 1/900$. Moreover, assume a N(μ, λ^{-1}) prior for m with $\mu = 161$ and $\lambda = 1/70$.

a) Plot the prior predictive distribution for one observation y and compute its expectation and standard deviation. Estimate P[y > 200] for one future observation of Height.

The prior predictive distribution of one future observation y is N(μ, λ^{-1}+κ^{-1}).

```
my_mu <- 161
my_lambda <- 1/70
my_kappa <- 1/900
my_seq <- seq(50, 270, by = 0.01)
plot(my_seq, dnorm(my_seq, mean = my_mu, sd = sqrt(1/my_lambda+1/my_kappa)), type = "l", main = "Prior p
```




```
expect <- my_mu
stand_dev <- sqrt(1/my_lambda + 1/my_kappa)
prob <- 1-pnorm(200, mean = my_mu, sd = sqrt(1/my_lambda + 1/my_kappa))
```

The expectation is 161 and the standard deviation is 31.145.

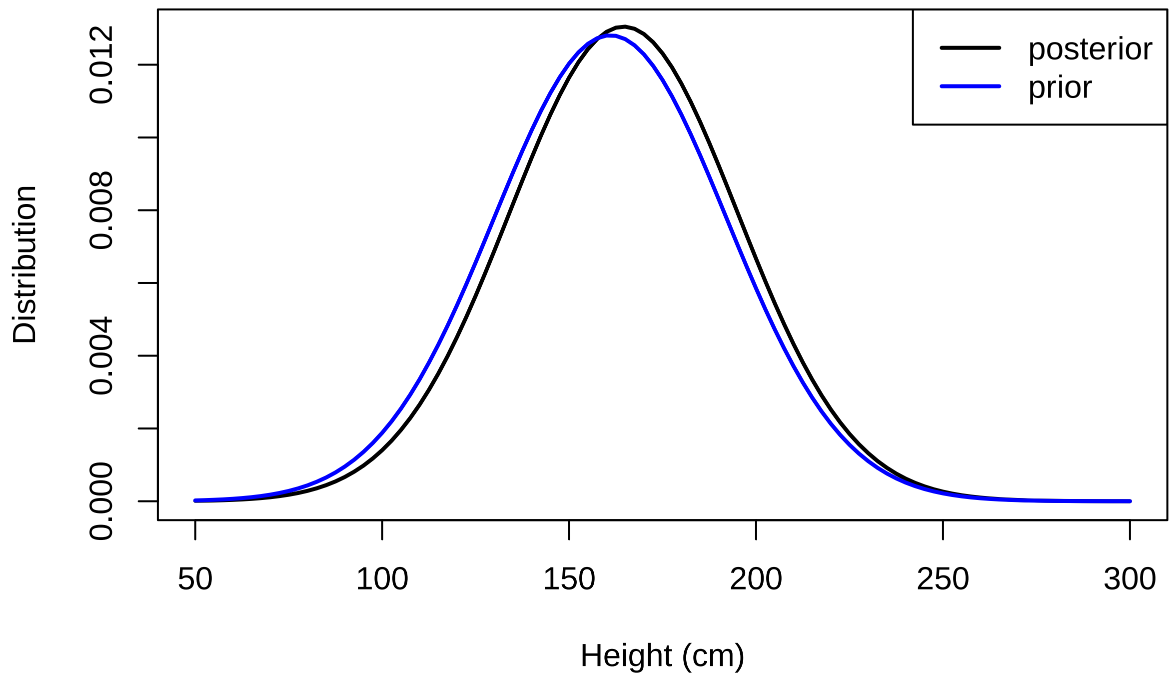
P[y >200] for one future observation of Height is 0.105.

b) Plot the posterior predictive distribution for one future observation y_{n+1} given that $y_1, ..., y_n$ have been observed and compute its expectation and standard deviation. Estimate P[$y_{n+1} >$ 200$y_1, ..., y_n$] for one future observation y_{n+1} of Height.

```
obs <- c(166, 168, 168, 177, 160, 170, 172, 159, 175, 164, 175, 167, 164)
n <- length(obs)
mean_obs <- mean(obs)

mu_post <- (my_kappa*n*mean_obs+my_lambda*my_mu)/(n*my_kappa+my_lambda)
lambda_post <- n*my_kappa+my_lambda

curve(dnorm(x, mean = mu_post, sd = sqrt(1/lambda_post+1/my_kappa)), xlab = "Height (cm)", ylab = "Dist
curve(dnorm(x, mean = my_mu, sd = sqrt(1/my_lambda+1/my_kappa)), from = 50, to = 300, lwd = 2, add = TRU
legend("topright", legend = c("posterior", "prior"), col = c("black", "blue"), lwd = 2)
```




```
prob_post <- 1-pnorm(200, mean = mu_post, sd = sqrt(1/lambda_post+1/my_kappa))
```

The expectation is 164.558 and the standard deviation is 30.575.

P[y_{n+1} >200] for one future observation of Height is 0.123.

c) Compare the results obtained for predictive distribution with those obtained for the posterior in Exercise 4 of Worksheet 2. Discuss how much posterior, prior predictive, and posterior predictive distributions differ.

In Exercise 4 of Worksheet 2 the posterior distribution of m$|y_1, ..., y_n$ was derived. Whereas in this worksheet's exercise 4 the posterior distribution of $y_{n+1}|y_1, ..., y_n$ was derived. The values for the mean of the posterior predictive distribution are the same (164.558). However, the standard deviations differ with much larger values for the posterior distribution of y_{n+1}.