Advanced Topics in Artificial Intelligence - Exercise 04

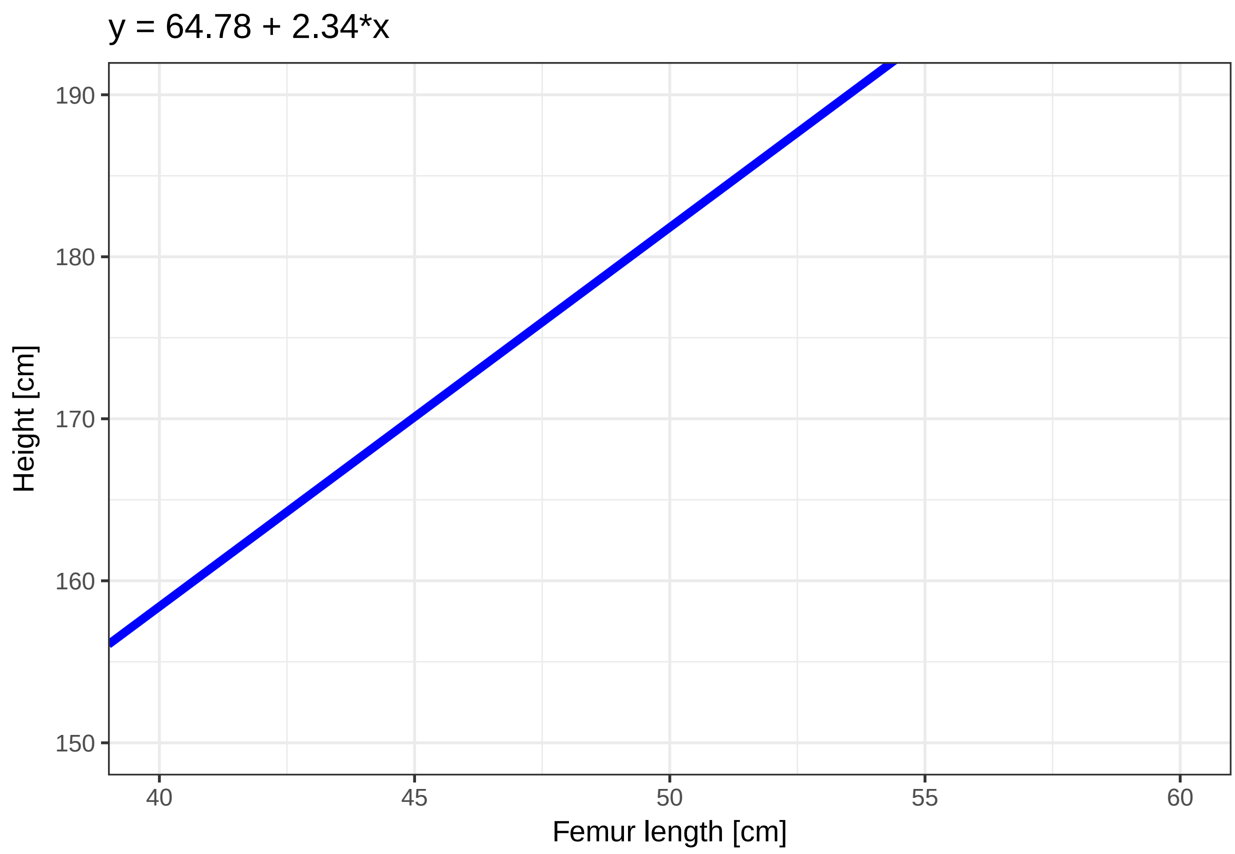
Augusto Magalhães, 335766

```
##       femur   height
## 1: 44.00332 167.9342
## 2: 49.40933 181.1709
## 3: 46.46925 173.1320
## 4: 49.40001 180.4553
## 5: 46.05390 172.5348
## 6: 44.23563 167.9825
## 7: 51.99498 186.1566
## 8: 41.86718 162.9707
```

Maximum likelihood estimate

$\mathbf{w}_{ML}$ equals to

```
## [1] 64.781587  2.340669
```



The result above came from solving the maximum likelihood to determine $\mathbf{w}$. The following equation taken from Bishop (2006) was used, where the *design matrix* was a 2 column matrix with $\phi_0(\mathbf{x}) = 1$ and $\phi_1(\mathbf{x}) = \mathbf{x}$.

$$\mathbf{w}_{ML} = (\mathbf{\Phi}^T \Phi)^{-1} \Phi^T \mathbf{t}$$

Batch Bayesian estimate

In the Forensic Arthropology field, anthropologists attempt to identify aspects of an individual such as race, sex and height from human skeletal.

One of the estimations done is when using the femur, where measuring the length of the femur bone in centimetres, then multiplying this legth by 2.6 and adding 65 to it approximates the height of the person, as descrbibed in *Forensic Anthropology: Height Estimation* (link at the end).

In contrast to classical calibration (no prior), using a informative prior distribution for the relationship between the femur length and the person's height can produce a Bayes estimator, so we take:

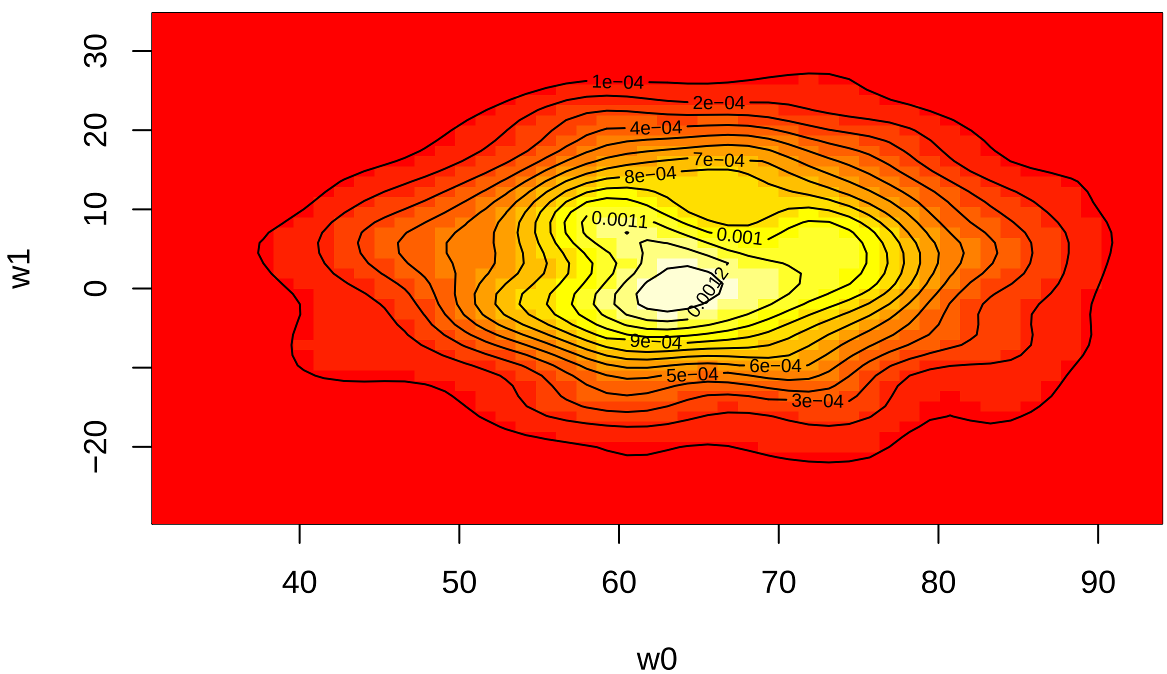
$$\mathbf{w} \sim \mathcal{N}(\mathbf{m}_0, \mathbf{S}_0 = \alpha^{-1}\mathbf{I})$$

$$\mathbf{w} = (65, 2.6); \alpha = 0.01$$

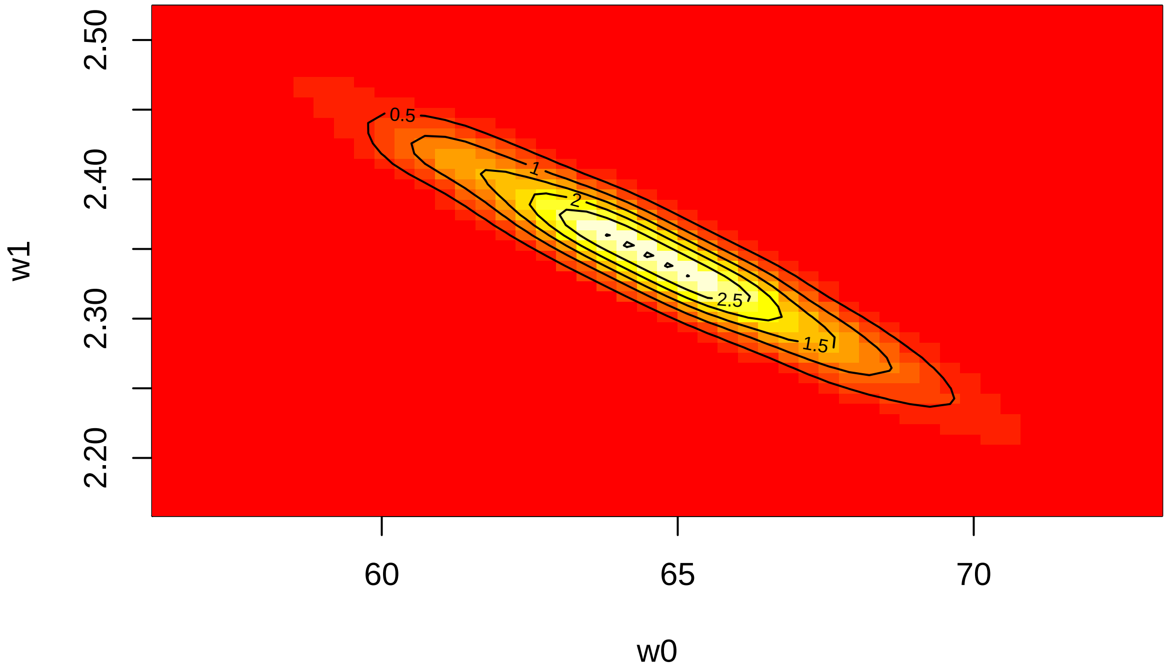
Since no value for α is described in that reference, we set it to 0.01 at a first glance.

Many other aspects can take informative priors from individuals from the broadly used Terry, WWII and Forensic Databank samples, depending on the problem being addressed.

Our distribution of the prior for the coefficients is drawn bellow.



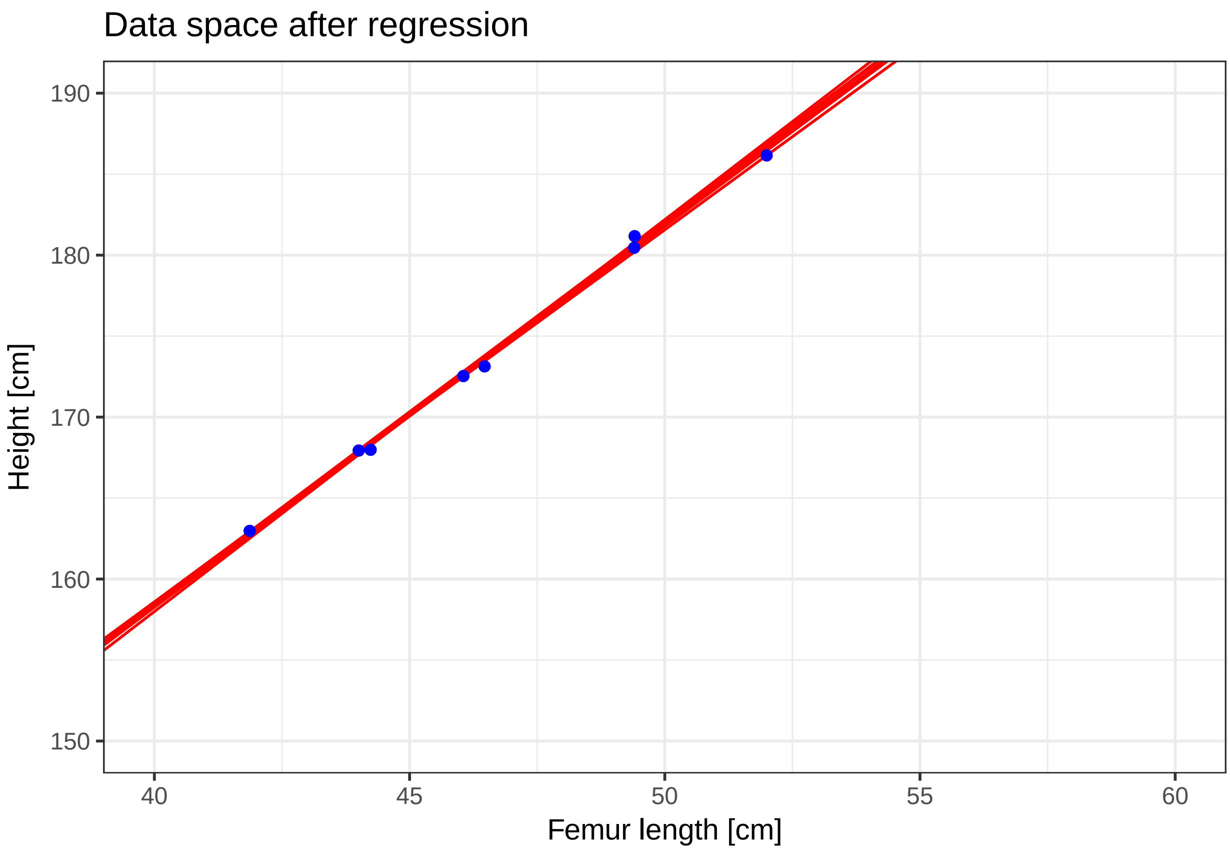
After using the 8 points provided, the posterior distribution is as follows:



Notice how much the range for w1 has shortened after the Bayesian linear regression. The plot scale had to be zoomed in order to allow further visualization.

The points observed fitted well the resultant regression line. The parameter α, when changed to a greater value, results in a data space much more stricted for the regression lines after fitting them with data.

From that, one can describe the data space for which there can be drawn regression lines for the problem.

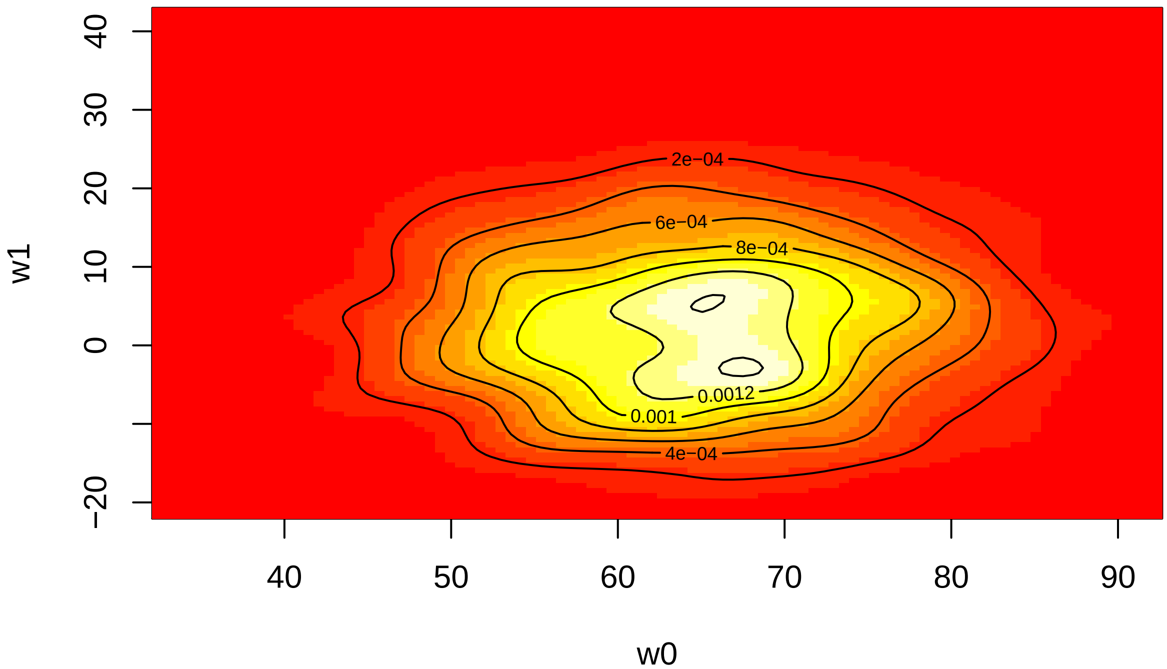

Data space after regression
Height [cm]
190
180
170
160
150
40
45
50
55
60
Femur length [cm]

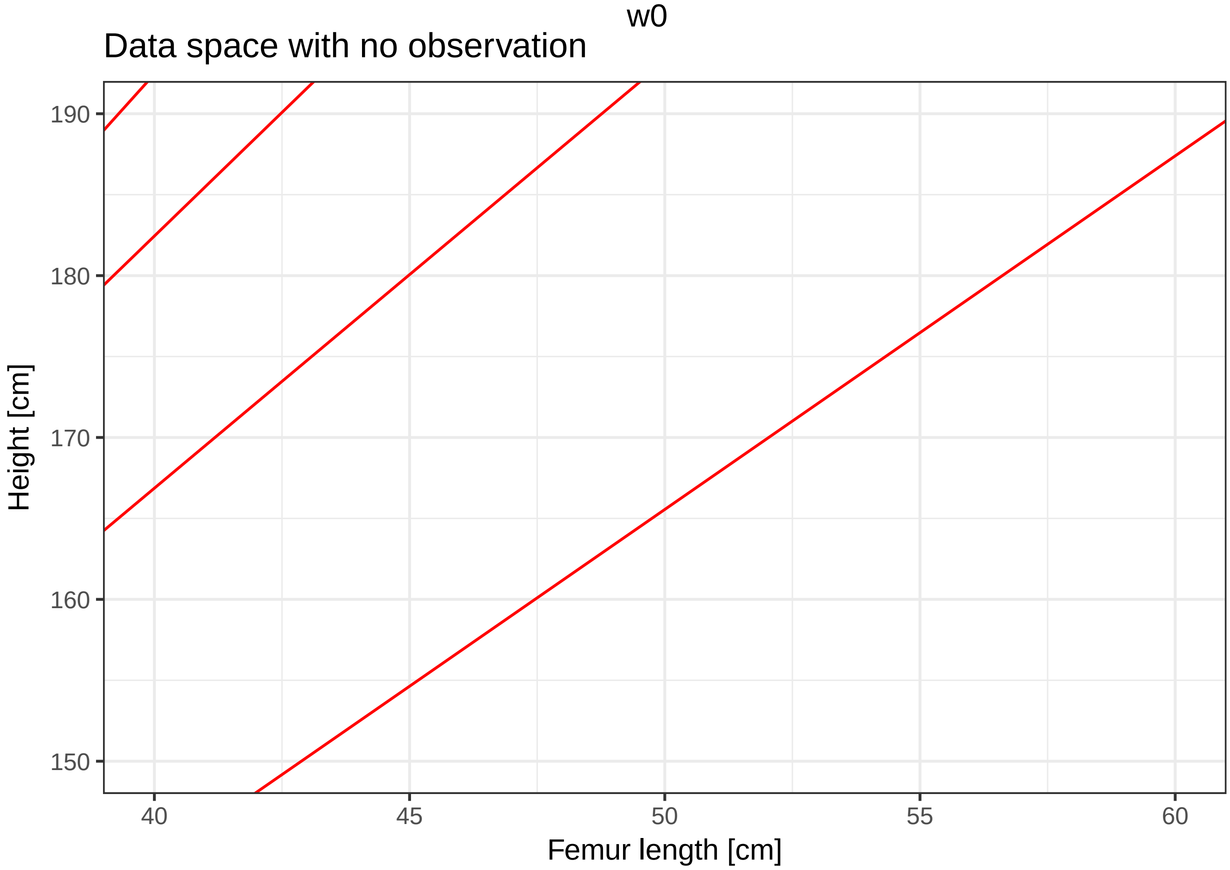
Sequential Bayesian estimate

After performing the Sequential Bayesian estimate, the same output from the Batch approach was reached.

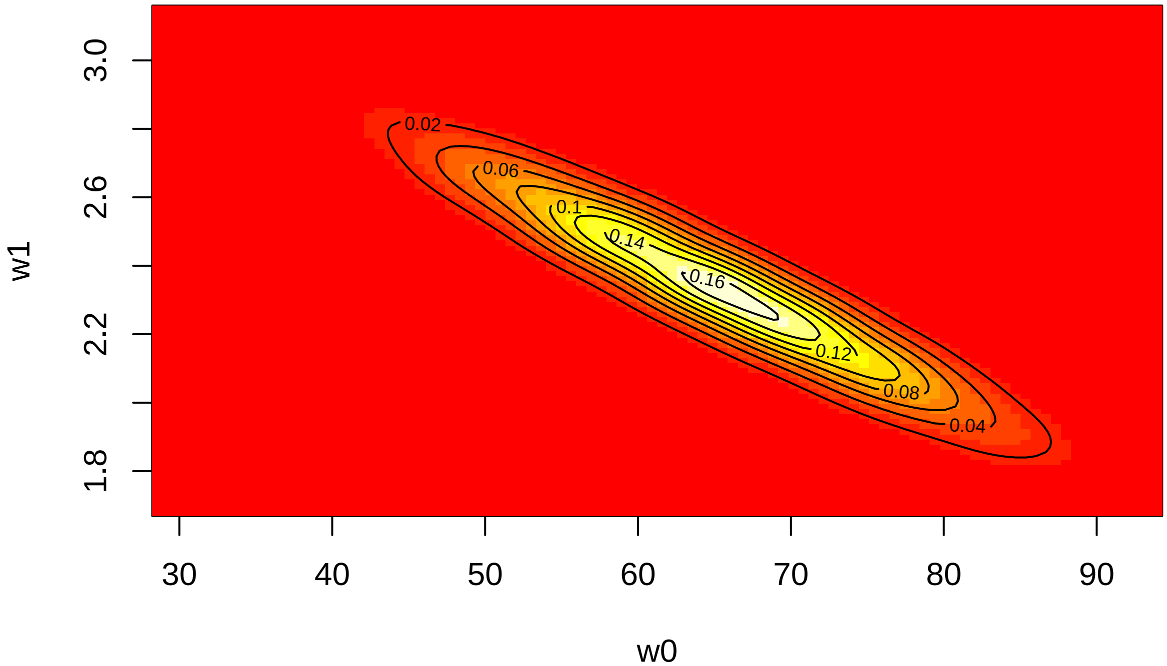
Some of te intermediate solutions were:

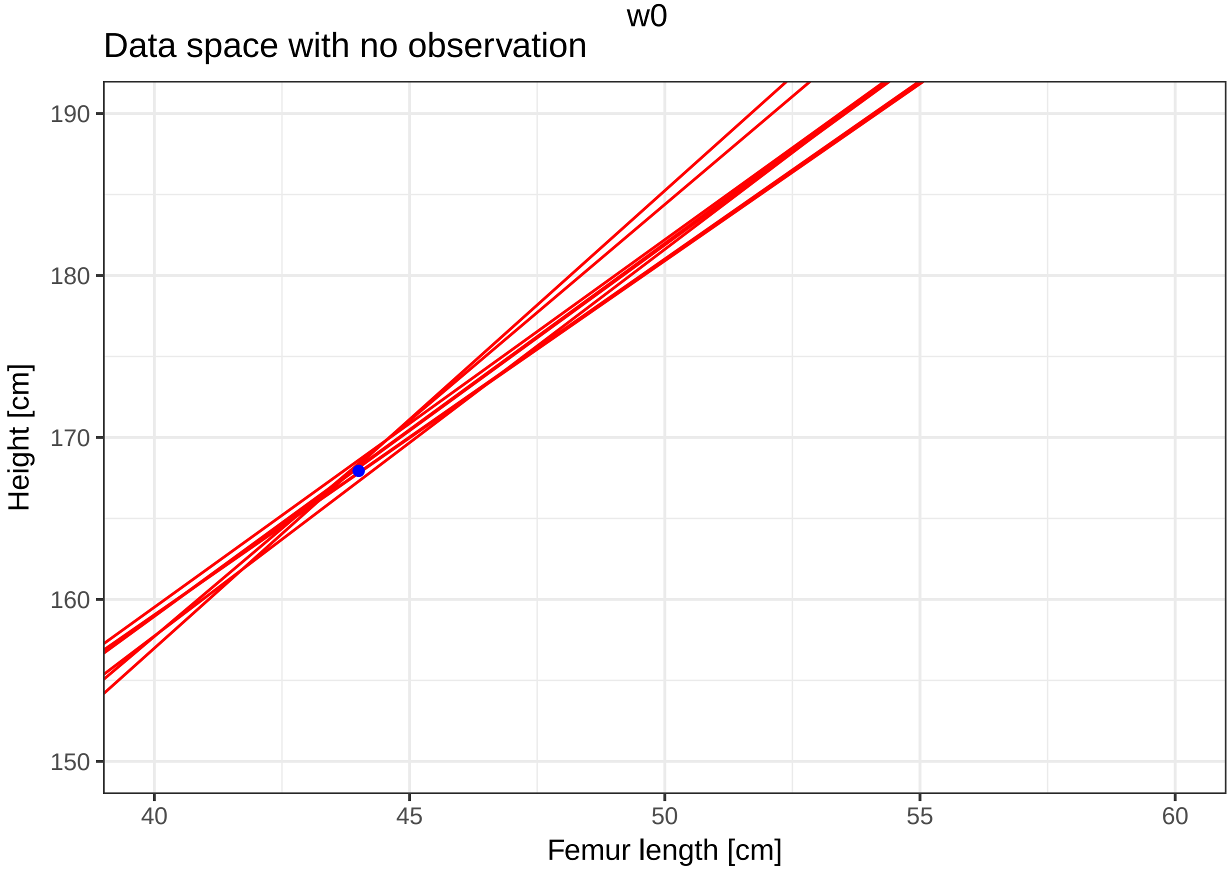
0 observation prior



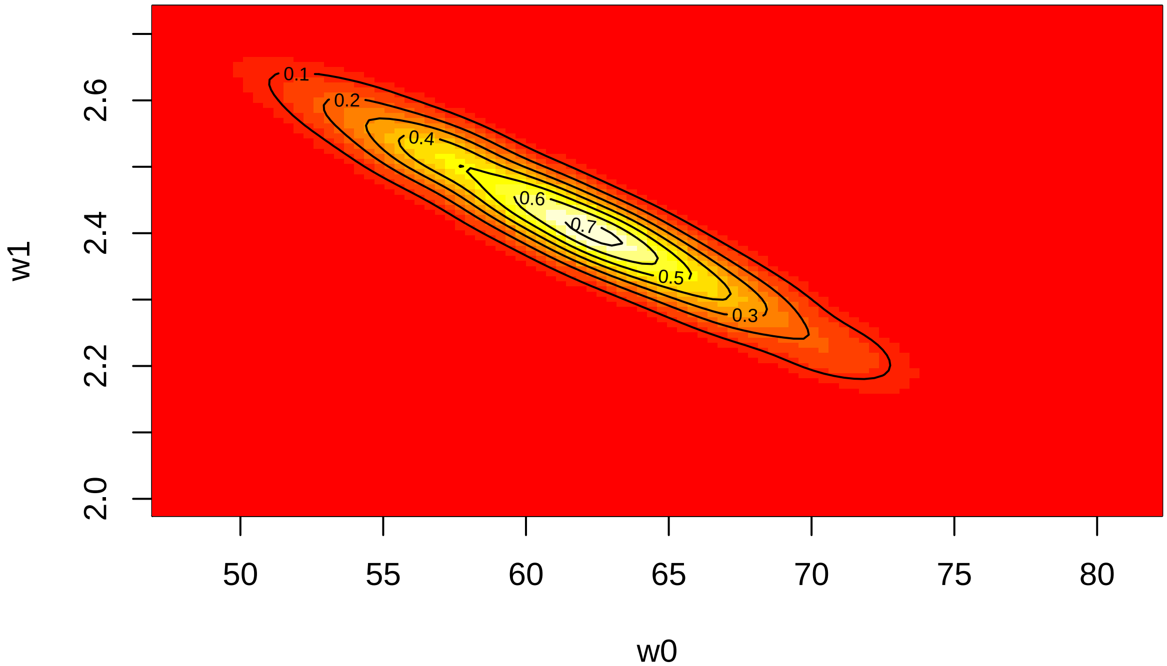


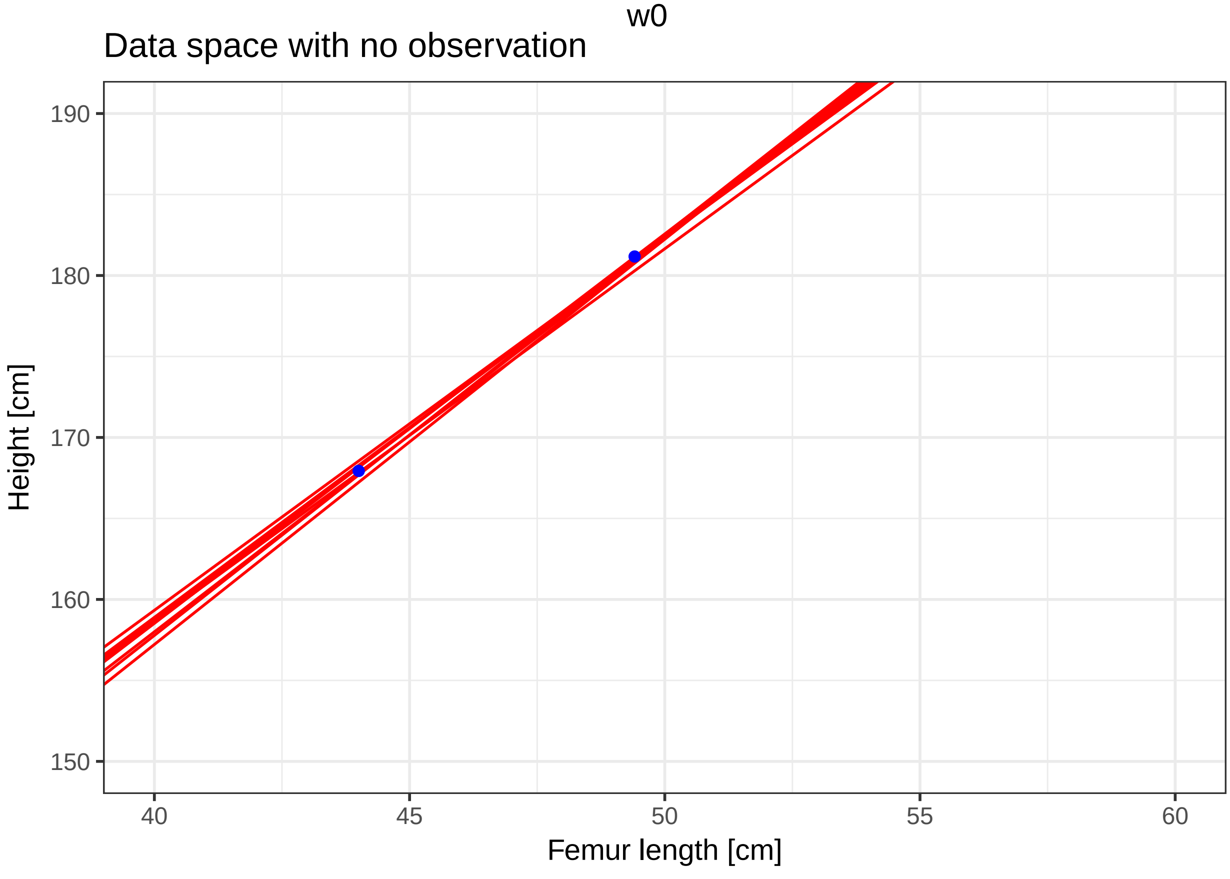
1 observation posterior distribution





2 observations posterior distribution





What is the estimated (posterior distriution of) height of a person whose femur has length 48cm?

Using predictive distribution, for the value of femur length above and using the equations below, the estimated height can be plotted as follows.

$$p(t \mid \mathbf{x}, \mathbf{t}, \alpha, \beta) = \mathcal{N}(t \mid \mathbf{m}_N^T \phi(\mathbf{x}), \sigma_N^2(\mathbf{x}))$$

$$\sigma_N^2(\mathbf{x}) = \frac{1}{\beta} + \phi(\mathbf{x})^T \mathbf{S}_N \phi(\mathbf{x}).$$

So that for $\beta = 4$, $\sigma_N^2(\mathbf{48cm})$ equals to:

```
## [1] 0.2864676
```

Thus, the predictive distribution is: