

Logistic regression

Load the `iris` data from the MASS library. We will only consider the two last species in this exercise, which you can find as `iris[,51:150]`.

Question 3.1. Use `glm` to fit a logistic regression model to this data set – modeling species as a binary variable conditionally on the four other measurements. Predict the species of the new measurement

<i>Sepal Length</i>	<i>Sepal Width</i>	<i>Petal Length</i>	<i>Petal Width</i>
5.9	2.7	4.6	1.3

Can you do any simplifications of the model?

Question 3.2. Compute the LDA classifier using only the two species. Compute and compare the conditional probabilities of the species on the dataset when we use logistic regression as above with the conditional probabilities found by LDA.

Recall the parameter function τ that gives the β parameters in the conditional distribution from the parameters in the LDA setup.

Question 3.3. Compute the plug-in estimate of τ from the LDA-method with $\hat{\beta}$ from logistic regression.

Hint: The group means can be extracted from the `lda`-object in R. Subtracting these from the observations you can form the covariance matrix and compute the value of the parameter function.