Exercise Set 2

This report contains my answers for exercise set 2.

Problem 8

Task a

Coefficients for penguins train data fit.

```
    (Intercept)     bill_length_mm     bill_depth_mm flipper_length_mm
  -380.81882000        14.30428414      -11.68758439        0.48138045
    body_mass_g
    -0.03094901
```

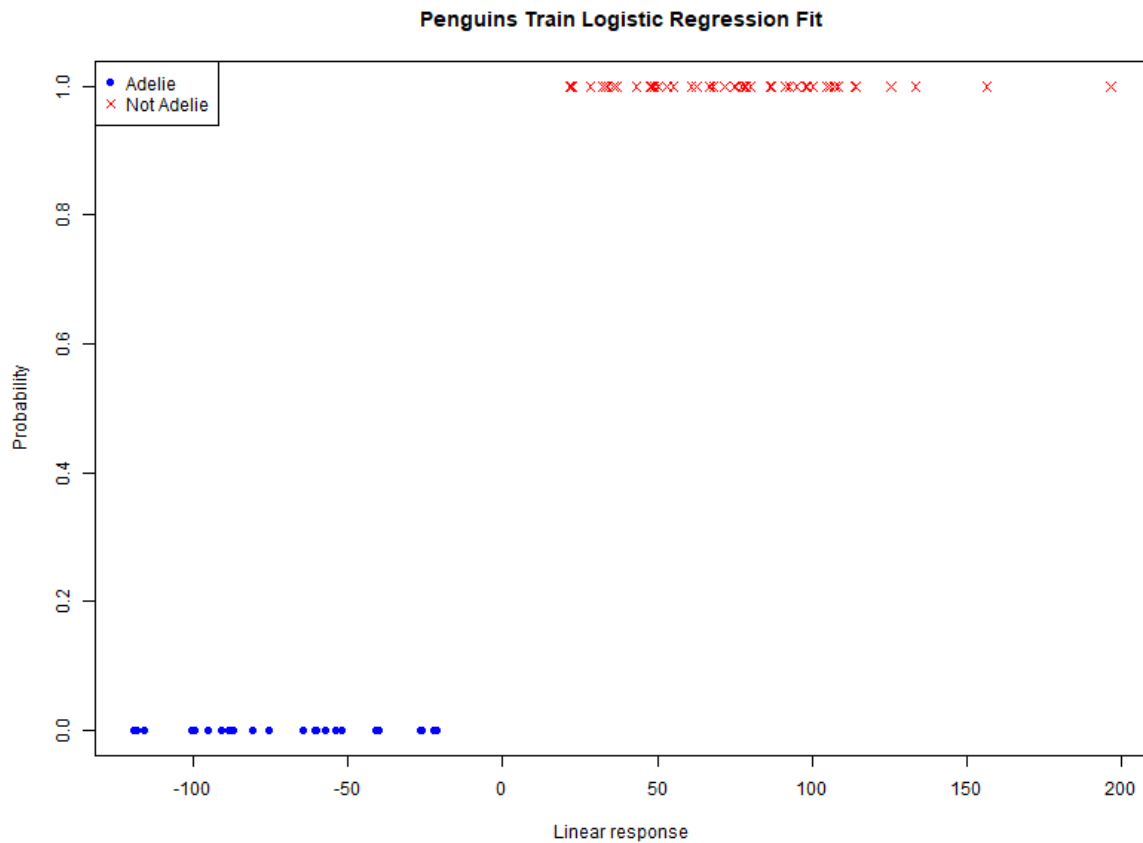Accuracy on training and test sets without regularisation:

```
[1] "Train accuracy:"
```

```
[1] 1
```

```
[1] "Test accuracy:"
```

```
[1] 0.9466667
```

Plot of penguins train data predictions.

Penguins Train Logistic Regression Fit

Task b

Coefficients for penguins train data fit.

```
Loading required package: Matrix

5 x 1 sparse Matrix of class "dgCMatrix"
                              s0
(Intercept)       -16.082633193
bill_length_mm      0.662337141
bill_depth_mm      -0.733855638
flipper_length_mm   0.004030675
body_mass_g             .
```

Accuracy on training and test sets using lasso with regularisation:

```
[1] "Train accuracy:"
```

```
[1] 1
```

```
[1] "Test accuracy:"
```

```
[1] 0.9866667
```

Task c

R gives warnings because the model is fitting the data too well and because there's no regulation, it will try to push probabilities closer to 0 or 1 forever if the programme didn't stop itself and give the warning.

Problem 9

According to the textbook:

The discriminant function is linear when p=1 and all classes have a shared variance:

$$\delta_k(x) = \frac{x \cdot \mu_k}{\sigma^2} - \frac{\mu_k^2}{2\sigma^2} + \log(\pi_k) \tag{4.18}$$

When each class has its own variance the discriminant function becomes quadratic (non-linear):

$$\delta_k(x) = -\frac{1}{2}(x - \mu_k)^\top \Sigma_k^{-1}(x - \mu_k) - \frac{1}{2}\log|\Sigma_k| + \log \pi_k \tag{4.28}$$

Expanded form:

$$\delta_k(x) = -\frac{1}{2}x^\top \Sigma_k^{-1} x + x^\top \Sigma_k^{-1} \mu_k - \frac{1}{2}\mu_k^\top \Sigma_k^{-1} \mu_k - \frac{1}{2}\log|\Sigma_k| + \log \pi_k \tag{4.28}$$

Problem 10

Task a

Tables of each attribute's means, standard deviations and class probabilities using Laplace smoothing:

species	variable	value
Adelie	bill_length_mm_mean	38.1240000
notAdelie	bill_length_mm_mean	47.8180000
Adelie	bill_depth_mm_mean	18.3360000
notAdelie	bill_depth_mm_mean	15.8900000
Adelie	flipper_length_mm_mean	188.8800000
notAdelie	flipper_length_mm_mean	211.3000000
Adelie	body_mass_g_mean	3576.0000000
notAdelie	body_mass_g_mean	4657.0000000
Adelie	bill_length_mm_sd	2.7815284
notAdelie	bill_length_mm_sd	3.5994722
Adelie	bill_depth_mm_sd	1.2041179
notAdelie	bill_depth_mm_sd	1.9654412
Adelie	flipper_length_mm_sd	6.3200738
notAdelie	flipper_length_mm_sd	11.7928550
Adelie	body_mass_g_sd	461.3431478
notAdelie	body_mass_g_sd	787.5310166
Adelie	class_prob	0.3418255
notAdelie	class_prob	0.6581745