```
layout: post
title: "ISL_Chapter4_Logistic Regression"
author: "hyeju.kim"
categories: facts
tags: [ISL]
image: LinearRegression.png
```

"""" """"

```
<script type="text/javascript" src="http://cdn.mathjax.org/mathjax/latest/MathJax.js?config=TeX-AMS-MML_HTMLorMML"></script>

<script type="text/x-mathjax-config">

MathJax.Hub.Config({

    displayAlign: "center"

});

</script>

$p(X) = \frac{e^{\beta_0 + \beta_1X}}{1+e^{\beta_0 + \beta_1X} }$
p(X) = \frac{e^{\beta_0 + \beta_1X}}{1+e^{\beta_0 + \beta_1X} }
```

Chapter 4. Classification

- What is classification?

  predicting a qualitative reponse

4.1 An Overview of Classification

Dataset Introduction

- default(Y) : Yes / No
- balance(X_1)
- income(X_2)

4.2 Why Not Linear Regression?

$$Y = \begin{cases} 1 & \text{if } \texttt{stroke}; \\ 2 & \text{if } \texttt{drug overdose}; \\ 3 & \text{if } \texttt{epileptic seizure}. \end{cases}$$

- The gap between levels are not exactly same

  Then, binary variable? (dummy variable)

- Estimates can be outside the [0,1] interval

4.3 Logistic Regression

- Logistic regression model predicts **the probability that Y belongs to a particular category**, rather than the reponse Y directly

  ex.

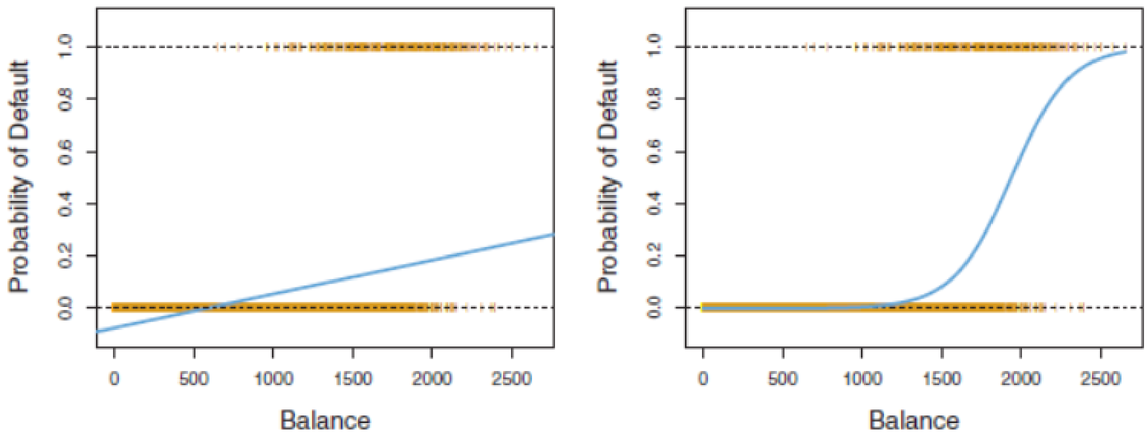

FIGURE 4.2. *Classification using the* `Default` *data.* Left: *Estimated probability of* `default` *using linear regression. Some estimated probabilities are negative! The orange ticks indicate the 0/1 values coded for* `default`*(*`No` *or* `Yes`*).* Right: *Predicted probabilities of* `default` *using logistic regression. All probabilities lie between* 0 *and* 1*.*

The Logistic Model

- how to set output values between [0,1]?

  logistic function

  $p(X) = \frac{e^{\beta_0+\beta_1 X}}{1+e^{\beta_0+\beta_1 X}}$ (4.2)

$$p(X) = \frac{e^{\beta_0+\beta_1 X}}{1 + e^{\beta_0+\beta_1 X}}. \tag{4.2}$$

  - S - shaped curve

After a bit of manipulation of (4.2), we find that

$$\frac{p(X)}{1 - p(X)} = e^{\beta_0+\beta_1 X}. \tag{4.3}$$

  - p(X) / [1-p(X)] is called **odds**, between 0(very low possibility) and infinite(very high possibility)

By taking the logarithm of both sides of (4.3), we arrive at

$$\log\left(\frac{p(X)}{1 - p(X)}\right) = \beta_0 + \beta_1 X. \tag{4.4}$$

  - log(p(X) / [1-p(X)]) is called the log-odds or **logit**.
  - β_1 does not correspond to the change in p(X) associated with a one-unit increase in X
  - The amount that p(X) changed due to one-unit change in X will depend on **the current value of X**
  - **INCREASING X BY ONE UNIT CHANGES MULTIPLIES THE ODDS BY** e^{β_1}

    S