Problem Set 7

Logistic Regression

Consider the following model

$$P(Y_i = y_i) = \pi_i^{y_i}(1-\pi_i)^{1-y_i} \tag{1}$$

where

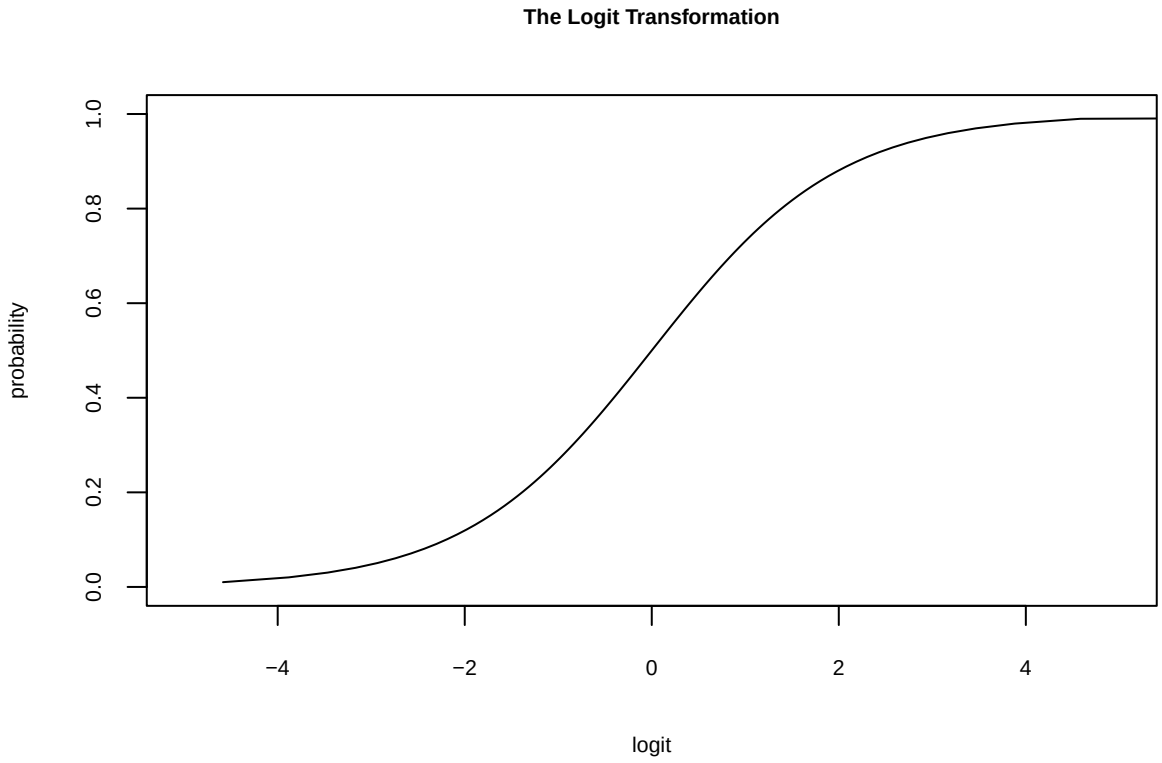
$$\pi_i = \frac{exp\{\mathbf{x}_i'\boldsymbol{\beta}\}}{1+exp\{\mathbf{x}_i'\boldsymbol{\beta}\}} \tag{2}$$

1. Simulate this model with the probabilities as described above with the following values: - $n = 1000$ - $\beta_0 = -2$, $\beta_1 = 0.1$, $\beta 2 = 1$. - $x_{0i} = 1 \, \forall \, i$, $x_{1i} \sim \mathcal{U}(18, 60)$, $x_{2i} \sim \mathcal{B}(0.5)$.
2. Write down the likelihood function analytically and in code.
3. Estimate $\beta_0, \beta_1, \beta_2$ via maximum likelihood using the code you have written above and calculate the standard errors.
4. Propose and calculate a suitable method for the interpretation of the coefficients as discussed in the lecture.
5. Bonus: graphically illustrate the problem when using linear regression to calculate the predicted probabilities.

```
##The logit transformation#

####Generate a vector of probabilities########
pi=seq(0,0.99999999,le=100)
logit<-function(x)
{
  logit<-log(x/(1-x))
  return(logit)
}

plot(logit(pi),pi,type="l", xlim=c(-5,5), main="The Logit Transformation",xlab="logit",ylab="probabilit
```




```r
#Maximum Likelihood estimation: Logit###

##General Syntax###

library(miscTools)#the maxLik package acts as a wrapper for the more
#basic "optim", the library miscTools is required.
library(maxLik)
loglike<-function(beta)#Define a function that takes only the parameter vector as arguments.
{
  ll <- "my log likelihood function" #depending on your optimization routine,
  #check whether you need the negative or the positive log likelihood!
  return(ll)
}
estim<-maxBFGS(loglike,finalHessian=TRUE,start=c(.,.))###initialize the optimization,
#pass on starting values and store the results in estim
estim.par<-estim$estimate ### store the paramter estimates in a variable "estim.par"
```