Exercises: Week 2

Prof. Conlon

Due: 2/8/21

1. Let's load the Boston HMDA data.

The function should take the following arguments:

- dir: debt to income ratio
- hir : housing to income ratio
- single : dummy for single borrower
- self : dummy for self-employed

```
library("Ecdat")
```

```
##
## Attaching package: 'Ecdat'

## The following object is masked from 'package:datasets':
##
##     Orange
```

```
library("fixest")
data("Hmda")
Hmda$deny <- 1*(Hmda$deny=="yes")
probit <- feglm(deny ~ dir + hir + single + self, data = Hmda,
                family = binomial(link = "probit"))
```

```
## NOTE: 1 observation removed because of NA values (RHS: 1).
```

```
logit <- feglm(deny ~ dir + hir + single + self, data = Hmda,
               family = binomial(link = "logit"))
```

```
## NOTE: 1 observation removed because of NA values (RHS: 1).
```

```
etable(probit,logit)
```

```
##                               probit              logit
## Dependent Var.:                 deny               deny
##
## Constant        -2.268*** (0.1470) -4.154*** (0.2848)
## dir              3.183*** (0.4960)  6.118*** (0.9181)
## hir               -0.4892 (0.5729)    -0.7501 (1.043)
## singleyes       0.2407*** (0.0692) 0.4503*** (0.1307)
## selfyes           0.1971. (0.1023)   0.3609. (0.1887)
## _______________ __________________ __________________
## Family                      Probit              Logit
## S.E. type                      IID                IID
## Observations                 2,380              2,380
## Squared Cor.               0.05771            0.05933
## Pseudo R2                  0.05552            0.05696
```

```
## BIC                              1,686.2            1,683.7
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

2. Consider the regression model of the logit regression:

$$deny_i = F(\beta_1 \cdot dir_i + \beta_2 \cdot hir_i + \beta_3 \cdot single_i + \beta_4 \cdot self_i)$$

   For a single observation compute the contribution to the log-likelihood (analytically)

3. For a single observation compute the Score (analytically).
4. Compute the Hessian Matrix and Fisher information (analytically).
5. Code up the Fisher Information for the logit model above $I(\widehat{\beta})$ using the Hessian Matrix.
6. Code up the Fisher Information for the logit model above $I(\widehat{\beta})$ using the score method.
7. Compute the standard errors from the Fisher information and compare them to the standard errors reported from the regression. How do they compare?
8. Generate $n = 100$ observations where $\lambda = 15$ from a poisson model:

$$Y_i \sim Pois(\lambda)$$

9. The poisson distribution is a discrete distribution for count data where the p.m.f. is given by:

$$Pr(Y_i = k) = \frac{\lambda^k e^{-\lambda}}{k!}a$$

10. Write the log-likelihood $\ell(y_1, \ldots, y_n; \lambda)$ (analytically).
11. Write the Score contribution $\mathcal{S}_i(y_i; \lambda)$ (analytically).
12. Write the Hessian Contribution $\mathcal{H}_i(y_i; \lambda)$ (analytically).
13. Code up the log-likelihood function

```r
pois_log_lik <- function(lambda,y){

  return(ll)
}
```

14. Find the value of λ that maximizes your log likelihood using `optim` in R.
15. Write a function that returns the standard error of $\hat{\lambda}$:

```r
pois_se <- function(lambda_hat,y){

  return(se)
}
```