Logistic Regression

Brandon Wolff

March 20, 2017

Bronx Logit

```
stop_frisk <- read.csv("2015_stopandfrisk_CLEAN_w_counties.CSV")

sf_bronx <- stop_frisk[stop_frisk$city == 'Bronx',]

for(i in 1:4582){
  if(nchar(sf_bronx$datestop[i]) == 3){
    if(substr(sf_bronx$datestop[i],1,1) == "1"){
      sf_bronx$datestop[i] <- paste0("-01-
",substring(sf_bronx$datestop[i],2))
    } else if (substr(sf_bronx$datestop[i],1,1) == "2"){
      sf_bronx$datestop[i] <- paste0("-02-
",substring(sf_bronx$datestop[i],2))
    } else if (substr(sf_bronx$datestop[i],1,1) == "3"){
      sf_bronx$datestop[i] <- paste0("-03-
",substring(sf_bronx$datestop[i],2))
    } else if (substr(sf_bronx$datestop[i],1,1) == "4"){
      sf_bronx$datestop[i] <- paste0("-04-
",substring(sf_bronx$datestop[i],2))
    } else if (substr(sf_bronx$datestop[i],1,1) == "5"){
      sf_bronx$datestop[i] <- paste0("-05-
",substring(sf_bronx$datestop[i],2))
    } else if (substr(sf_bronx$datestop[i],1,1) == "6"){
      sf_bronx$datestop[i] <- paste0("-06-
",substring(sf_bronx$datestop[i],2))
    } else if (substr(sf_bronx$datestop[i],1,1) == "7"){
      sf_bronx$datestop[i] <- paste0("-07-
",substring(sf_bronx$datestop[i],2))
    } else if (substr(sf_bronx$datestop[i],1,1) == "8"){
      sf_bronx$datestop[i] <- paste0("-08-
",substring(sf_bronx$datestop[i],2))
    } else if (substr(sf_bronx$datestop[i],1,1) == "9"){
      sf_bronx$datestop[i] <- paste0("-09-
",substring(sf_bronx$datestop[i],2))
    }
  } else if (nchar(sf_bronx$datestop[i]) == 4){
    if(substr(sf_bronx$datestop[i],1,2) == "11"){
      sf_bronx$datestop[i] <- paste0("-11-
",substring(sf_bronx$datestop[i],3))
    } else if (substr(sf_bronx$datestop[i],1,2) == "12"){
      sf_bronx$datestop[i] <- paste0("-12-
```

```
",substring(sf_bronx$datestop[i],3))
    } else if (substr(sf_bronx$datestop[i],1,2) == "10"){
      sf_bronx$datestop[i] <- paste0("-10-
",substring(sf_bronx$datestop[i],3))
    }
  }
}

sf_bronx$datestop <- paste0("2015", sf_bronx$datestop)

sf_bronx$race <- as.factor(sf_bronx$race)
sf_bronx <- within(sf_bronx, race <- relevel(race, ref = 4))

#sf_bronx$datestop <- as.Date(sf_bronx$datestop)

sf_bronx$pforce[sf_bronx$pforce == 1] <- "Physical Force"
sf_bronx$pforce[sf_bronx$pforce == 0] <- "No Physical Force"

sf_bronx$pforce <- as.factor(sf_bronx$pforce)

summary(fit <- glm(pforce ~ race + sex, data = sf_bronx, family =
"binomial"))

##
## Call:
## glm(formula = pforce ~ race + sex, family = "binomial", data =
sf_bronx)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -1.851  -1.087  -1.047   1.270   1.367
##
## Coefficients:
##                                        Estimate Std. Error z value
Pr(>|z|)
## (Intercept)                           -0.327376   0.184058  -1.779
0.0753
## raceAmerican Indian / Alaskan Native  1.722921   0.795612   2.166
0.0303
## raceAsian / Pacific Islander          0.302883   0.461234   0.657
0.5114
## raceBlack                             -0.008028   0.152246  -0.053
0.9579
## raceWhite - Hispanic                  -0.107047   0.157289  -0.681
0.4961
## sexMale                               0.119804   0.118210   1.013
0.3108
##
## (Intercept)                           .
## raceAmerican Indian / Alaskan Native *
```

```
## raceAsian / Pacific Islander
## raceBlack
## raceWhite - Hispanic
## sexMale
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 6108.8  on 4455  degrees of freedom
## Residual deviance: 6098.0  on 4450  degrees of freedom
##   (126 observations deleted due to missingness)
## AIC: 6110
##
## Number of Fisher Scoring iterations: 4
```