

Stat 245 – Housing Analysis

Trey Tipton

October 28, 2021

```
hdma_mi_20 <- read_csv('https://sldr.netlify.app/data/hdma-mi-20.csv')
```

Choose Response and Predictors

The first variable I chose as a predictor is race, for this is the main variable of interest as to possible biases for mortgage loan acceptance. Other things that might affect our response variable of 'action_taken' are income, the amount of the loan, and Debt-to-Income Ratio. The Markup mentioned these as well as showing graphics of the debt-to-income ratio, so I believe it would be an interesting variable to look at.

Fixing Variable Labels and Names

```
hdma_mi_20 <- hdma_mi_20 %>%  
  mutate(derived_race = case_when(derived_race == "Free Form Text Only" ~ "Other",  
                                   derived_race == "Joint" ~ "Other",  
                                   derived_race == "2 or more minority races" ~ "Multiracial",  
                                   derived_race == "White" ~ "White",  
                                   derived_race == "Black or African American" ~ "Black", derived_race ==  
                                   derived_race == "Native Hawaiian or Other Pacific Islander" ~ "Pacific  
                                   derived_race == "American Indian or Alaska Native" ~ "Native American  
                                   TRUE ~ "Other"))%>%  
  filter(derived_race != "Race Not Available")
```

Fit Model

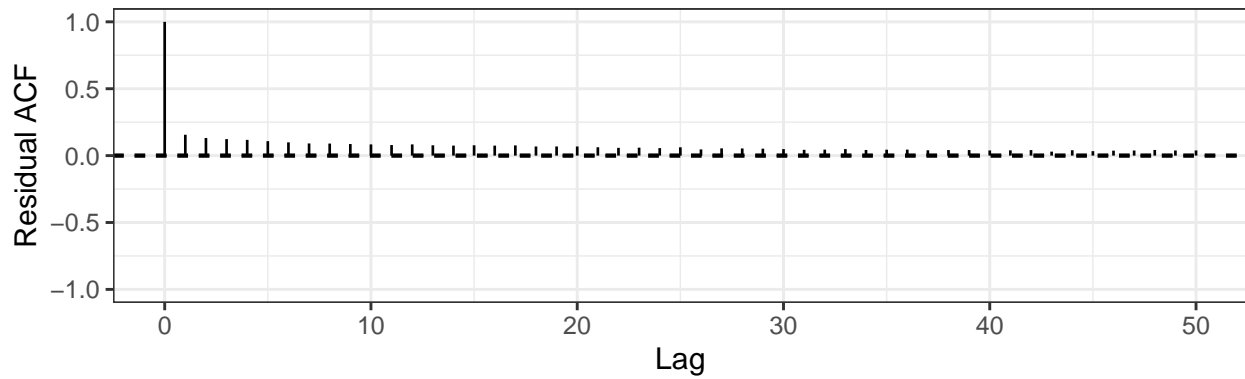
```
housing_model <- glm(factor(action_taken) ~ derived_race  
  + income + debt_to_income_ratio  
  + loan_amount, data = hdma_mi_20,  
  family = binomial(link = 'logit'))  
  
summary(housing_model)  
  
##  
## Call:  
## glm(formula = factor(action_taken) ~ derived_race + income +  
##     debt_to_income_ratio + loan_amount, family = binomial(link = "logit"),  
##     data = hdma_mi_20)  
##  
## Deviance Residuals:  
##      Min       1Q   Median       3Q      Max
```

```
## -2.6294 -0.3663 -0.3102 -0.2857 4.2690
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.966e+00  7.638e-02 -25.746 < 2e-16 ***
## derived_raceBlack      7.890e-01  7.824e-02  10.084 < 2e-16 ***
## derived_raceMultiracial  4.646e-01  3.898e-01   1.192  0.23329
## derived_raceNative American  3.322e-01  2.217e-01   1.498  0.13405
## derived_raceOther      1.840e-01  7.031e-02   2.617  0.00887 **
## derived_racePacific     4.195e-01  4.172e-01   1.006  0.31460
## derived_raceWhite     -3.381e-01  6.499e-02  -5.202 1.97e-07 ***
## income           4.103e-05  5.056e-05   0.811  0.41714
## debt_to_income_ratio>60%  4.886e+00  1.141e-01  42.831 < 2e-16 ***
## debt_to_income_ratio20%-<30% -5.849e-01  4.908e-02 -11.917 < 2e-16 ***
## debt_to_income_ratio30%-<36% -5.478e-01  5.059e-02 -10.828 < 2e-16 ***
## debt_to_income_ratio36     -4.313e-01  8.312e-02  -5.189 2.12e-07 ***
## debt_to_income_ratio37     -5.813e-01  8.639e-02  -6.729 1.71e-11 ***
## debt_to_income_ratio38     -4.545e-01  8.281e-02  -5.489 4.05e-08 ***
## debt_to_income_ratio39     -4.871e-01  8.298e-02  -5.870 4.37e-09 ***
## debt_to_income_ratio40     -4.884e-01  8.318e-02  -5.871 4.33e-09 ***
## debt_to_income_ratio41     -4.574e-01  8.096e-02  -5.650 1.60e-08 ***
## debt_to_income_ratio42     -3.073e-01  7.596e-02  -4.045 5.23e-05 ***
## debt_to_income_ratio43     -4.539e-01  8.147e-02  -5.571 2.53e-08 ***
## debt_to_income_ratio44     -3.048e-01  7.512e-02  -4.058 4.95e-05 ***
## debt_to_income_ratio45     -1.641e-01  8.713e-02  -1.883 0.05973 .
## debt_to_income_ratio46     -1.429e-01  9.533e-02  -1.499 0.13398
## debt_to_income_ratio47     -3.646e-02  9.285e-02  -0.393 0.69456
## debt_to_income_ratio48     -1.044e-01  9.663e-02  -1.080 0.28011
## debt_to_income_ratio49      4.999e-02  8.560e-02   0.584 0.55925
## debt_to_income_ratio50%-60%  2.752e+00  6.989e-02  39.375 < 2e-16 ***
## debt_to_income_ratioExempt  2.846e-01  6.984e-02   4.075 4.61e-05 ***
## loan_amount       -1.399e-06  1.082e-07 -12.929 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 54967  on 104976  degrees of freedom
## Residual deviance: 46306  on 104949  degrees of freedom
## (3694 observations deleted due to missingness)
## AIC: 46362
##
## Number of Fisher Scoring iterations: 6
```

Check Conditions

Independence Condition

```
s245::gf_acf(~housing_model)%>%
  gf_lims(y = c(-1,1))
```



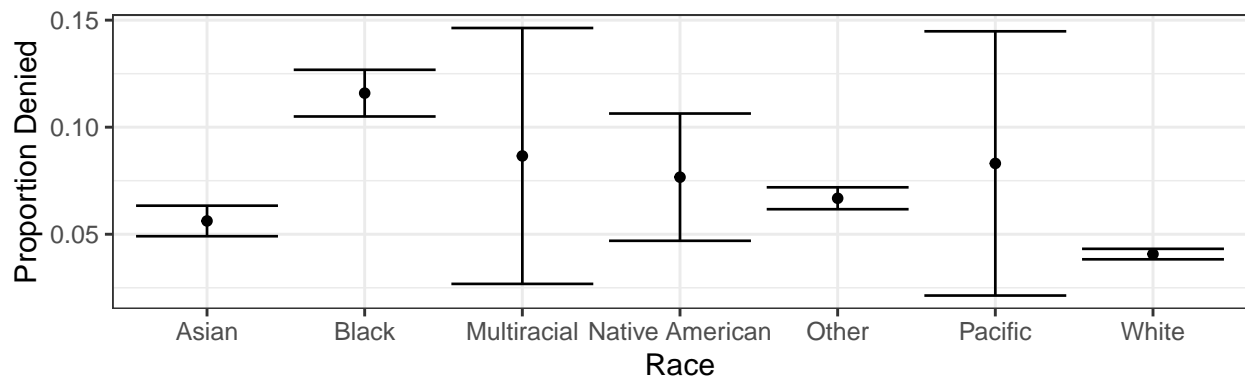
I believe that this model does not pass the independence condition, as the range is so small it can hardly be seen. Albeit, the Residual ACF values are outside of the allowed range.

Prediction Plot

```
require(s245)

get_fixed(housing_model)

##   factor.action_taken. derived_race income debt_to_income_ratio loan_amount
## 2                    1         White      79                20%-<30%    195000
pred_plot(housing_model, 'derived_race', data = hdma_mi_20)%>%
  gf_labs(y = "Proportion Denied", x = "Race")
```



The prediction plot does seem consistent with The Markup's nationwide analysis. Minorities that apply for mortgage loans are clearly denied at a higher rate.