BST210 Project Check-in 2: Appendix

Linear Regression

All of our potential dependent variables are in the form of counts, rates and percentages, however for the purposes of this assignment, we treated the emergency department visits per 1,000 beneficiaries as a linear variable and fit a multiple linear regression to determine the effects of Medicaid expansion on this rate. For our analysis, we plan to run this model as a Poisson model.

To evaluate the effects of Medicaid Expansion on health outcomes among the Medicare population in 2018, we fit a multiple linear regression model:

$$E(\text{Number of ED visits/1,000 beneficiaries}) = \beta_0 + \beta_1 * \text{Average Age} +$$
$$\beta_2 * I(\text{State Expanded in 2014}) +$$
$$\beta_3 * I(\text{State Expanded after 2014}) +$$
$$\beta_4 * I(\text{State has not Expanded}) +$$
$$\beta_5 * \text{Percent Eligible for Medicaid} +$$
$$\beta_6 * \text{Percent Eligible for Medicaid}^2$$

```
mod.lm <- lm(emergency_department_visits_per_1000_beneficiaries ~ average_age +
               expansion_status + percent_eligible_for_medicaid +
               percent_eligible_for_medicaid_2, data = medicare %>% mutate(percent_eligible_for_medicaid

# Model interpretation
summary(mod.lm)
```

```
##
## Call:
## lm(formula = emergency_department_visits_per_1000_beneficiaries ~
##     average_age + expansion_status + percent_eligible_for_medicaid +
##         percent_eligible_for_medicaid_2, data = medicare %>%
##     mutate(percent_eligible_for_medicaid_2 = percent_eligible_for_medicaid^2))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -655.61  -66.17    3.30   69.00 1056.20
##
## Coefficients:
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     1479.308    128.124  11.546  < 2e-16 ***
## average_age                      -15.953      1.688  -9.451  < 2e-16 ***
## expansion_status2                 36.771      8.627   4.262 2.08e-05 ***
## expansion_status3                 16.459      9.581   1.718   0.0859 .
## expansion_status4                 55.867      8.260   6.763 1.60e-11 ***
## percent_eligible_for_medicaid   2139.196     94.964  22.526  < 2e-16 ***
## percent_eligible_for_medicaid_2 -2710.022    165.888 -16.336  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

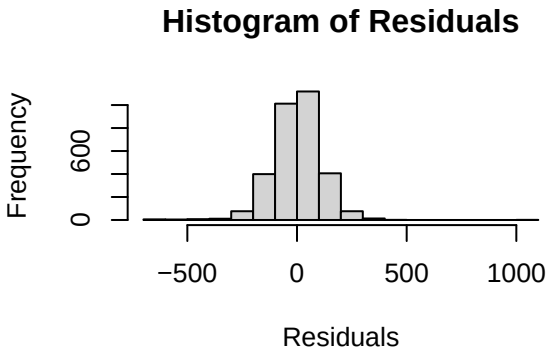
```
##
## Residual standard error: 112.9 on 3122 degrees of freedom
##   (15 observations deleted due to missingness)
## Multiple R-squared:  0.4149, Adjusted R-squared:  0.4138
## F-statistic:   369 on 6 and 3122 DF,  p-value: < 2.2e-16
```

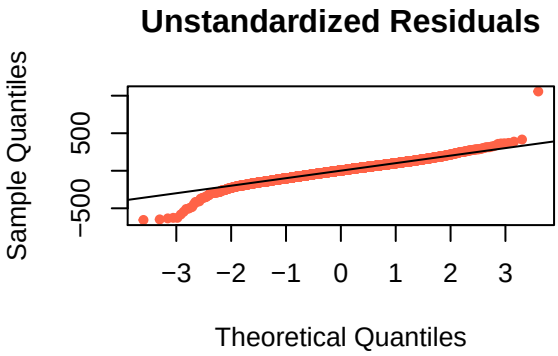
```
confint(mod.lm)
```

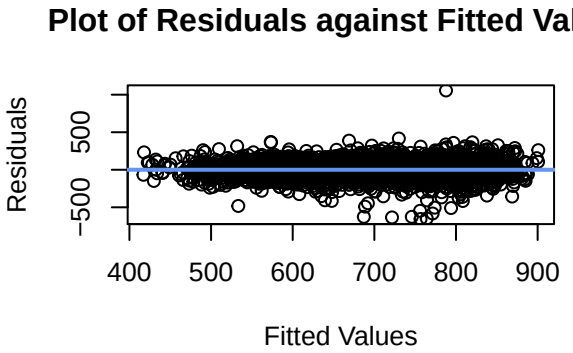
```
##                                    2.5 %      97.5 %
## (Intercept)                     1228.09271  1730.52411
## average_age                      -19.26243   -12.64340
## expansion_status2                 19.85531    53.68695
## expansion_status3                 -2.32591    35.24445
## expansion_status4                 39.67072    72.06293
## percent_eligible_for_medicaid   1952.99814  2325.39465
## percent_eligible_for_medicaid_2 -3035.28232 -2384.76099
```

According to this model, for counties with the same expansion status level and percent eligible for Medicaid, a one year increase in the average age of Medicare beneficiaries results in about 16 less emergency department visits per 1,000 beneficiaries. For counties with the same average age of Medicare beneficiaries and percent eligible for Medicaid, being in a state that expanded Medicaid in 2014 results in about 37 (95% CI: (19.86, 53.69)) more emergency department visits per 1,000 beneficiaries than counties that are in states that expanded Medicaid before 2014. For counties with the same average age of Medicare beneficiaries and percent eligible for Medicaid, being in a state that expanded Medicaid in years after 2014 results in about 16 (95% CI: (-2.326, 35.24)) more emergency department visits per 1,000 beneficiaries than counties that are in states that expanded Medicaid before 2014. For counties with the same average age of Medicare beneficiaries and percent eligible for Medicaid, being in a state that has not expanded Medicaid results in about 56 (95% CI: (39.67, 72.06)) more emergency department visits per 1,000 beneficiaries than counties that are in states that expanded Medicaid before 2014. For counties in states with the same expansion category and average age of Medicare beneficiaries, a one percent increase in the percent eligible for Medicaid results in about 571 less emergency department visits per 1,000 beneficiaries. All p-values, with the exception of that for the expansion category 3 (state expanded Medicaid after 2014), is < 0.001.

```
# Model evaluation
par(mfrow = c(2, 2))
hist(mod.lm$residuals, main = "Histogram of Residuals", xlab = "Residuals")
qqnorm(residuals(mod.lm), pch = 20, col = "tomato", main = "Unstandardized Residuals")
qqline(residuals(mod.lm)) #NJP Added line
plot(fitted(mod.lm), residuals(mod.lm), main = "Plot of Residuals against Fitted Values", xlab = "Fitte
abline(h = 0, col = "cornflowerblue", lwd = 2)
```







In evaluating this linear model, we have determined that the linearity assumption holds, as there is no pattern in the residuals. The residuals are pretty normally distributed, although there is some evidence of left skew in the qqplot.

Poisson Regression

FQHC/RHC centers are centers that provide services to Medicare beneficiaries in geographic areas that have a shortage of health services. To determine the effect of Medicaid expansion on the rate of FQHC/RHC visits per 1,000 beneficiaries, we fit a Poisson model. The form of the model is:

$$
\begin{aligned}
&log(\lambda) = \beta_0 + \beta_1 * AverageAge + \\
&\beta_2 * I(\text{State Expanded Medicaid in 2014}) + \\
&\beta_3 * I(\text{State Expanded Medicaid after 2014}) + \\
&\beta_4 * I(\text{State has not expanded Medicaid}) + \\
&\beta_5 * I(\text{Percent Eligible for Medicaid}) + \\
&\beta_6 * I(\text{Percent Eligible for Medicaid}^2)
\end{aligned}
$$

where λ is the incidence rate or FQHRC/RHC visits per 1,000 beneficiaries.

```
mod.pois <-  glm(`fqhc/rhc_visits_per_1000_beneficiaries` ~ average_age + expansion_status + percent_el
summary(mod.pois)
```

```
##
## Call:
## glm(formula = `fqhc/rhc_visits_per_1000_beneficiaries` ~ average_age +
##     expansion_status + percent_eligible_for_medicaid + percent_eligible_for_medicaid_2,
##     family = poisson(), data = na.omit(medicare[, c("county",
```

```
##         "percent_eligible_for_medicaid", "average_age", "expansion_status",
##         "fqhc/rhc_visits_per_1000_beneficiaries")]) %>% mutate(percent_eligible_for_medicaid_2 = per
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -94.57  -32.26  -14.06   17.87  110.44
##
## Coefficients:
##                                  Estimate Std. Error  z value Pr(>|z|)
## (Intercept)                    -1.143e+01  3.186e-02 -358.690  < 2e-16 ***
## average_age                     2.405e-01  4.147e-04  579.943  < 2e-16 ***
## expansion_status2               2.031e-01  2.260e-03   89.835  < 2e-16 ***
## expansion_status3              -1.871e-02  2.620e-03   -7.141 9.29e-13 ***
## expansion_status4               1.689e-01  2.178e-03   77.545  < 2e-16 ***
## percent_eligible_for_medicaid   6.664e+00  2.298e-02  290.015  < 2e-16 ***
## percent_eligible_for_medicaid_2 -4.219e+00  3.683e-02 -114.544  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 3866643  on 3108  degrees of freedom
## Residual deviance: 3429374  on 3102  degrees of freedom
## AIC: 3454731
##
## Number of Fisher Scoring iterations: 5
```

```
exp(confint(mod.pois))
```

```
## Waiting for profiling to be done...
##                                      2.5 %       97.5 %
## (Intercept)                    1.023444e-05 1.159580e-05
## average_age                    1.270841e+00 1.272908e+00
## expansion_status2              1.219736e+00 1.230592e+00
## expansion_status3              9.764351e-01 9.865171e-01
## expansion_status4              1.178939e+00 1.189047e+00
## percent_eligible_for_medicaid  7.492959e+02 8.199245e+02
## percent_eligible_for_medicaid_2 1.368768e-02 1.581387e-02
```

By our model, for counties that have the same average age of Medicare beneficiaries and same percentage of beneficiaries who are eligible for Medicaid, counties in states that expanded in 2014 have an incidence rate of FQHC/RHC visits that is 1.225 (95% CI: (1.219, 1.231)) times the incidence rate of FQHC/RHC visits in counties whose state expanded before 2014. For counties that have the same average age of Medicare beneficiaries and same percentage of beneficiaries who are eligible for Medicaid, counties in states that expanded after 2014 have an incidence rate of FQHC/RHC visits that is 0.981 (95% CI: (.976, .987)) times the incidence rate of FQHC/RHC visits in counties whose state expanded before 2014. For counties that have the same average age of Medicare beneficiaries and same percentage of beneficiaries who are eligible for Medicaid, counties in states that have not expanded have an incidence rate of FQHC/RHC visits that is 1.184 (95% CI: (1.179, 1.189)) times the incidence rate of FQHC/RHC visits in counties whose state expanded before 2014.

Overdispersion

```r
# Checking for Overdispersion (Latitude-only model)
deviance(mod.pois)/mod.pois$df.residual
```

```
## [1] 1105.537
```

```r
pearson.stat1 <- sum((na.omit(medicare[, c("county", "percent_eligible_for_medicaid", "average_age", "e
pearson.stat1/mod.pois$df.residual
```

```
## [1] 1173.745
```

The deviance/degrees of freedom is about 1105, and the Pearson χ^2 statistic divided by degrees of freedom is around 1174. This indicates that there is overdisperson in the model. To account for this, we will need to try fitting a negative binomial model.