

# stat431 q3 q3

Yiming Shen 20891774

18/11/2023

```
# Access the NMES1988 dataset and print the first few observations
library(AER) # warnings are OK
```

```
## Loading required package: car
## Loading required package: carData
## Loading required package: lmtest
## Loading required package: zoo
##
## Attaching package: 'zoo'
## The following objects are masked from 'package:base':
##
##   as.Date, as.Date.numeric
## Loading required package: sandwich
## Loading required package: survival
```

```
data(NMES1988)
```

```
head(NMES1988)
```

```
##   visits nvisits ovisits novisits emergency hospital health chronic   adl
## 1     5      0      0      0         0         1 average      2 normal
## 2     1      0      2      0         2         0 average      2 normal
## 3    13      0      0      0         3         3   poor      4 limited
## 4    16      0      5      0         1         1   poor      2 limited
## 5     3      0      0      0         0         0 average      2 limited
## 6    17      0      0      0         0         0   poor      5 limited
##   region age afam gender married school income employed insurance medicaid
## 1  other 6.9  yes  male      yes      6 2.8810      yes      yes      no
## 2  other 7.4   no female     yes     10 2.7478      no      yes      no
## 3  other 6.6  yes female     no     10 0.6532      no      no      yes
## 4  other 7.6   no  male     yes      3 0.6588      no      yes      no
## 5  other 7.9   no female     yes      6 0.6588      no      yes      no
## 6  other 6.6   no female     no      7 0.3301      no      no      yes
```

```
help(NMES1988)
```

(c)

```
# hist for visits
hist(NMES1988$visits,
```

```

xlab="Number of visits",
main="The number of visits for health care",
breaks=40,
xlim=c(0,40),
ylim=c(0,0.2),
probability = T)

```

```

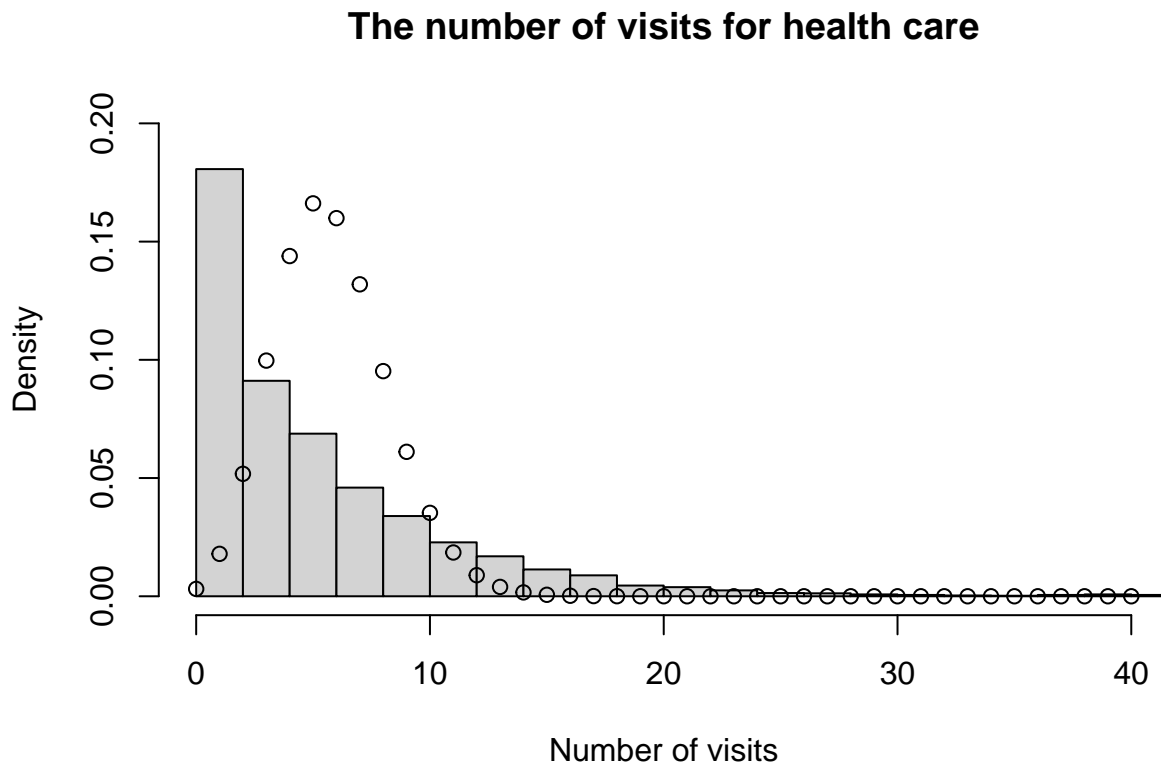
# Poisson Model

```

```

points(x=seq(0,40),y=dpois(x=seq(0,40),lambda = mean(NMES1988$visits)))

```



Comments: Based on the graphs, we found that the observation of 0 has much higher density than what we expected under Poisson model. Therefore, there are more respondents with 0 visits than might be expected under a Poisson model.

(d)

```
# code for part (e)
require(pscl)

## Loading required package: pscl

## Classes and Methods for R developed in the
## Political Science Computational Laboratory
## Department of Political Science
## Stanford University
## Simon Jackman
## hurdle and zeroinfl functions by Achim Zeileis

myZIP = zeroinfl(visits ~ chronic + health + insurance | chronic + insurance, data = NMES1988)
summary(myZIP)

##
## Call:
## zeroinfl(formula = visits ~ chronic + health + insurance | chronic +
## insurance, data = NMES1988)
##
## Pearson residuals:
##      Min       1Q   Median       3Q      Max
## -3.9221 -1.2195 -0.4316  0.5598 24.1031
##
## Count model coefficients (poisson with log link):
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   1.55878    0.01762  88.448  <2e-16 ***
## chronic       0.11868    0.00462  25.691  <2e-16 ***
## healthpoor    0.29470    0.01729  17.043  <2e-16 ***
## healthexcellent -0.30482    0.03115  -9.786  <2e-16 ***
## insuranceyes  0.14467    0.01631   8.870  <2e-16 ***
##
## Zero-inflation model coefficients (binomial with logit link):
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.37426    0.09213  -4.062 4.86e-05 ***
## chronic      -0.56112    0.04334 -12.948  < 2e-16 ***
## insuranceyes -0.88314    0.09464  -9.332  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Number of iterations in BFGS optimization: 11
## Log-likelihood: -1.651e+04 on 8 Df

# chronic in Poisson part
estimate_chronic_poi <- 0.11868
exp(estimate_chronic_poi)

## [1] 1.12601

Interpretation: When number of chronic conditions increases one and holding other variables unchanged, the
expected number of physician office visits (rate) with current number of chronic plus one will be 1.12601
times of the expected number of physician office visits (rate) with current number of chronic.

# poor health in Poisson part
estimate_poorhealth_poi <- 0.29470
```

```
exp(estimate_poorhealth_poi)
```

```
## [1] 1.342723
```

Interpretation: The expected number of physician office visits (rate) under poor self-perceived health status is 1.342723 times of the expected number of physician office visits (rate) under average self-perceived health status, while holding other variables unchanged.

```
# intercept in logistic part
```

```
estimate_intercept_log <- -0.37426
```

```
exp(estimate_intercept_log)
```

```
## [1] 0.6877981
```

Interpretation: When the number of chronic conditions is 0 and without private insurance coverage, the odds of 0 physician office visit is 0.6877981, while holding other variables unchanged.

```
# insurance in logistic part
```

```
estimate_insurance_log <- -0.88314
```

```
exp(estimate_insurance_log)
```

```
## [1] 0.4134825
```

Interpretation: The odds ratio of 0 physician office visit with private insurance coverage VS. without private insurance coverage is 0.4134825, while holding other variables unchanged.