Chicago Models

```
library(DT)
library(data.table)
```

```
## Warning: package 'data.table' was built under R version 3.5.2
```

```
library(magrittr)
library(sf)
```

```
## Warning: package 'sf' was built under R version 3.5.2
```

```
## Linking to GEOS 3.7.2, GDAL 2.4.2, PROJ 5.2.0
```

```
library(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
```

```
library(countreg)
```

```
## Loading required package: MASS
```

```
##
## Attaching package: 'countreg'
```

```
## The following objects are masked from 'package:pscl':
##
##     hurdle, hurdle.control, hurdletest, zeroinfl, zeroinfl.control
```

```
library(GISTools)
```

```
## Loading required package: maptools
```

```
## Warning: package 'maptools' was built under R version 3.5.2
```

```
## Loading required package: sp
```

```
## Checking rgeos availability: TRUE
```

```
## Loading required package: RColorBrewer
```

```
## Loading required package: rgeos
```

```
## Warning: package 'rgeos' was built under R version 3.5.2
```

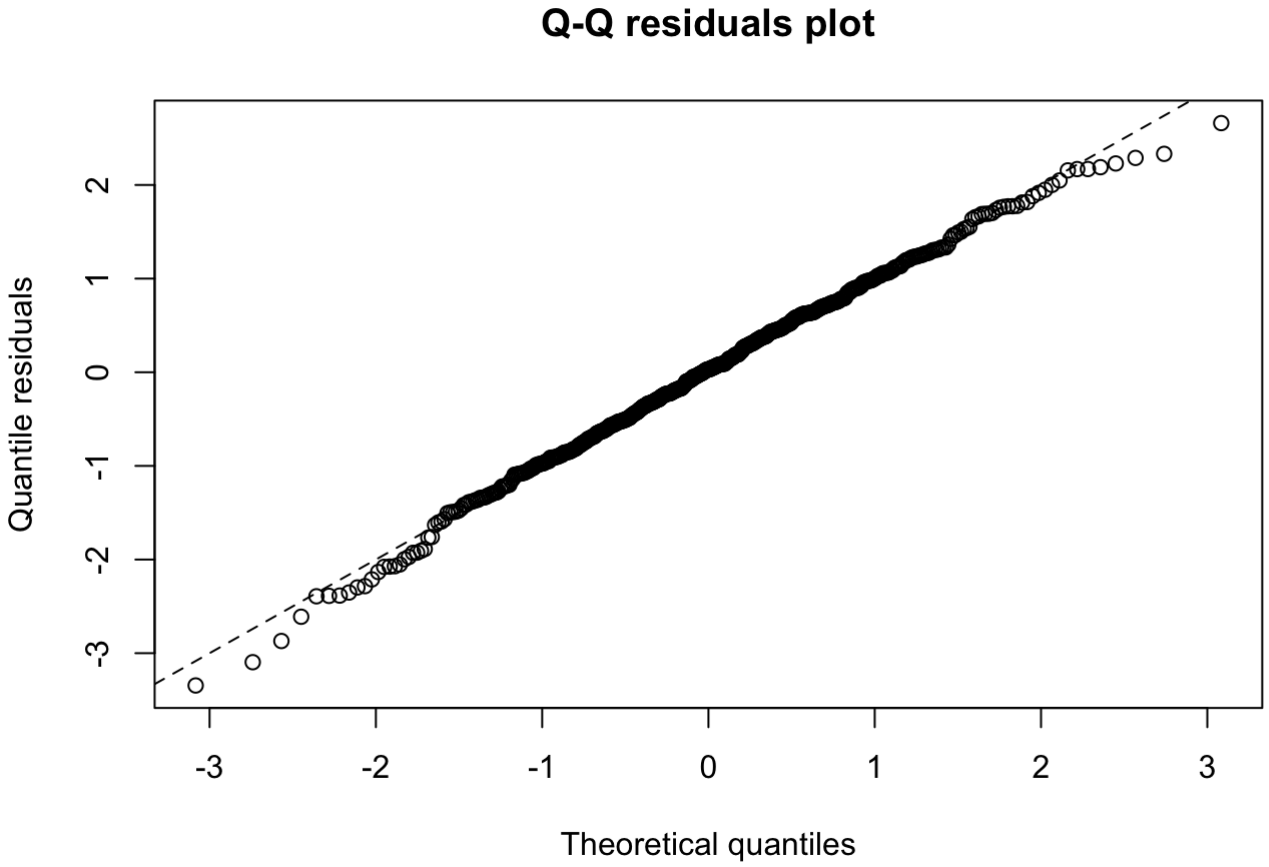
```
## rgeos version: 0.5-2, (SVN revision 621)
##  GEOS runtime version: 3.7.2-CAPI-1.11.2
##  Linking to sp version: 1.3-1
##  Polygon checking: TRUE
```

```
library(gtfsr)
library(sp)
```

```
chicago_final<-fread("/Users/11kolop/Desktop/chicago_final.csv")
dat.hom.chicago<-chicago_final[chicago_final$ofns_desc=="criminal homicide",]
pca <- princomp(na.omit(dat.hom.chicago)[,c(4:15)], cor = TRUE)

mod.zero.inflated.poisson<-zeroinfl(n ~ foreign_share2010 + share_black2010 + share_hisp
2010 + singleparent_share2010+mail_return_rate2010 + scale(transp/sqmi), data = na.omit
(dat.hom.chicago),dist="poisson")

mod.poisson<-glm(n ~ foreign_share2010 + share_black2010 + share_hisp2010 + singleparent
_share2010+mail_return_rate2010 + scale(transp/sqmi), data = na.omit(dat.hom.chicago),fa
mily="poisson")

mod.pca.zero.inflated.poisson<-zeroinfl(na.omit(dat.hom.chicago)$n~scale(na.omit(dat.ho
m.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)+pca$scores[,1] + pca$scores[,2]+pca$sco
res[,3] + pca$scores[,4],dist="poisson")

mod.pca.poisson<-glm(na.omit(dat.hom.chicago)$n~scale(na.omit(dat.hom.chicago)$transp/n
a.omit(dat.hom.chicago)$sqmi)+pca$scores[,1] + pca$scores[,2]+pca$scores[,3] + pca$score
s[,4],family="poisson")

mod.zero.inflated.nb<-zeroinfl(n ~ foreign_share2010 + share_black2010 + share_hisp2010
 + singleparent_share2010+mail_return_rate2010 + scale(transp/sqmi), data = na.omit(dat.
hom.chicago),dist="negbin")

mod.pca.zero.inflated.nb<-zeroinfl(na.omit(dat.hom.chicago)$n~scale(na.omit(dat.hom.chic
ago)$transp/na.omit(dat.hom.chicago)$sqmi)+pca$scores[,1] + pca$scores[,2]+pca$scores[,3
] + pca$scores[,4],dist="negbin")

BIC(mod.poisson,mod.zero.inflated.poisson,mod.pca.poisson,mod.pca.zero.inflated.poisson,
mod.zero.inflated.nb,mod.pca.zero.inflated.nb)
```

```
##                                df       BIC
## mod.poisson                     7 2844.371
## mod.zero.inflated.poisson      14 2741.713
## mod.pca.poisson                 6 2488.937
## mod.pca.zero.inflated.poisson 12 2399.338
## mod.zero.inflated.nb           15 2252.208
## mod.pca.zero.inflated.nb       13 2169.982
```

```
qqrplot(mod.pca.zero.inflated.nb)
```

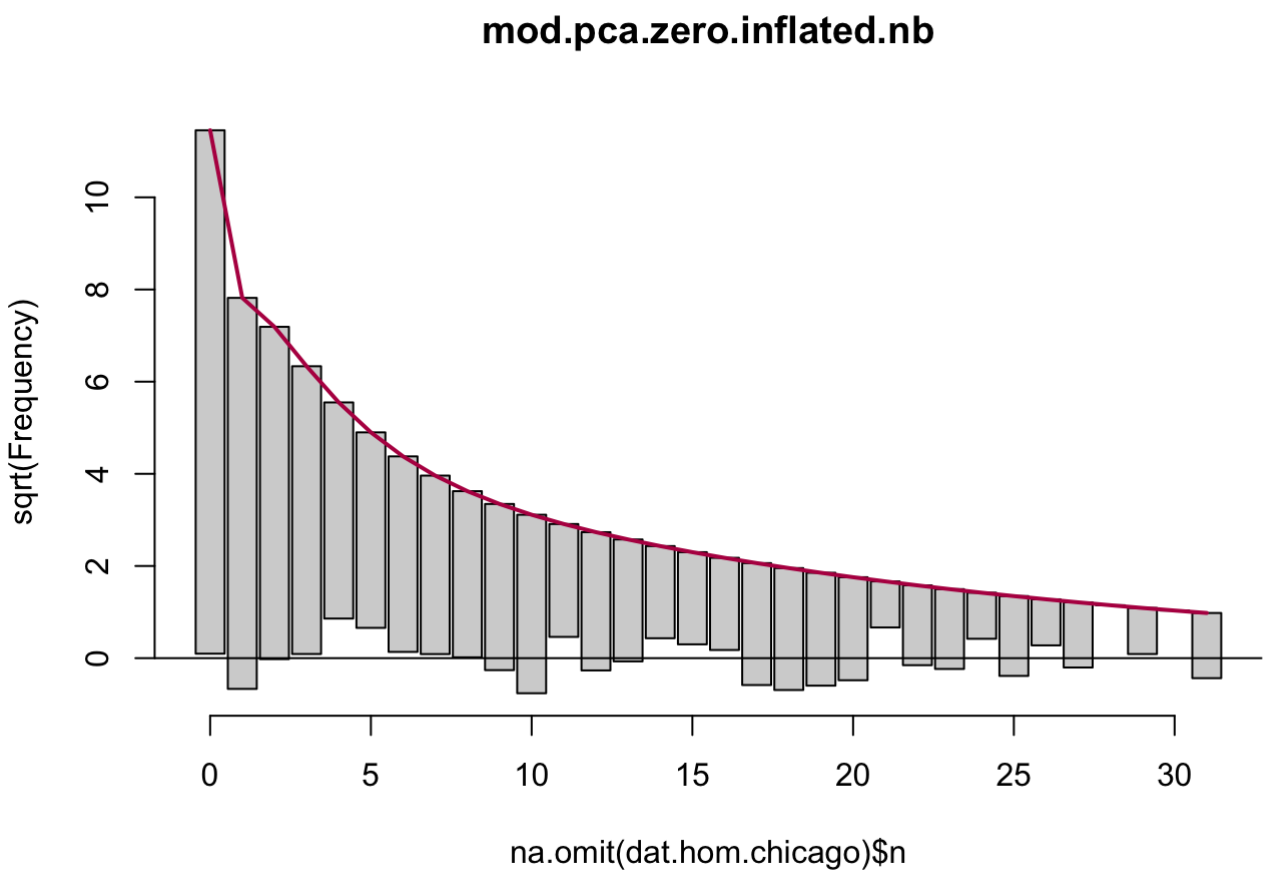



```
summary(mod.pca.zero.inflated.nb)
```

```
##
## Call:
## zeroinfl(formula = na.omit(dat.hom.chicago)$n ~ scale(na.omit(dat.hom.chicago)$trans
p/na.omit(dat.hom.chicago)$sqmi) +
##     pca$scores[, 1] + pca$scores[, 2] + pca$scores[, 3] + pca$scores[,
##     4], dist = "negbin")
##
## Pearson residuals:
##     Min      1Q  Median      3Q     Max
## -1.7443 -0.6702 -0.2405  0.4849  3.8149
##
## Count model coefficients (negbin with log link):
##                                                                          Estimate
## (Intercept)                                                              1.234267
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi) -0.030214
## pca$scores[, 1]                                                          0.458097
## pca$scores[, 2]                                                         -0.019313
## pca$scores[, 3]                                                         -0.055115
## pca$scores[, 4]                                                          0.009071
## Log(theta)                                                               1.372075
##                                                                         Std. Error
## (Intercept)                                                               0.050660
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)   0.052432
## pca$scores[, 1]                                                           0.022164
## pca$scores[, 2]                                                           0.024382
## pca$scores[, 3]                                                           0.040820
## pca$scores[, 4]                                                           0.050176
## Log(theta)                                                                0.146019
##                                                                         z value
## (Intercept)                                                              24.364
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)  -0.576
## pca$scores[, 1]                                                          20.669
## pca$scores[, 2]                                                          -0.792
## pca$scores[, 3]                                                          -1.350
## pca$scores[, 4]                                                           0.181
## Log(theta)                                                                9.397
##                                                                         Pr(>|z|)
## (Intercept)                                                               <2e-16
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)    0.564
## pca$scores[, 1]                                                           <2e-16
## pca$scores[, 2]                                                            0.428
## pca$scores[, 3]                                                            0.177
## pca$scores[, 4]                                                            0.857
## Log(theta)                                                                <2e-16
##
## (Intercept)                                                              ***
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)
## pca$scores[, 1]                                                          ***
## pca$scores[, 2]
## pca$scores[, 3]
## pca$scores[, 4]
## Log(theta)                                                               ***
##
```

```
## Zero-inflation model coefficients (binomial with logit link):
##                                                                    Estimate
## (Intercept)                                                         -3.8318
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)  -0.9916
## pca$scores[, 1]                                                     -1.1458
## pca$scores[, 2]                                                      0.9134
## pca$scores[, 3]                                                     -0.4569
## pca$scores[, 4]                                                      0.4405
##                                                                    Std. Error
## (Intercept)                                                            0.8073
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)     0.4498
## pca$scores[, 1]                                                        0.2834
## pca$scores[, 2]                                                        0.3761
## pca$scores[, 3]                                                        0.4554
## pca$scores[, 4]                                                        0.4329
##                                                                    z value
## (Intercept)                                                         -4.746
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)  -2.205
## pca$scores[, 1]                                                     -4.043
## pca$scores[, 2]                                                      2.428
## pca$scores[, 3]                                                     -1.003
## pca$scores[, 4]                                                      1.018
##                                                                    Pr(>|z|)
## (Intercept)                                                        2.07e-06
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)   0.0275
## pca$scores[, 1]                                                    5.27e-05
## pca$scores[, 2]                                                      0.0152
## pca$scores[, 3]                                                      0.3158
## pca$scores[, 4]                                                      0.3089
##
## (Intercept)                                                        ***
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi) *
## pca$scores[, 1]                                                    ***
## pca$scores[, 2]                                                    *
## pca$scores[, 3]
## pca$scores[, 4]
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Theta = 3.9435
## Number of iterations in BFGS optimization: 23
## Log-likelihood: -1045 on 13 Df
```

```
rootogram(mod.pca.zero.inflated.nb)
```


mod.pca.zero.inflated.nb
sqrt(Frequency)
0
2
4
6
8
10
0
5
10
15
20
25
30
na.omit(dat.hom.chicago)$n