

# Baltimore 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, hurdttest, 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)
library(stringi)
```

```
chicago_final<-fread("/Users/11kolop/Desktop/baltimore_final.csv")[,-c(1:2)]
dat.hom.chicago<-chicago_final[chicago_final$ofns_desc=="criminal homicide",]
dat.hom.chicago$transp<-as.numeric(as.character(dat.hom.chicago$transp))
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")
```

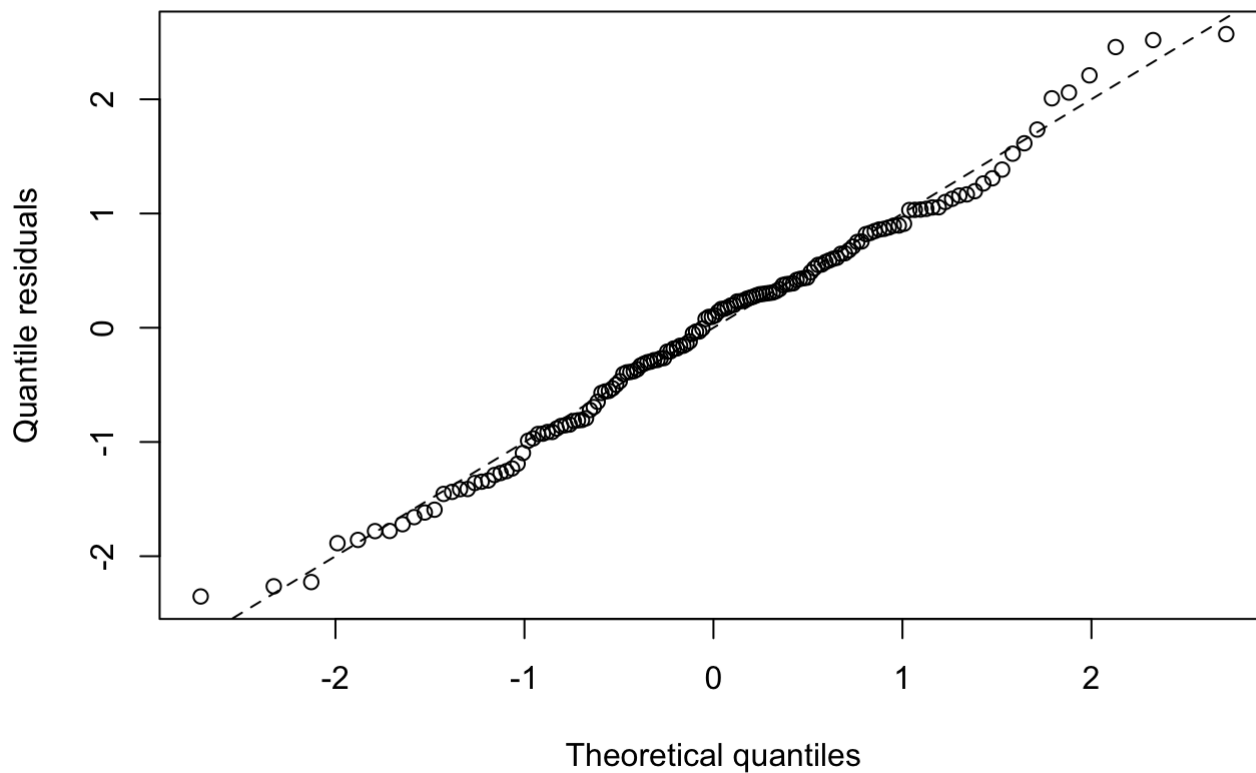
```
mod.pca.zero.inflated.nb<-zeroinfl(na.omit(dat.hom.chicago)$n~scale(na.omit(dat.hom.chic
ago)$transp/na.omit(dat.hom.chicago)$sqmi)+pca$scores[,1] + pca$scores[,2]+pca$scores[,3
] + pca$scores[,4],dist="negbin")
```

```
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 1093.1421
## mod.zero.inflated.poisson 14 1112.4539
## mod.pca.poisson   6 1107.9387
## mod.pca.zero.inflated.poisson 12 1116.4972
## mod.zero.inflated.nb 15 916.2901
## mod.pca.zero.inflated.nb 13 917.1289
```

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

### Q-Q residuals plot



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

```
##
## Call:
## zeroinfl(formula = na.omit(dat.hom.chicago)$n ~ scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi) +
##      pca$scores[, 1] + pca$scores[, 2] + pca$scores[, 3] + pca$scores[,
##      4], dist = "negbin")
##
## Pearson residuals:
##      Min      1Q  Median      3Q      Max
## -1.4290 -0.7308 -0.1786  0.4728  3.5342
##
## Count model coefficients (negbin with log link):
##
##                                     Estimate
## (Intercept)                        1.93931
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)  0.05616
## pca$scores[, 1]                    -0.37256
## pca$scores[, 2]                    -0.10052
## pca$scores[, 3]                    -0.09266
## pca$scores[, 4]                    -0.11498
## Log(theta)                         1.16956
##
##                                     Std. Error
## (Intercept)                        0.06525
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)  0.08318
## pca$scores[, 1]                    0.03690
## pca$scores[, 2]                    0.04365
## pca$scores[, 3]                    0.06129
## pca$scores[, 4]                    0.05942
## Log(theta)                        0.17454
##
##                                     z value
## (Intercept)                        29.721
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)  0.675
## pca$scores[, 1]                   -10.098
## pca$scores[, 2]                    -2.303
## pca$scores[, 3]                    -1.512
## pca$scores[, 4]                    -1.935
## Log(theta)                        6.701
##
##                                     Pr(>|z|)
## (Intercept)                        < 2e-16
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)  0.4996
## pca$scores[, 1]                    < 2e-16
## pca$scores[, 2]                    0.0213
## pca$scores[, 3]                    0.1306
## pca$scores[, 4]                    0.0530
## Log(theta)                        2.07e-11
##
## (Intercept)                        ***
## scale(na.omit(dat.hom.chicago)$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):
##
## (Intercept) Estimate
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi) -4.1005
## pca$scores[, 1] 0.6483
## pca$scores[, 2] -0.7765
## pca$scores[, 3] 0.9147
## pca$scores[, 4] -0.3287
## Std. Error
## (Intercept) 3.7116
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi) 3.8713
## pca$scores[, 1] 0.4483
## pca$scores[, 2] 0.5291
## pca$scores[, 3] 0.6570
## pca$scores[, 4] 0.5619
## z value
## (Intercept) -1.915
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi) -1.059
## pca$scores[, 1] 1.446
## pca$scores[, 2] -1.468
## pca$scores[, 3] 1.392
## pca$scores[, 4] -0.585
## Pr(>|z|)
## (Intercept) 0.0555
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi) 0.2895
## pca$scores[, 1] 0.1482
## pca$scores[, 2] 0.1422
## pca$scores[, 3] 0.1639
## pca$scores[, 4] 0.5585
##
## (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 = 3.2206
## Number of iterations in BFGS optimization: 29
## Log-likelihood: -426 on 13 Df
```

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

mod.pca.zero.inflated.nb

