

ZERO INFLATED MODELS

BY IAN CUMMINGS

WHAT IS ZERO INFLATED DATA?

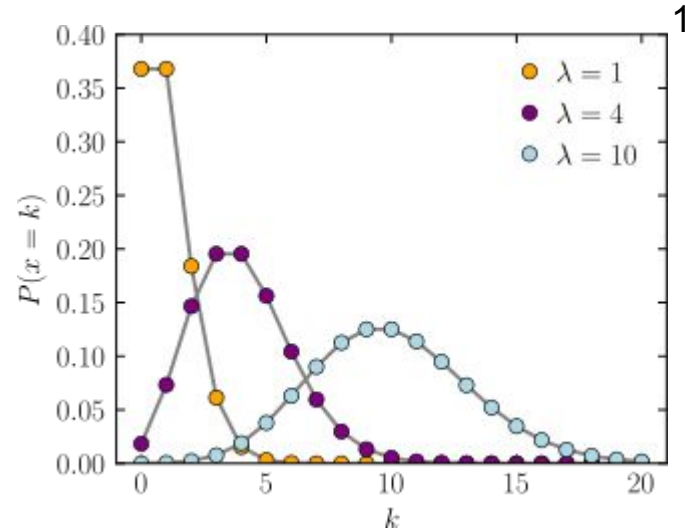
DATA WITH AN ABUNDANCE OF
ZERO-VALUED OBSERVATIONS.

EXAMPLES

NUMBER OF INSURANCE CLAIMS FILED

NUMBER OF KIDS PER COUPLE

POISSON DISTRIBUTION



MEET THE MODELS

ZERO-INFLATED POISSON MODEL

HURDLE MODEL

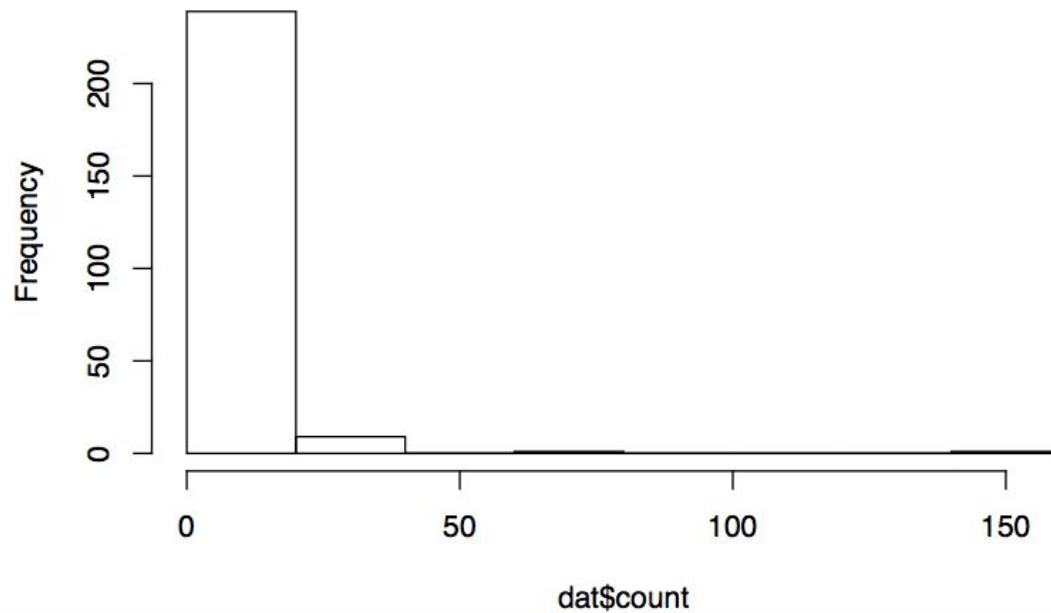
AN EXAMPLE IN R



<https://stats.idre.ucla.edu/stat/data/fish.csv>

HISTOGRAM

Histogram of dat\$count



SUMMARY

```
summary(dat)
```

```
##  nofish  livebait  camper      persons      child      xb
##  0:176    0: 34    0:103  Min.    :1.000  Min.    :0.000  Min.    : -3.275050
##  1: 74    1:216    1:147  1st Qu.:2.000  1st Qu.:0.000  1st Qu.: 0.008267
##                                     Median :2.000  Median :0.000  Median : 0.954550
##                                     Mean    :2.528  Mean    :0.684  Mean    : 0.973796
##                                     3rd Qu.:4.000  3rd Qu.:1.000  3rd Qu.: 1.963855
##                                     Max.    :4.000  Max.    :3.000  Max.    : 5.352674
##
##      zg      count
##  Min.    : -5.6259  Min.    : 0.000
##  1st Qu.: -1.2527  1st Qu.: 0.000
##  Median : 0.6051  Median : 0.000
##  Mean    : 0.2523  Mean    : 3.296
##  3rd Qu.: 1.9932  3rd Qu.: 2.000
##  Max.    : 4.2632  Max.    :149.000
```


MODEL COMPARISONS

```
#the models  
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
```

```
zip <- zeroinfl(count ~ child + camper + persons, data = dat)  
AIC(zip)
```

```
## [1] 1521.463
```

```
hurdle <- hurdle(count ~ child + camper + persons, data = dat)  
AIC(hurdle)
```

```
## [1] 1519.236
```

THANK YOU