PLSC 503 - Fall 2020
Event Counts, II

April 23, 2020


Seattle
WASHINGTON
MONTANA
NORTH DAKOTA
MINNESOTA
Ottawa
NB
OREGON
IDAHO
WYOMING
SOUTH DAKOTA
WISCONSIN
MICHIGAN
Toronto
VT
MAINE
NH
NEW YORK
MA
Boston
CT
RI
Chicago
IOWA
NEBRASKA
ILLINOIS
OHIO
PENN
New York
NEVADA
UTAH
United States
COLORADO
KANSAS
MISSOURI
INDIANA
MD
DE
NJ
WEST VIRGINIA
VIRGINIA
San Francisco
KENTUCKY
CALIFORNIA
Las Vegas
OKLAHOMA
TENNESSEE
NORTH CAROLINA
Los Angeles
ARIZONA
NEW MEXICO
ARKANSAS
MISSISSIPPI
SOUTH CAROLINA
ALABAMA
San Diego
TEXAS
GEORGIA
LOUISIANA
Houston
Gulf of California
FLORIDA
Gulf of Mexico
Mexico


Lisco
26
Oshkosh
27
Lewellen
26

Heterogeneity, Contagion, and Dispersion

Cats:

$$
\begin{aligned}
Y_{cats} &= \{0,1,1,0,2,0,1,0,3,1,2,1,0,2\} \\
\bar{Y}_{cats} &= 1.0, \\
\sigma_{cats} &= 0.92.
\end{aligned}
$$

Heterogeneity, Contagion, and Dispersion

$$\mathrm{E}(Y_{cats}) = \lambda_{cats}$$

Assumes:

- $Y = 0$ at $t = 0$,
- Exclusive events
- $t_j = t_k \, \forall \, j \neq k$
- Constant, independent Pr(Event) over t

Antelope

$$\begin{aligned} Y_{antelope} &= \{0,0,0,0,0,0,0,0,0,0,0,0,7,7\} \\ \bar{Y}_{antelope} &= 1.0, \\ \sigma_{antelope} &= 6.46. \end{aligned}$$

Positive contagion $\rightarrow$ *overdispersion*.

$$
\begin{aligned}
Y_{foxes} &= \{1,0,1,1,1,1,1,2,1,1,1,1,1,1\} \\
\bar{Y}_{foxes} &= 1.0, \\
\sigma_{foxes} &= 0.15.
\end{aligned}
$$

Negative contagion $\rightarrow$ *underdispersion*.

Aggregation & Cross-Period Effects

$$
\begin{aligned}
Y_{cats} &= \{1, 1, 2, 1, 4, 3, 2\} \\
Y_{antelope} &= \{0, 0, 0, 0, 0, 0, 14\} \\
Y_{foxes} &= \{1, 2, 2, 3, 2, 2, 2\}
\end{aligned}
$$

- Correct specification
- Correct distribution for ϵ
- Constant $\mathrm{E}(Y|\mathbf{X}, \boldsymbol{\beta})$

$$\lambda_i \equiv \mathrm{E}(Y_i) = f[\mathbf{X}_i\boldsymbol{\beta} + Z_i\theta]$$

Overdispersion: A Test

Examine:

$$\hat{u}_i = \delta\hat{\lambda}_i + \epsilon_i$$

where

$$\hat{u}_i = \frac{(Y_i - \hat{\lambda}_i)^2 - Y_i}{\hat{\lambda}_i\sqrt{2}}$$

- Estimate a Poisson regression of Y_i on $\mathbf{X}_i$, and generate predicted counts $\hat{\lambda}_i$.
- Calculate $\hat{u}_i$ according to the equation above.
- Estimate δ using OLS, and test $H_0 : \hat{\delta} = 0$.

$$\begin{aligned} \mathrm{E}(Y_i) \equiv \lambda_i &= \exp(\mathbf{X}_i\boldsymbol{\beta} + u_i) \\ &= \exp(\mathbf{X}_i\boldsymbol{\beta})\exp(u_i) \\ &= \lambda_i\nu_i \end{aligned}$$

$$\nu_i \sim \text{gamma}\left(1, \frac{1}{\alpha}\right)$$

$$\Pr(Y_i = y|\lambda_i, \alpha) = \left(\frac{\Gamma(\alpha^{-1} + Y_i)}{\Gamma(\alpha^{-1})\Gamma(Y_i + 1)}\right)\left(\frac{\alpha^{-1}}{\alpha^{-1} + \lambda_i}\right)^{\alpha^{-1}}\left(\frac{\lambda_i}{\lambda_i + \alpha^{-1}}\right)^{Y_i}$$

where

$$\Gamma(a) = \int_0^\infty \exp(-t)t^{a-1}dt$$

Negative Binomial

Basis:

$$\lambda_i = \exp(\mathbf{X}_i \boldsymbol{\beta})$$

Model has

$$\mathrm{E}(Y) = \lambda$$

$$\mathrm{Var}(Y) = \lambda(1 + \alpha\lambda), \quad \alpha > 0$$

or (equivalently):

$$\mathrm{Var}(Y) = \lambda + \frac{\lambda^2}{\theta}, \text{ where } \theta = \frac{1}{\alpha}.$$

Negative Binomial (log-)Likelihood

$$\ln L_{NB} = \sum_{i=1}^{N} \left\{ \left(\sum_{j=0}^{Y_i - 1} \ln(j + \alpha^{-1}) \right) - \ln Y_i! \right.$$
$$\left. - (Y_i - \alpha^{-1}) \ln[1 + \alpha \exp(\mathbf{X}_i \boldsymbol{\beta})] + Y_i \ln \alpha + Y_i \mathbf{X}_i \boldsymbol{\beta} \right\}$$

Note that:

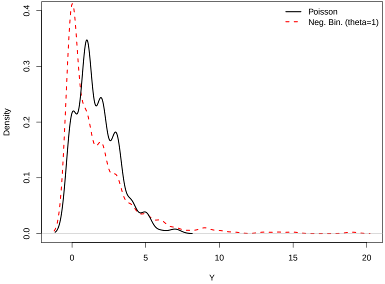
- $\alpha = 0 \iff \mathrm{E}(Y) = \mathrm{Var}(Y)$
- LR test for overdispersion:

$$-2 \times (\widehat{\ln L_{Poisson}} - \widehat{\ln L_{NB}}) \sim \chi^2_1$$

- $\widehat{E(Y_i)} \equiv \hat{\lambda}_i = \exp(\mathbf{X}_i \hat{\boldsymbol{\beta}})$

What Difference Does It Make?

```
> N<-400
> set.seed(7222009)
> X <- runif(N,min=0,max=1)
> YPois <- rpois(N,exp(0+1*X))            # Poisson
> YNB <- rnbinom(N,size=1,mu=exp(0+1*X)) # NB with theta=1.0
>
> describe(cbind(YPois,YNB))
      vars   n mean   sd median trimmed  mad min max range skew kurtosis   se
YPois    1 400 1.72 1.41      1    1.56 1.48   0   7     7 0.92     0.84 0.07
YNB      2 400 1.71 2.44      1    1.22 1.48   0  19    19 2.76    11.15 0.12
```



What Difference Does It Make (cont'd)?

```
> # Regressions:
>
> summary(glm(YPois~X,family="poisson")) # Poisson

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.009637   0.085337  -0.113     0.91
X            1.030573   0.131992   7.808 5.82e-15 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 516.06  on 399  degrees of freedom
Residual deviance: 453.55  on 398  degrees of freedom
AIC: 1274.4



> summary(glm.nb(YPois~X))               # NB

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.009629   0.085345  -0.113     0.91
X            1.030557   0.132007   7.807 5.86e-15 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for Negative Binomial(7837.699) family taken to be 1)

    Null deviance: 515.96  on 399  degrees of freedom
Residual deviance: 453.46  on 398  degrees of freedom
AIC: 1276.5

              Theta:  7838
          Std. Err.:  135342
Warning while fitting theta: iteration limit reached

 2 x log-likelihood:  -1270.451
```

What Difference Does It Make (cont'd)?

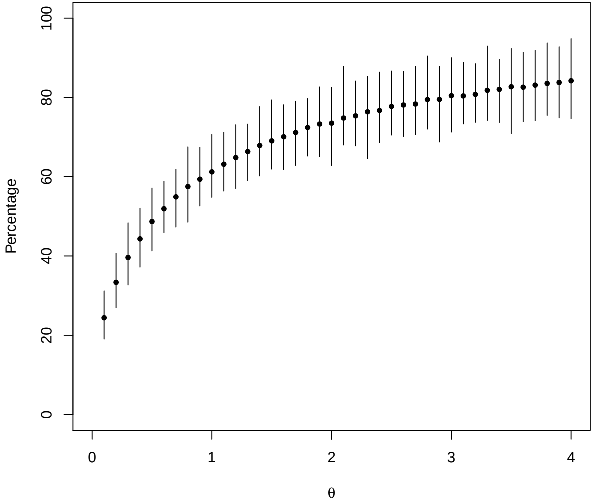
```
> # More regressions:
>
> summary(glm(YNB~X,family="poisson"))  # Poisson

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.03170    0.08593  -0.369    0.712
X            1.06109    0.13248   8.009 1.15e-15 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 1118.0  on 399  degrees of freedom
Residual deviance: 1052.1  on 398  degrees of freedom
AIC: 1698.6
```

```
> summary(glm.nb(YNB~X))               # NB

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.03525    0.13650  -0.258    0.796
X            1.06773    0.22809   4.681 2.85e-06 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for Negative Binomial(0.8499) family taken to be 1)

    Null deviance: 436.92  on 399  degrees of freedom
Residual deviance: 414.81  on 398  degrees of freedom
AIC: 1407.4

              Theta:  0.850
          Std. Err.:  0.109

 2 x log-likelihood:  -1401.354
```

Poisson Regression Underestimates N.B. Variances

```
Sims <- 250 # (250 sims each)
theta <- seq(0.1,4,by=0.1) # values of theta
diffs<-matrix(nrow=Sims,ncol=length(theta))

set.seed(7222009)
for(j in 1:length(theta)) {
  for(i in 1:Sims) {
    X<-runif(N,min=0,max=1)
    Y<-rnbinom(N,size=theta[j],mu=exp(0+1*X))
    p<-glm(Y~X,family="poisson")
    nb<-glm.nb(Y~X)
    diffs[i,j]<- ((sqrt(vcov(p))[2,2]) / sqrt(vcov(nb))[2,2])*100
  }
}
```

Percentage of True (Negative Binomial) S.E. From Fitted Poisson Model, by $\theta = \frac{1}{\alpha}$



Negative Binomial In Practice

Model fitting (in R):

- `glm.nb` (in `MASS`)
- `negbinomial` (in `VGAM`)
- `negbin` (in `aod`)
- `glmnb.fit` (in `statmod`)
- probably others...

Model interpretation + diagnostics:

- `fitNBP` (in `statmod`) (dispersion parameter estimation)
- `negbinirr` (in `mfx`) (IRRs)
- `negbinmfx` (in `mfx`) (marginal effects)
- Predicted values / probabilities via `predict`

Negative Binomial In Practice

Model fitting (in R):

- `glm.nb` (in `MASS`)
- `negbinomial` (in `VGAM`)
- `negbin` (in `aod`)
- `glmnb.fit` (in `statmod`)
- Probably others...

Model interpretation + diagnostics:

- `fitNBP` (in `statmod`) (dispersion parameter estimation)
- `negbinirr` (in `mfx`) (IRRs)
- `negbinmfx` (in `mfx`) (marginal effects)

Underdispersion / CPB

"Continuous parameter binomial":

$$\Pr(Y_i = y|\lambda_i, \alpha) = \frac{\frac{\Gamma\left(\frac{-\lambda_i}{\alpha-1}+1\right)}{Y_i!\Gamma\left(\frac{-\lambda_i}{\alpha-1}-Y_i+1\right)}(1-\alpha)^{Y_i}(\alpha)^{\frac{-\lambda_i}{\alpha-1}-Y_i}}{D_i}$$

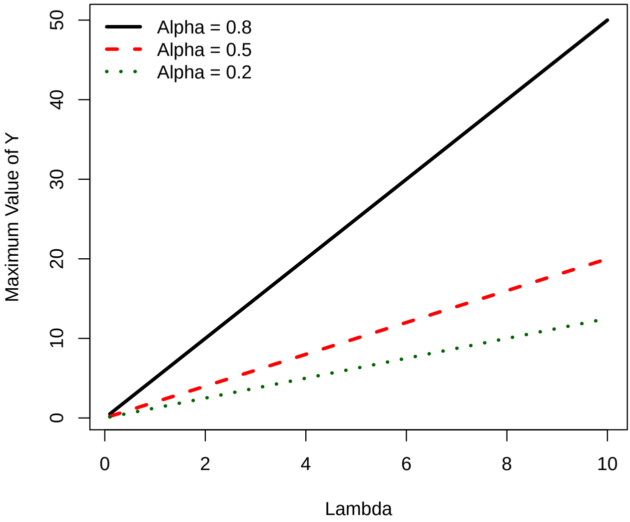
where $D_i = \sum_0^{\frac{-\lambda_i}{\alpha-1}+1}$ of the binomial distribution...

Are You Down With The CPB?

CPB:

- ...also has $\mathrm{E}(Y_i) = \lambda_i$ [with $\mu_i = \exp(\mathbf{X}_i\boldsymbol{\beta})$]
- ...has $\mathrm{Var}(Y) = \lambda_i\alpha$ with $0 < \alpha < 1$
- ... reduces to the standard Poisson when $\alpha = 1$
- ...imposes a theoretical "upper limit" on the count variable. In particular,

$$\max(Y_i) = \frac{-\lambda_i}{\alpha - 1}.$$



Alpha = 0.8
Alpha = 0.5
Alpha = 0.2
Maximum Value of Y
Lambda
0
10
20
30
40
50
0
2
4
6
8
10

CPB (log-)Likelihood

$$\begin{aligned} \ln L_{CPB} &= \sum_{i=1}^{N} \Big\{ \ln\Gamma\left(\frac{-\lambda_i}{\alpha-1}+1\right) - \ln\Gamma\left(\frac{-\lambda_i}{\alpha-1} - Y_i + 1\right) \\ &\quad + Y_i \ln(1-\alpha) + \left(\frac{-\lambda_i}{\alpha-1} - Y_i\right)\ln(\alpha) - \ln(D_i) \Big\} \end{aligned}$$

Example: SCOTUS amicus curiae (1953-85)

"Friend of the Court" briefs...

- $N = 7157$
- `namici` is the number of amicus curiae briefs filed in each case,
- `term` is the term (i.e., year) of the Court,
- `civlibs` is whether (=1) or not (=0) the case involved a civil rights and liberties issue.

```
> describe(amici)
        vars    n  mean   sd median trimmed   mad min max range  skew kurtosis   se
namici     1 7157  1.03 2.54      0    0.42  0.00   0  53    53  5.63    56.36 0.03
term       2 7157 71.12 9.19     72   71.46 11.86  53  85    32 -0.24    -1.07 0.11
civlibs    3 7157  0.50 0.50      1    0.50  0.00   0   1     1  0.00    -2.00 0.01
```

Amicus Example: Poisson

```
> amici.poisson<-glm(namici~term+civlibs,data=amici,family="poisson")
> summary(amici.poisson)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.51196    0.11190  -40.32   <2e-16 ***
term         0.06361    0.00147   43.27   <2e-16 ***
civlibs     -0.29797    0.02350  -12.68   <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: 22875  on 7156  degrees of freedom
Residual deviance: 20675  on 7154  degrees of freedom
  (4 observations deleted due to missingness)
AIC: 26862

Number of Fisher Scoring iterations: 6
```

Overdispersion Test: "By Hand"

```
> Phats<-fitted.values(amici.poisson)
> Uhats<-((amici$namici-Phats)^2 - amici$namici) / (Phats * sqrt(2))
> summary(lm(Uhats~Phats))

Call:
lm(formula = Uhats ~ Phats)

Residuals:
   Min     1Q Median     3Q    Max
  -5.9   -3.0   -2.3   -1.9 1707.0

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.579      0.693    2.28    0.023 *
Phats          1.466      0.591    2.48    0.013 *
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 28.4 on 7155 degrees of freedom
Multiple R-squared:  0.000858,Adjusted R-squared:  0.000718
F-statistic: 6.14 on 1 and 7155 DF,  p-value: 0.0132
```

Negative Binomial Regression

```
> library(MASS)
> amici.NB<-glm.nb(namici~term+civlibs,data=amici)
> summary(amici.NB)

Call:
glm.nb(formula = namici ~ term + civlibs, data = amici, init.theta = 0.256657474,
    link = log)

Coefficients:
            Estimate Std. Error z value             Pr(>|z|)
(Intercept) -4.68314    0.22058  -21.23 < 0.0000000000000002 ***
term         0.06573    0.00304   21.60 < 0.0000000000000002 ***
civlibs     -0.26777    0.05403   -4.96           0.00000072 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for Negative Binomial(0.2567) family taken to be 1)

    Null deviance: 5442  on 7156  degrees of freedom
Residual deviance: 4968  on 7154  degrees of freedom
AIC: 17378

Number of Fisher Scoring iterations: 1
              Theta:  0.25666
          Std. Err.:  0.00838
```

Poisson vs. NB

```
> # alpha:
>
> 1 / amici.NB$theta
[1] 3.896


> # Coefficient estimates:
>
> cbind(amici.poisson$coefficients,coef(amici.NB))
                [,1]     [,2]
(Intercept) -4.51196 -4.68314
term         0.06361  0.06573
civlibs     -0.29797 -0.26777


> # Estimated standard errors:
>
> cbind(diag(sqrt(vcov(amici.poisson))),diag(sqrt(vcov(amici.NB))))
               [,1]     [,2]
(Intercept) 0.11190 0.220582
term        0.00147 0.003043
civlibs     0.02350 0.054032
```

Predicted Values: Poisson and NB

```
> plot(amici.poisson$fitted.values,amici.NB$fitted.values,pch=20,
    xlab="Poisson",ylab="Negative Binomial",main="",
    xlim=c(0,3),ylim=c(0,3))
> abline(a=0,b=1,lwd=1,lty=2)
```