PLSC 503 – Spring 2021
Event Counts

April 21, 2021

Things that are not counts:

- Ordinal scales/variables
- Grouped Binary Data
  - $\frac{N \text{ of ``successes''}}{N \text{ of ``trials''}}$
  - Binomial data
  - $=$ counts only if Pr("success") is small

Count properties:

- Discrete / integer-values
- Non-negative
- "Cumulative"

$$\text{Arrival Rate} = \lambda$$

$$\Pr(\text{Event})_{t,t+h} = \lambda h$$

$$\Pr(\text{No Event})_{t,t+h} = 1 - \lambda h$$

$$\begin{aligned} \Pr(Y_t = y) &= \frac{\exp(-\lambda h)\lambda h^y}{y!} \\ &= \frac{\exp(-\lambda)\lambda^y}{y!} \end{aligned}$$

Poisson Assumptions

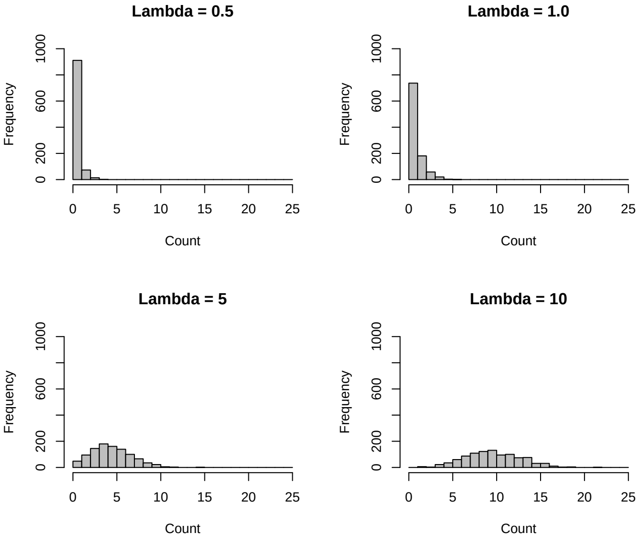
- No Simultaneous Events
- Constant Arrival Rate
- Independent Event Arrivals

For M independent Bernoulli trials with (sufficiently small) probability of success π and where $M\pi \equiv \lambda > 0$,

$$\begin{aligned} \Pr(Y_i = y) &= \lim_{M\to\infty}\left[\binom{M}{y}\left(\frac{\lambda}{M}\right)^y\left(1-\frac{\lambda}{M}\right)^{M-y}\right] \\ &= \frac{\lambda^y \exp(-\lambda)}{y!} \end{aligned}$$

- Discrete
- $\mathrm{E}(Y) = \mathrm{Var}(Y) = \lambda$
- Is not preserved under affine transformations...
- For $X \sim \text{Poisson}(\lambda_X)$ and $Y \sim \text{Poisson}(\lambda_Y)$, $Z = X + Y \sim \text{Poisson}(\lambda_{X+Y})$ *iff* X and Y are *independent* but
- ...same is not true for differences.
- $\lambda \to \infty \iff Y \sim N$

Poissons: Examples



Suppose

$$\mathrm{E}(Y_i) \equiv \lambda_i = \exp(\mathbf{X}_i\boldsymbol{\beta})$$

then

$$\Pr(Y_i = y|\mathbf{X}_i, \boldsymbol{\beta}) = \frac{\exp[-\exp(\mathbf{X}_i\boldsymbol{\beta})][\exp(\mathbf{X}_i\boldsymbol{\beta})]^y}{y!}$$

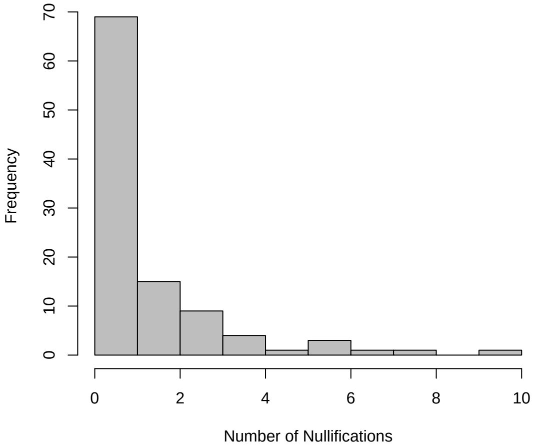
$$L = \prod_{i=1}^{N} \frac{\exp[-\exp(\mathbf{X}_i\boldsymbol{\beta})][\exp(\mathbf{X}_i\boldsymbol{\beta})]^{Y_i}}{Y_i!}$$

$$\ln L = \sum_{i=1}^{N}[-\exp(\mathbf{X}_i\boldsymbol{\beta}) + Y_i\mathbf{X}_i\boldsymbol{\beta} - \ln(Y_i!)]$$

- Y_i = number of Acts of Congress overturned by the Supreme Court in each Congress,
- The *mean tenure* (`tenure`) of the Supreme Court's justices ($\bar{X} = 10.4, \sigma = 3.4, \mathrm{E}(\hat{\beta}) > 0$).
- Whether (1) or not (0) there was *unified government* (`unified`) ($\bar{X} = 0.83, \mathrm{E}(\hat{\beta}) < 0$).

Supreme Court Nullifications, 1789-1996

```
> Nulls<-read.dta("nulls.dta")
> hist(Nulls$nulls,main="",xlab="Number of Nullifications",
    col="grey")
```



```
> nulls.poisson<-glm(nulls~tenure+unified,family="poisson",data=Nulls)
> summary(nulls.poisson)

Call:
glm(formula = nulls ~ tenure + unified, family = "poisson", data = Nulls)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
-2.367  -1.503  -0.623   0.561   4.153

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -0.8776     0.3713   -2.36  0.01809 *
tenure        0.0959     0.0256    3.74  0.00018 ***
unified       0.1435     0.2327    0.62  0.53747
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 251.80  on 103  degrees of freedom
Residual deviance: 237.52  on 101  degrees of freedom
AIC: 385.1

Number of Fisher Scoring iterations: 6
```

Interpretation: Incidence Rate Ratios

$$
\begin{aligned}
\frac{\hat{\lambda}|X_D = 1}{\hat{\lambda}|X_D = 0} &= \frac{\exp(\hat{\beta}_0 + \bar{\mathbf{X}}\hat{\boldsymbol{\beta}} + (X_D = 1)\hat{\beta}_{X_D})}{\exp(\hat{\beta}_0 + \bar{\mathbf{X}}\hat{\boldsymbol{\beta}} + (X_D = 0)\hat{\beta}_{X_D})} \\
&= \exp(\hat{\beta}_{X_D})
\end{aligned}
$$

- Like ORs
- `unified`: IRR$=\exp(0.143) = 1.15$

Incidence Rate Ratios, continued

$$\text{IRR}_{X_k, X_{k+\delta}} = \exp(\delta \hat{\beta}_k)$$

So, a ten-year difference in `tenure`:

$$\begin{aligned} \text{IRR} &= \exp(10 \times 0.096) \\ &= \exp(0.96) \\ &= 2.61 \end{aligned}$$

```
> library(mfx)
> nulls.poisson.IRR<-poissonirr(nulls~tenure+unified,
                    data=Nulls)
> nulls.poisson.IRR

Call:
poissonirr(formula = nulls ~ tenure + unified, data = Nulls)

Incidence-Rate Ratio:
           IRR Std. Err.    z   P>|z|
tenure  1.1006    0.0282 3.74 0.00018 ***
unified 1.1543    0.2686 0.62 0.53747
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

Mean predicted Y:

$$\begin{aligned} E(Y|\bar{\mathbf{X}}_i) &= \exp[-0.878 + (0.096 \times 10) + (0.143 \times 1)] \\ &= \exp(0.225) \\ &= 1.25 \end{aligned}$$

In-Sample

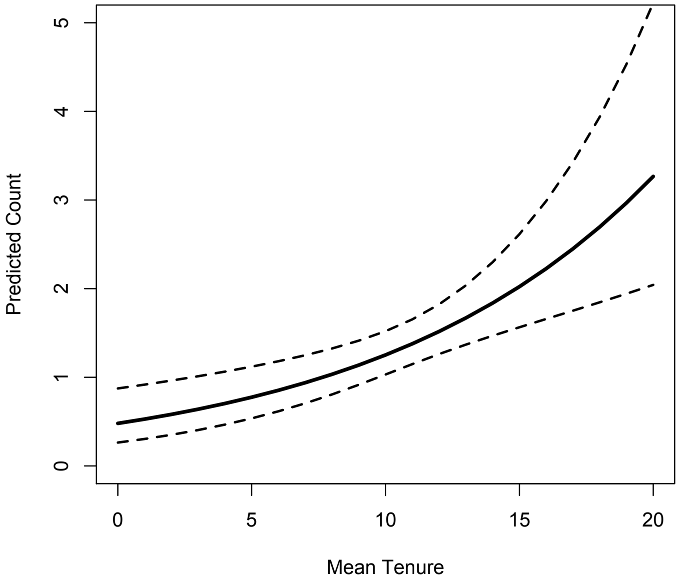
- R : in `$fitted.values`
- Stata : use `predict`

Out-of-Sample: use `predict`

Example: Out-Of-Sample Predicted Values

```
> tenure<-seq(0,20,1)
> unified<-1
> simdata<-as.data.frame(cbind(tenure,unified))
> nullhats<-predict(nulls.poisson,newdata=simdata,se.fit=TRUE)
>
> # NOTE: These are XBs, not predicted counts.
> # Transforming:
> nullhats$Yhat<-exp(nullhats$fit)
> nullhats$UB<-exp(nullhats$fit + 1.96*(nullhats$se.fit))
> nullhats$LB<-exp(nullhats$fit - 1.96*(nullhats$se.fit))
> plot(simdata$tenure,nullhats$Yhat,t="l",lwd=3,ylim=c(0,5),ylab=
+           "Predicted Count", xlab="Mean Tenure")
> lines(simdata$tenure,nullhats$UB,lwd=2,lty=2)
> lines(simdata$tenure,nullhats$LB,lwd=2,lty=2)
>
> plot(simdata$tenure,nullhats$Yhat,t="l",lwd=3,ylim=c(0,5),ylab=
+           "Predicted Count", xlab="Mean Tenure")
> lines(simdata$tenure,nullhats$UB,lwd=2,lty=2)
> lines(simdata$tenure,nullhats$LB,lwd=2,lty=2)
```

Plotting Out-Of-Sample Predicted Values



$$\widehat{\Pr(Y_i = y}|\mathbf{X}_i, \hat{\boldsymbol{\beta}}) = \frac{\exp[-\exp(\mathbf{X}_i\hat{\boldsymbol{\beta}})][\exp(\mathbf{X}_i\hat{\boldsymbol{\beta}})]^y}{y!}$$

$$\begin{aligned}
\rightarrow \ \widehat{\Pr(Y_i = 0}|\bar{\mathbf{X}}_i, \hat{\boldsymbol{\beta}}) &= \frac{[\exp(-1.25)](1.25)^0}{0!} \\
&= \frac{(0.287)(1)}{1} \\
&= 0.287
\end{aligned}$$

$$\begin{aligned}
\widehat{\Pr(Y_i = 1}|\bar{\mathbf{X}}_i, \hat{\boldsymbol{\beta}}) &= \frac{[\exp(-1.25)](1.25)^1}{1!} \\
&= \frac{(0.287)(1.25)}{1} \\
&= 0.359
\end{aligned}$$

$$
\begin{aligned}
\Pr(\widehat{Y_i = 2}|\bar{\mathbf{X}}_i, \hat{\beta}) &= \frac{[\exp(-1.25)](1.25)^2}{2!} \\
&= \frac{(0.287)(1.563)}{2} \\
&= 0.224
\end{aligned}
$$

$$
\begin{aligned}
\Pr(\widehat{Y_i = 3}|\bar{\mathbf{X}}_i, \hat{\beta}) &= \frac{[\exp(-1.25)](1.25)^3}{3!} \\
&= \frac{(0.287)(1.953)}{6} \\
&= 0.093
\end{aligned}
$$

"Exposure" and "Offsets"

$$E(Y_i|\mathbf{X}_i, M_i) = \lambda_i M_i$$

Same as including $\ln(M_i)$ in $\mathbf{X}$ with $\beta_{\ln M} = 1$.

Example: Data on numbers of interstate disputes by country, 1950-1985...

- $N = 102$, but
- `Ndyads` = number of dyad-years which were aggregated to create each observation, ranging from five to 3249
- `disputes` = number of (interstate) dispute-years that country experienced during 1950-1985
- `allies` = number of (dyadic) ally-years each country had during 1950-1985
- `openness` $= \frac{1}{36}\left(\frac{\text{Imports}_t + \text{Exports}_t}{\text{GDP}_t}\right)$ across all 36 years in the data.

“Exposure” and “Offsets”: Data

```
# Data are aggregated dyadic data, 1950-1985...

> IR<-read.csv("Data/offsetIR.csv")
> summary(IR)
     ccode          Ndyads         disputes          allies          openness         exposure
 Min.   :  2   Min.   :   5   Min.   : 0.00   Min.   :   0.0   Min.   :0.032   Min.   :1.61
 1st Qu.:214   1st Qu.:  44   1st Qu.: 0.00   1st Qu.:   0.0   1st Qu.:0.185   1st Qu.:3.79
 Median :436   Median :  92   Median : 1.00   Median :  26.0   Median :0.296   Median :4.52
 Mean   :418   Mean   : 179   Mean   : 3.55   Mean   :  63.9   Mean   :0.392   Mean   :4.42
 3rd Qu.:598   3rd Qu.: 146   3rd Qu.: 4.00   3rd Qu.:  81.0   3rd Qu.:0.535   3rd Qu.:4.98
 Max.   :900   Max.   :3249   Max.   :52.00   Max.   :1283.0   Max.   :1.659   Max.   :8.09
                                                               NA's   :12

> cor(IR,use="complete.obs")
            ccode   Ndyads disputes  allies openness exposure
ccode     1.00000 -0.29623  -0.1399 -0.3983  0.02744  -0.6544
Ndyads   -0.29623  1.00000   0.8626  0.9200 -0.07511   0.6988
disputes -0.13989  0.86257   1.0000  0.8255 -0.16819   0.6335
allies   -0.39826  0.92004   0.8255  1.0000 -0.12548   0.7003
openness  0.02744 -0.07511  -0.1682 -0.1255  1.00000  -0.1433
exposure -0.65442  0.69878   0.6335  0.7003 -0.14325   1.0000
```

Ignoring Exposure

```
> IR.fit1<-glm(disputes~allies+openness,data=IR,family="poisson")
> summary(IR.fit1)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  1.1559498  0.1117581  10.343  < 2e-16 ***
allies       0.0025184  0.0001159  21.734  < 2e-16 ***
openness    -1.1144132  0.2773631  -4.018 5.87e-05 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 732.68  on 101  degrees of freedom
Residual deviance: 392.97  on  99  degrees of freedom
  (12 observations deleted due to missingness)
AIC: 588.29

Number of Fisher Scoring iterations: 6
```

```
> IR.fit2<-glm(disputes~allies+openness,data=IR,family="poisson",
  offset=log(Ndyads))
> summary(IR.fit2)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.2906055  0.1194616 -27.545  < 2e-16 ***
allies      -0.0006058  0.0001333  -4.544 5.52e-06 ***
openness    -1.6040587  0.3167415  -5.064 4.10e-07 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

    Null deviance: 320.19  on 101  degrees of freedom
Residual deviance: 277.79  on  99  degrees of freedom
  (12 observations deleted due to missingness)
AIC: 473.11

Number of Fisher Scoring iterations: 5
```

Correcting for Exposure (continued)

```
> IR.fit3<-glm(disputes~allies+openness+log(Ndyads),data=IR,
+              family="poisson")
> summary(IR.fit3)

Call:
glm(formula = disputes ~ allies + openness + log(Ndyads), family = "poisson",
    data = IR)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
-2.838  -1.390  -0.758   0.605   4.731

Coefficients:
               Estimate  Std. Error z value             Pr(>|z|)
(Intercept) -2.42656676  0.34345252   -7.07      0.0000000000016 ***
allies      -0.00000948  0.00025687   -0.04                 0.97
openness    -1.44462460  0.31193821   -4.63      0.0000036368547 ***
log(Ndyads)  0.81097748  0.07095243   11.43 < 0.0000000000000002 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 732.68  on 101  degrees of freedom
Residual deviance: 270.59  on  98  degrees of freedom
  (12 observations deleted due to missingness)
AIC: 467.9

Number of Fisher Scoring iterations: 5
```

Test $\beta_{\texttt{exposure}} = 1.0$

```
> # z-test:

> 2*pnorm((0.811-1)/.071)
[1] 0.007768438
```

```
> # Wald test:

> wald.test(b=coef(IR.fit3),Sigma=vcov(IR.fit3),Terms=4,H0=1)

Wald test:
----------
Chi-squared test:
X2 = 7.1, df = 1, P(> X2) = 0.0077
```

Contagion, Heterogeneity, and Dispersion


Lisco
26
Oshkosh
27
Lewellen
26


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

Cats (daily values):

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

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

Daily values:

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

Positive contagion $\rightarrow$ *overdispersion*.

Daily values:

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

Negative contagion $\rightarrow$ *underdispersion*.

Aggregated two-day measures:

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

- 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 $H_0 : \hat{\delta} = 0$.

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

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

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

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

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

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

Negative Binomial (log-)Likelihood

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

So:

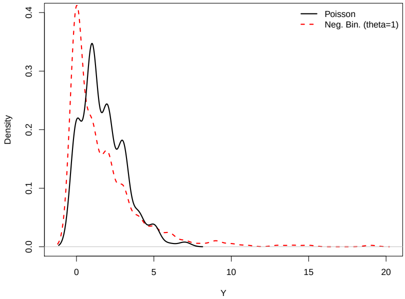
- $\alpha = 0 \iff \mathrm{E}(Y) = \mathrm{Var}(Y)$
- LR test for overdispersion:

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

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

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
```

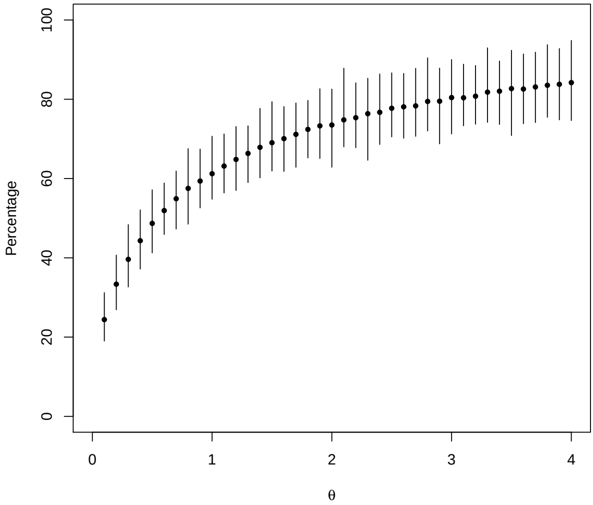
What Difference Does It Make (cont'd)?

```
> # 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)

"Continuous parameter binomial":

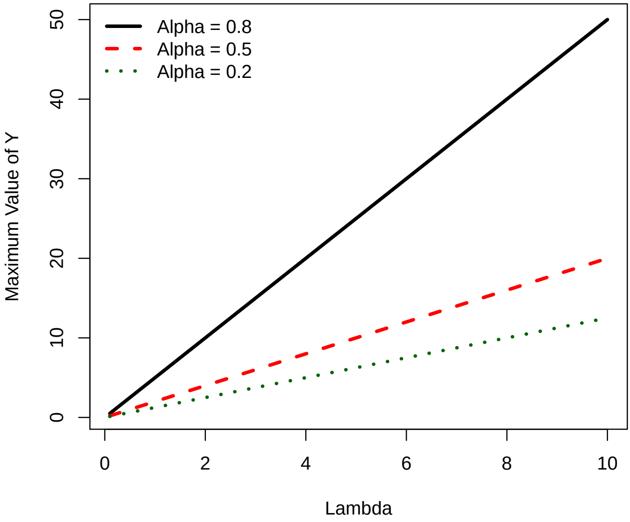
$$\Pr(Y_i = y|\lambda_i, \alpha) = \frac{\frac{\Gamma\left(\frac{-\lambda_i}{\alpha-1}+1\right)}{Y_i!\Gamma\left(\frac{-\lambda_i}{\alpha-1}-Y_i+1\right)}(1-\alpha)^{Y_i}(\alpha)^{\frac{-\lambda_i}{\alpha-1}-Y_i}}{D_i}$$

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

CPB:

- ...also has $\text{E}(Y_i) = \lambda_i$ [with $\mu_i = \exp(\mathbf{X}_i\boldsymbol{\beta})$]
- ...has $\text{Var}(Y) = \lambda_i\alpha$ 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

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

Example: SCOTUS Amicus Curiae (1953-85)

- $N = 7157$
- `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.

```
> summary(amici)
     namici              term           civlibs
 Min.   : 0.00   Min.   :53.0   Min.   :0.000
 1st Qu.: 0.00   1st Qu.:64.0   1st Qu.:0.000
 Median : 0.00   Median :72.0   Median :1.000
 Mean   : 1.03   Mean   :71.1   Mean   :0.501
 3rd Qu.: 1.00   3rd Qu.:79.0   3rd Qu.:1.000
 Max.   :53.00   Max.   :85.0   Max.   :1.000
```

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"

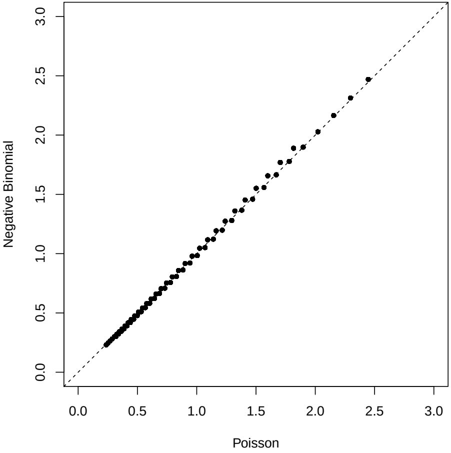
```
> Phats<-fitted.values(amici.poisson)
> Uhats<-((amici$namici-Phats)^2 - amici$namici) / (Phats * sqrt(2))
> summary(lm(Uhats~Phats))

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



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

Predicted Values: Poisson and NB

```
> plot(amici.poisson$fitted.values,amici.NB$fitted.values,xlab="Poisson",
  ylab="Negative Binomial",main="Predicted Counts")
> abline(a=0,b=1,lwd=2)
```



- Models where Over- / Underdispersion $= f(\mathbf{Z}_i\gamma)$
- Models for Censored / Truncated Counts
- "Zero-Inflated" and "Hurdle" Count Models
- Time Series of Counts
- Panel Data Models for Event Counts
- Spatial Count Models
- etc...