PLSC 503 – Spring 2024
Models For Event Counts

April 22, 2024

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-valued
- Non-negative
- "Cumulative"

Events:

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

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

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

Count of events:

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

Three assumptions:

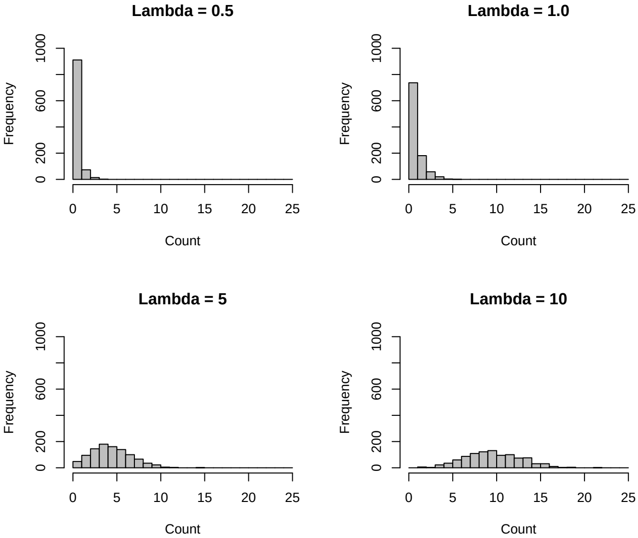
- No Simultaneous Events
- Constant Arrival Rate
- Independent Event Arrivals

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

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

A Poisson variate:

- 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.
- $\lambda \to \infty \iff Y \sim N$


Lambda = 0.5
Lambda = 1.0
Lambda = 5
Lambda = 10
Frequency
Count
0
200
600
1000
0
5
10
15
20
25

Suppose

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

then

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

Poisson Likelihood

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

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

Example: Federal Judicial Review

Dahl (1957), on SCOTUS overturning Acts of Congress:

- SCOTUS gets "out of step" with the other branches $\rightarrow$ judicial review
- Older / longer-serving justices will more likely to invalidate legislation
- BUT Court may fear retribution from other branches

Example: Federal Judicial Review, 1789-2021

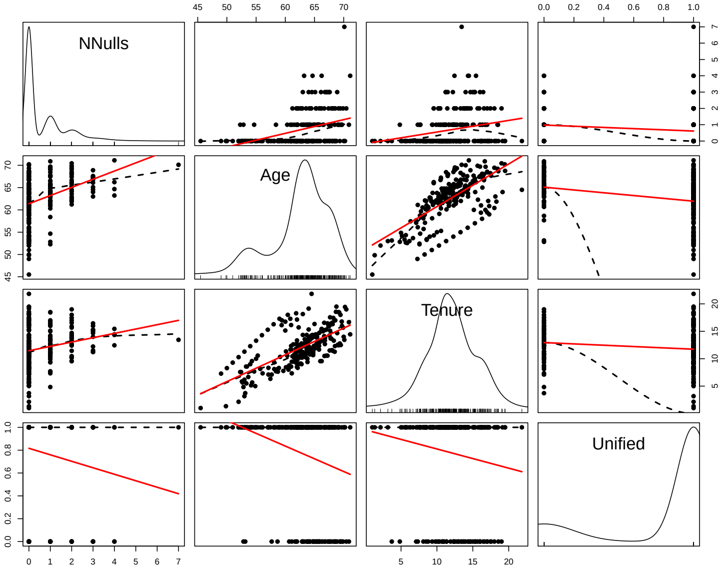
Data:

- Y_i = number of Acts of Congress overturned by the Supreme Court in each year (`NNulls`)

- Predictors:
  - The *mean age* (`Age`) of the Supreme Court's justices ($\bar{X} = 62.6, \sigma = 5, \text{E}(\hat{\beta}) > 0$)
  - The *mean tenure* (`Tenure`) of the Supreme Court's justices ($\bar{X} = 12.0, \sigma = 3.5, \text{E}(\hat{\beta}) > 0$)
  - Whether (1) or not (0) there was *unified government* (`Unified`) ($\bar{X} = 0.78, \text{E}(\hat{\beta}) < 0$)

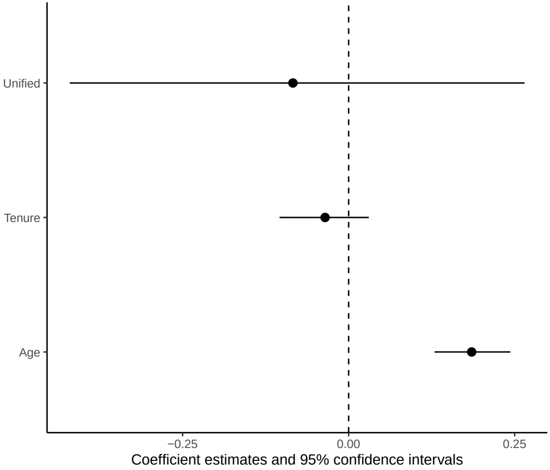
```
> describe(NewDahl)
               vars   n    mean    sd median trimmed   mad    min    max range  skew kurtosis   se
Year              1 233 1905.00 67.41 1905.0 1905.00 85.99 1789.0 2021.0 232.0  0.00    -1.22 4.42
NConstDecisions   2 233   17.96 19.11   12.0   14.63 14.83    0.0   85.0  85.0  1.38     1.48 1.25
NNulls            3 233    0.70  1.06    0.0    0.49  0.00    0.0    7.0   7.0  1.96     5.39 0.07
Age               4 233   62.65  4.96   63.6   63.11  3.95   45.5   71.1  25.6 -0.84     0.29 0.32
Tenure            5 233   12.00  3.54   11.9   12.06  3.15    1.0   21.8  20.8 -0.19     0.25 0.23
Unified           6 233    0.78  0.42    1.0    0.84  0.00    0.0    1.0   1.0 -1.32    -0.26 0.03
```

Federal Judicial Review, 1789-2021



```
> nulls.poisson<-glm(NNulls~Age+Tenure+Unified,family="poisson",
+                    data=NewDahl)
> summary(nulls.poisson)

Call:
glm(formula = NNulls ~ Age + Tenure + Unified, family = "poisson",
    data = NewDahl)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -11.7638     1.6993   -6.92  4.4e-12 ***
Age           0.1852     0.0291    6.36  2.0e-10 ***
Tenure       -0.0354     0.0343   -1.03     0.30
Unified      -0.0839     0.1743   -0.48     0.63
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 339.50  on 232  degrees of freedom
Residual deviance: 266.69  on 229  degrees of freedom
AIC: 497.7

Number of Fisher Scoring iterations: 5
```

Coefficient Plot (using `modelplot`)



Interpretation: Incidence Rate Ratios

IRRs:

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

- Like ORs
- `Age`: IRR=exp(0.19) = 1.21

For a δ-unit change in X_k:

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

So, a for ten-year difference in `Age`:

$$\begin{aligned} \mathrm{IRR} &= \exp(10 \times 0.190) \\ &= \exp(1.90) \\ &= 6.69 \end{aligned}$$

Via `mfx`:

```
> library(mfx)
> nulls.poisson.IRR<-poissonirr(NNulls~Age+Tenure+Unified,
+                                 data=NewDahl)
> nulls.poisson.IRR
Call:
poissonirr(formula = NNulls ~ Age + Tenure + Unified, data = NewDahl)

Incidence-Rate Ratio:
           IRR Std. Err.     z P>|z|
Age     1.2035    0.0350  6.36 2e-10 ***
Tenure  0.9652    0.0331 -1.03  0.30
Unified 0.9195    0.1603 -0.48  0.63
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Predicted Values ($\hat{Y}$s)

Mean predicted Y:

$$\mathrm{E}(Y|\bar{\mathbf{X}}_i) = \exp[\bar{\mathbf{X}}_i\hat{\beta}]$$

In-Sample:

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

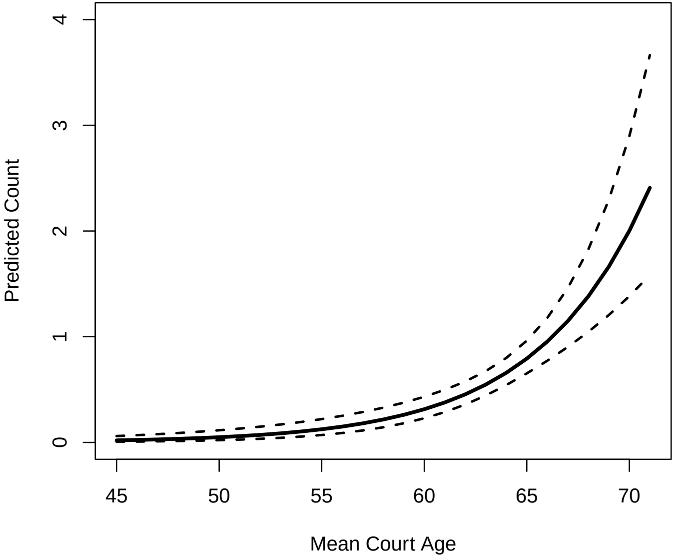
Out-of-Sample: use `predict`

Example: Out-Of-Sample Predicted Values

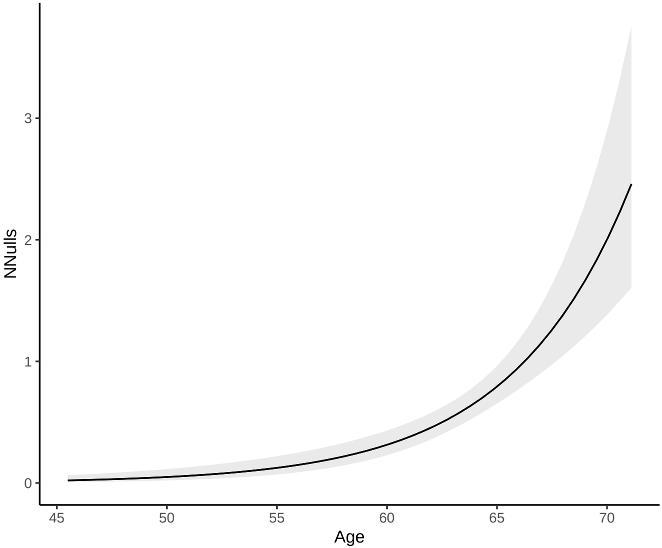
"By-hand" example:

```
> simdata<-data.frame(Age=seq(from=45,to=71,by=1),
+                     Tenure=mean(NewDahl$Tenure,na.rm=TRUE),
+                     Unified=1)
> 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$Age,nullhats$Yhat,t="l",lwd=3,ylim=c(0,5),ylab=
+        "Predicted Count", xlab="Mean Court Age")
> lines(simdata$Age,nullhats$UB,lwd=2,lty=2)
> lines(simdata$Age,nullhats$LB,lwd=2,lty=2)
```

Plotting Out-Of-Sample Predicted Values

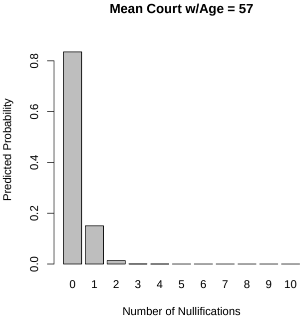


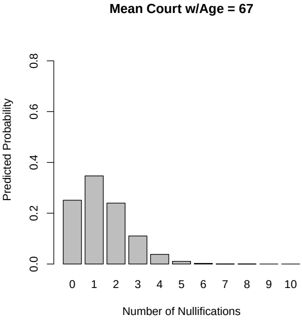
Same, Using `plot_predictions`



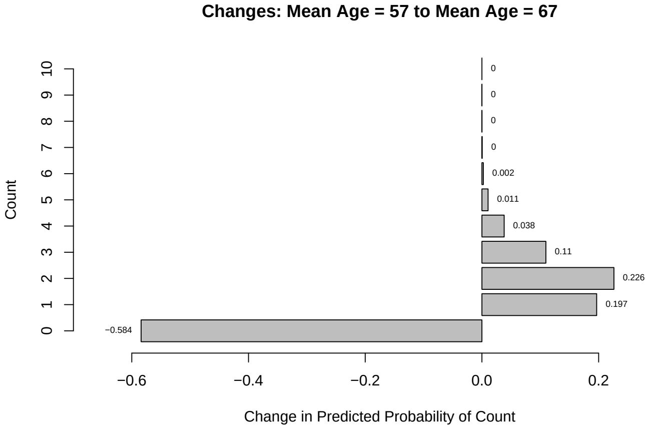
The predicted probability that $Y_i = y$ is:

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





Changes in Predicted Probabilities



"Exposure" and "Offsets"

If we relax the assumption of equal "exposure," we get:

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

i.e., the expected number of events is proportional to *exposure* M_i.

Note that now, instead of:

$$\ln[E(Y_i)] = \mathbf{X}_i\boldsymbol{\beta}$$

we have:

$$\ln\left[E\left(\frac{Y_i}{M_i}\right)\right] = \mathbf{X}_i\boldsymbol{\beta}$$

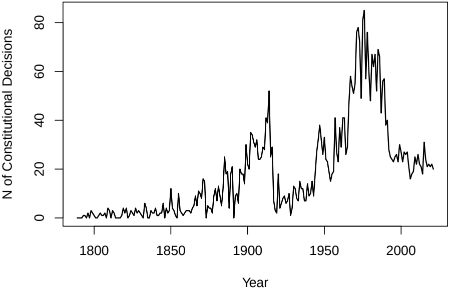
which is a *rate*, and the same as:

$$\ln[E(Y_i)] = \ln(M_i) + \mathbf{X}_i\boldsymbol{\beta}$$

that is, including $\ln(M_i)$ in $\mathbf{X}$ with $\beta_{\ln(M)} = 1$.

For the judicial review (1789-2021) data:

- SCOTUS (typically) reviews many constitutional cases per year
- The number of such cases is the *possible* number of nullifications



Adding an "offset":

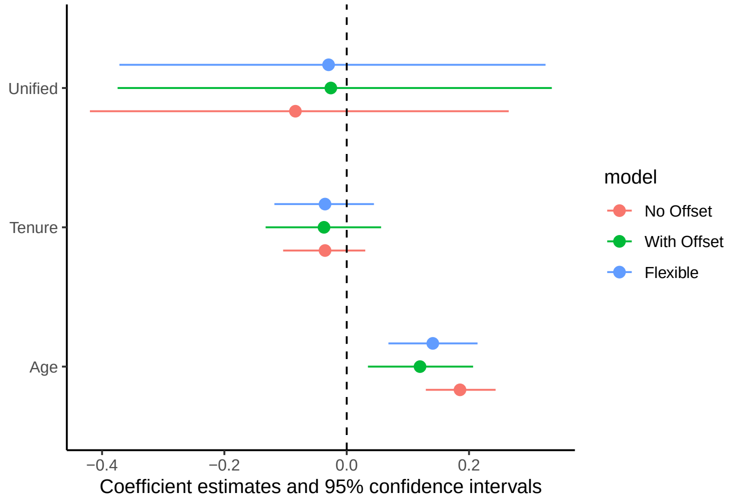
```
> nulls.poisson2<-glm(NNulls~Age+Tenure+Unified,family="poisson",
+                     offset=log(NConstDecisions+1),data=NewDahl)
> summary(nulls.poisson2)

Call:
glm(formula = NNulls ~ Age + Tenure + Unified, family = "poisson",
    data = NewDahl, offset = log(NConstDecisions + 1))

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -10.6120     2.4552   -4.32 0.000015 ***
Age           0.1199     0.0440    2.72   0.0065 **
Tenure       -0.0371     0.0483   -0.77   0.4420
Unified      -0.0259     0.1808   -0.14   0.8859
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 257.47  on 232  degrees of freedom
Residual deviance: 242.37  on 229  degrees of freedom
AIC: 473.3

Number of Fisher Scoring iterations: 6
```

Correcting for Exposure (continued)

Including the "offset" as a control:

```
> nulls.poisson3<-glm(NNulls~Age+Tenure+Unified+log(NConstDecisions+1),
+                     family="poisson",data=NewDahl)
> summary(nulls.poisson3)

Coefficients:
                        Estimate Std. Error z value      Pr(>|z|)
(Intercept)             -10.5838     2.0821   -5.08 0.00000037124 ***
Age                       0.1408     0.0371    3.79       0.00015 ***
Tenure                   -0.0354     0.0415   -0.85       0.39295
Unified                  -0.0296     0.1774   -0.17       0.86739
log(NConstDecisions + 1)   0.5744     0.0924    6.22 0.00000000051 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 339.50  on 232  degrees of freedom
Residual deviance: 223.14  on 228  degrees of freedom
AIC: 456.1

Number of Fisher Scoring iterations: 5


> # Wald test for b = 1.0:
>
> wald.test(b=coef(nulls.poisson3),Sigma=vcov(nulls.poisson3),Terms=4,H0=1)
Wald test:
----------

Chi-squared test:
X2 = 33.7, df = 1, P(> X2) = 0.0000000064
```

Model Comparisons



Contagion, Heterogeneity, and Dispersion


Lisco
26
Oshkosh
27
Lewellen
26


Seattle
WASHINGTON
MONTANA
NORTH DAKOTA
MINNESOTA
Ottawa
NB
MAINE
VT
NH
OREGON
IDAHO
WYOMING
SOUTH DAKOTA
WISCONSIN
MICHIGAN
Toronto
NEW YORK
MA
Boston
CT
RI
NEBRASKA
IOWA
Chicago
ILLINOIS
OHIO
PENN
New York
NEVADA
UTAH
United States
COLORADO
KANSAS
MISSOURI
INDIANA
MD
DE
NJ
WEST VIRGINIA
VIRGINIA
San Francisco
CALIFORNIA
Las Vegas
KENTUCKY
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):

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

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

Assumes:

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

Antelope

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:

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

Homogeneity implies:

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

E.g., omitted variables:

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

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

Heterogeneity $\rightarrow$ overdispersion:

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

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

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

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

Log-likelihood is:

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

So:

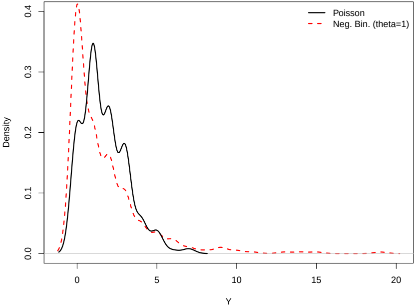
- $\alpha = 0 \iff \mathrm{E}(Y) = \mathrm{Var}(Y)$
- LR test for overdispersion:

$$-2 \times (\ln \widehat{L_{Poisson}} - \widehat{\ln L_{NB}}) \sim \chi_1^2$$

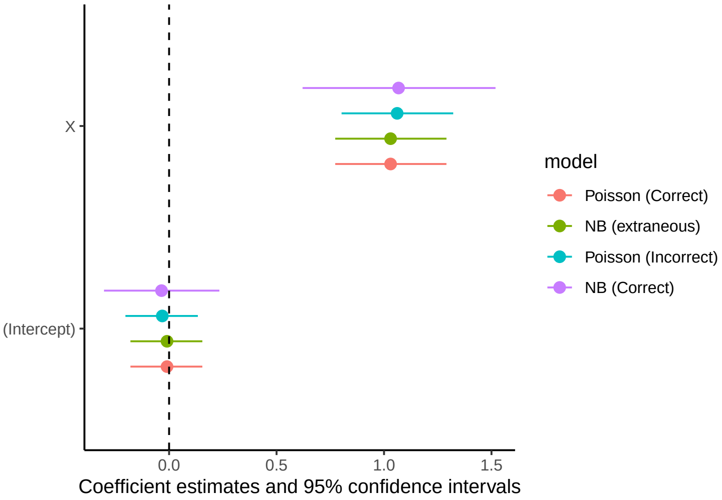
- $\widehat{E(Y_i)} \equiv \hat{\lambda}_i = \exp(\mathbf{X}_i \hat{\boldsymbol{\beta}})$

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)?

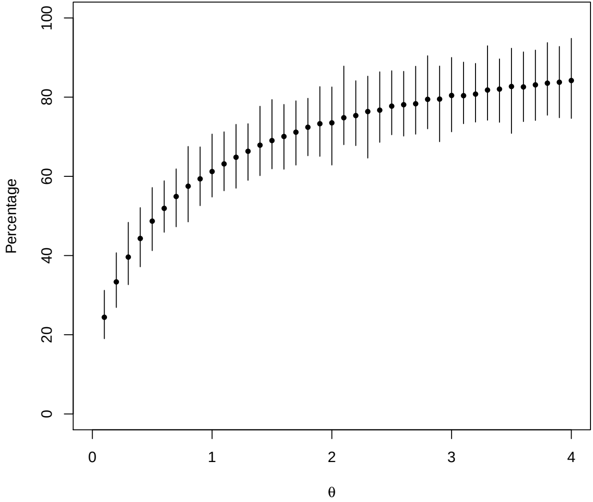


Poisson Regression Underestimates N.B. Variances

Simulation:

```
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)
- Also, `modelsummary` + `marginaleffects`

"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...

Yields a log-likelihood:

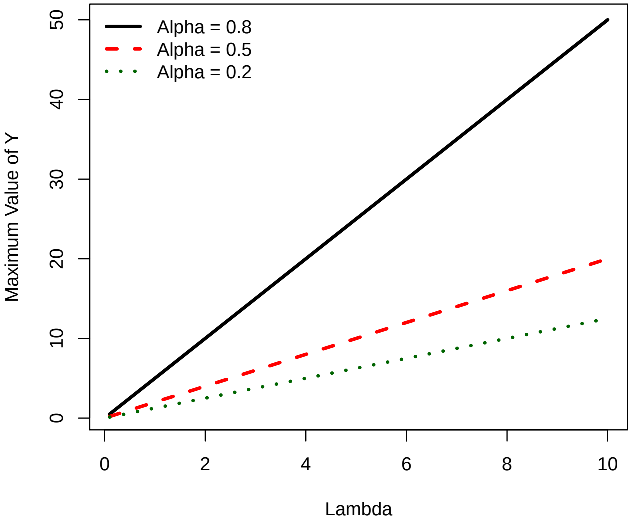
$$\begin{aligned}\ln L_{CPB} \quad = \quad & \sum_{i=1}^{N}\Big\{\ln\Gamma\left(\frac{-\lambda_i}{\alpha-1}+1\right) - \ln\Gamma\left(\frac{-\lambda_i}{\alpha-1} - Y_i + 1\right) \\ & + Y_i\ln(1-\alpha) + \left(\frac{-\lambda_i}{\alpha-1} - Y_i\right)\ln(\alpha) - \ln(D_i)\Big\}\end{aligned}$$

Are You Down With The CPB?

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
0
10
20
30
40
50
0
2
4
6
8
10
Lambda

Example Redux: Judicial Review

Recall:

```
> nulls.poisson<-glm(NNulls~Age+Tenure+Unified,family="poisson",
+                    data=NewDahl)
> summary(nulls.poisson)

Call:
glm(formula = NNulls ~ Age + Tenure + Unified, family = "poisson",
    data = NewDahl)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -11.7638     1.6993   -6.92  4.4e-12 ***
Age           0.1852     0.0291    6.36  2.0e-10 ***
Tenure       -0.0354     0.0343   -1.03     0.30
Unified      -0.0839     0.1743   -0.48     0.63
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 339.50  on 232  degrees of freedom
Residual deviance: 266.69  on 229  degrees of freedom
AIC: 497.7

Number of Fisher Scoring iterations: 5
```

Overdispersion Test: "By Hand"

Test:

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

Call:
lm(formula = Uhats ~ Phats)

Residuals:
   Min     1Q Median     3Q    Max
-1.421 -0.762  0.080  0.272  8.509

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.0532     0.1318    0.40     0.69
Phats         0.1828     0.1573    1.16     0.25

Residual standard error: 1.1 on 231 degrees of freedom
Multiple R-squared:  0.00581, Adjusted R-squared:  0.00151
F-statistic: 1.35 on 1 and 231 DF,  p-value: 0.246
```

$\rightarrow$ no particular evidence of overdispersion here. However...

Negative Binomial Regression

```
> library(MASS)
> nulls.NB<-glm.nb(NNulls~Age+Tenure+Unified,data=NewDahl)
> summary(nulls.NB)

Coefficients:
            Estimate Std. Error z value      Pr(>|z|)
(Intercept) -12.1775     1.9481   -6.25 0.00000000041 ***
Age           0.1930     0.0336    5.74 0.00000000932 ***
Tenure       -0.0413     0.0398   -1.04          0.30
Unified      -0.0956     0.2065   -0.46          0.64
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(2.5) family taken to be 1)

    Null deviance: 269.11  on 232  degrees of freedom
Residual deviance: 210.00  on 229  degrees of freedom
AIC: 492

Number of Fisher Scoring iterations: 1

              Theta:  2.53
          Std. Err.:  1.21

 2 x log-likelihood:  -482.00

> # alpha:
>
> 1 / nulls.NB$theta
[1] 0.4
```

Alternative NB Regression

```
> library(msme)
> nulls.nb2<-nbinomial(NNulls~Age+Tenure+Unified,data=NewDahl)
> summary(nulls.nb2)

Call:
ml_glm2(formula1 = formula1, formula2 = formula2, data = data,
    family = family, mean.link = mean.link, scale.link = scale.link,
    offset = offset, start = start, verbose = verbose)

Deviance Residuals:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  -1.70   -0.99   -0.51   -0.29    0.34    2.33

Pearson Residuals:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   -1.0    -0.7    -0.4     0.0     0.4     3.9

Coefficients (all in linear predictor):
              Estimate     SE      Z        p      LCL     UCL
(Intercept)   -12.0194 1.9867 -6.050 1.45e-09 -15.9135 -8.1254
Age             0.1901 0.0346  5.496 3.89e-08   0.1223  0.2579
Tenure         -0.0392 0.0418 -0.938    0.348  -0.1211  0.0427
Unified        -0.0975 0.2078 -0.469    0.639  -0.5048  0.3098
(Intercept)_s   0.3944 0.1903  2.072   0.0383   0.0213  0.7674

Null deviance: 269  on  231 d.f.
Residual deviance: 210  on  228 d.f.
Null Pearson: 295  on  231 d.f.
Residual Pearson: 220  on  228 d.f.
Dispersion: 0.97
AIC:  492

Number of optimizer iterations:  64
```

See here for details...

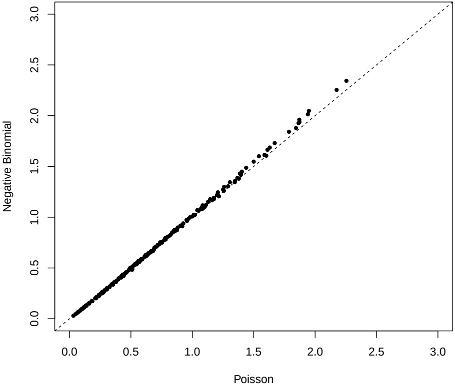
Estimated slopes and standard errors:

```
> # Coefficient estimates:
>
> cbind(nulls.poisson$coefficients,coef(nulls.NB))
               [,1]    [,2]
(Intercept) -11.764 -12.177
Age           0.185   0.193
Tenure       -0.035  -0.041
Unified      -0.084  -0.096
```

```
> # Estimated standard errors:
>
> cbind(diag(sqrt(vcov(nulls.poisson))),diag(sqrt(vcov(nulls.NB))))
             [,1]  [,2]
(Intercept) 1.699 1.948
Age         0.029 0.034
Tenure      0.034 0.040
Unified     0.174 0.206
```

Predicted Values: Poisson and NB

```
> plot(nulls.poisson$fitted.values,nulls.NB$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)
```



More Things

Topics we'll (possibly) cover in PLSC 504:

- 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...

Now That You're Done...

...read the following things:

- Breiman, Leo. 2001. **"Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author)."** *Statistical Science* 16(3):199-231.
- Daoud, Adel, and Devdatt Dubhashi. 2023. **"Statistical Modeling: The Three Cultures."** *Harvard Data Science Review* Issue 5.1 (Winter 2023).
- Molnar, Christopher. 2023. **Modeling Mindsets: The Many Cultures of Learning From Data**. MUCBOOK:Heidi Seibold.

Thank you!