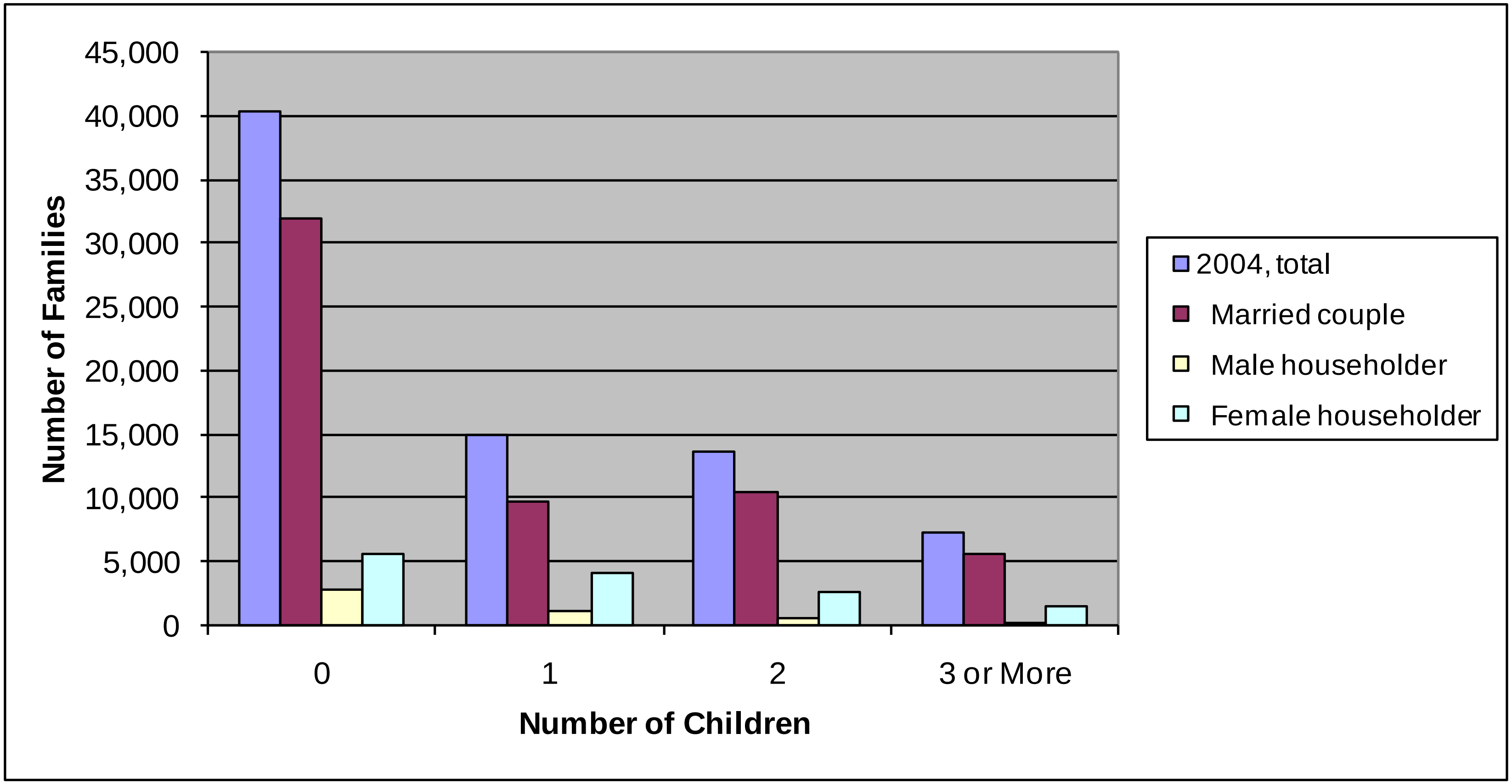
Poisson and Related Models

Categorical and Limited Dependent Variables

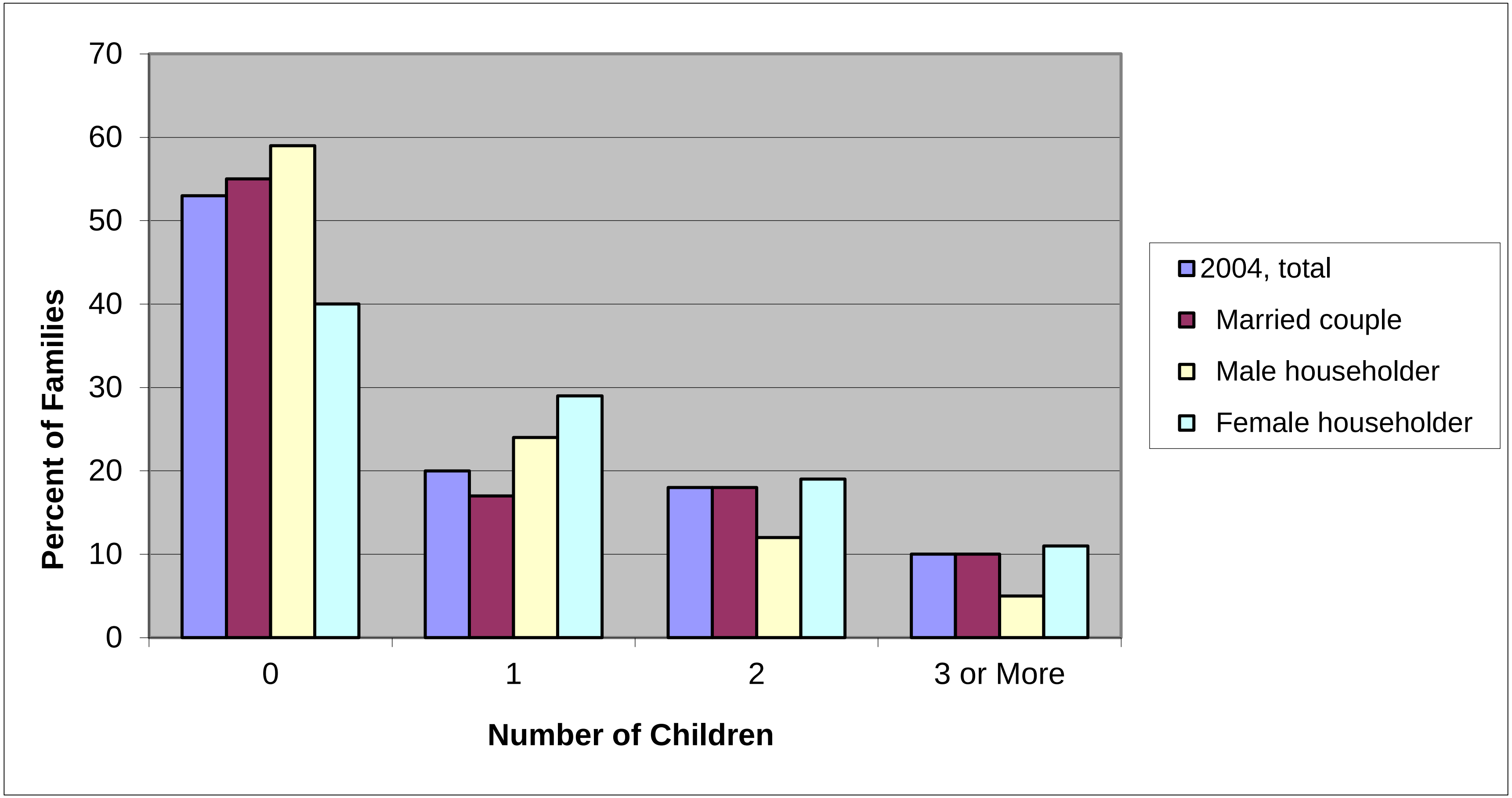
Paul A. Jargowsky

Number of Families by Number of Children, by Family Type, 2004



Source: 2006 U.S. Statistical Abstract, Table 62

The Probability of Each Outcome Varies by Family Type



Mean number of Children per Family: the Rate of Child Production Varies by Family Type

Families (1,000) by Number of Children, 2004

		Number of Children:				Mean Number
	Total	0	1	2	3+	Of Children*
Total	76,217	40,273	14,964	13,696	7,283	0.89
Married couple	57,719	31,926	9,763	10,481	5,548	0.87
Male Headed	4,716	2,786	1,146	550	235	0.65
Female Headed	13,781	5,560	4,055	2,665	1,501	1.06

*assuming an average of 3.5 for families in "3+" category.

$$\mu = \sum_{i=1}^{n} \Pr(c = i)(i) = \Pr(0)(0) + \Pr(1)(1) + \Pr(2)(2) + \Pr(3+)(3.5)$$

OLS on Counts?

- Not a continuous DV
- Not really censored or truncated or ordinal, which assume some continuous disturbance term
- Coefficients from OLS on counts are inefficient, inconsistent, and biased
- Need a different approach

Poisson Distribution for Counts

When an event occurs and reoccurs *randomly* and *independently*, the total number of occurrences of an event (Y) over *a fixed period of time* will have a Poisson Distribution, given by:

$$\Pr\left(Y=k\right)=\frac{e^{-\mu}\mu^{k}}{k!},\quad k=0,1,2,....$$

Note: this is a *discrete* probability distribution.

$$\mu=E\left[Y_i\right]$$

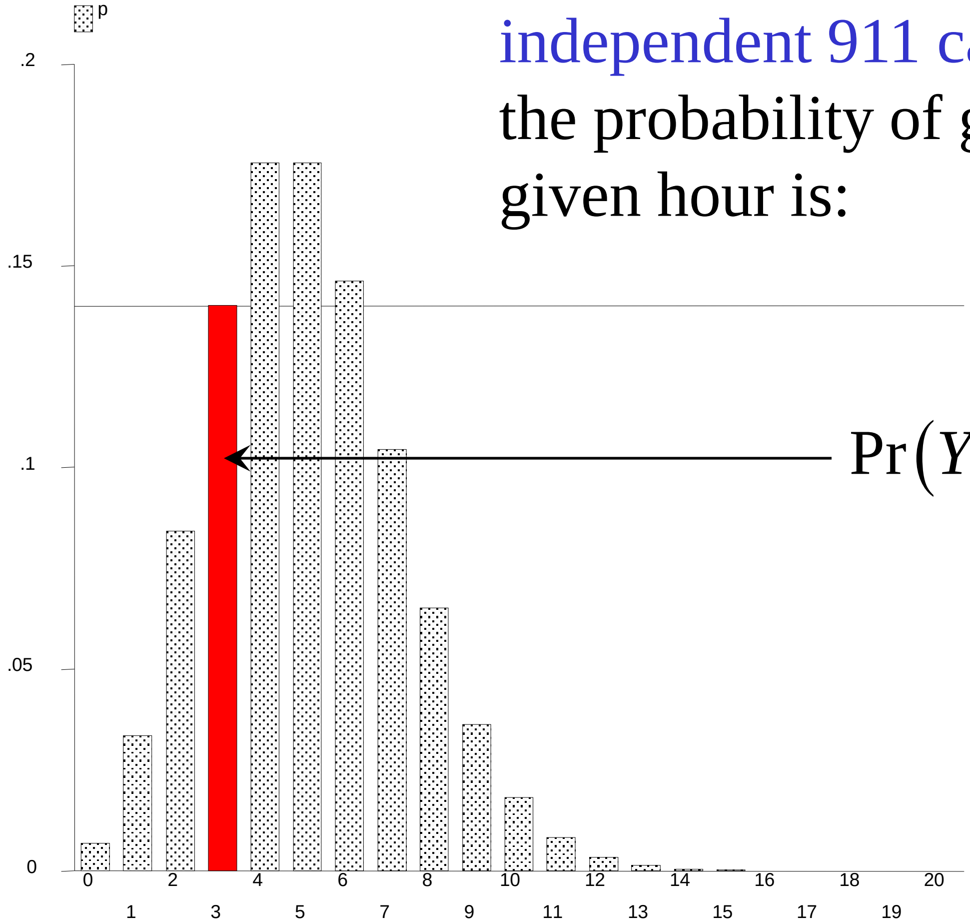
The expected arrival rate, or "incidence rate"

$$\sigma^2=E\left[\left(Y-\mu\right)^2\right]=\mu$$

Equidispersion, i.e. mean = variance

Poisson PDF Application

For example, if the mean number of independent 911 calls per hour is 5, the probability of getting 3 calls in a given hour is:

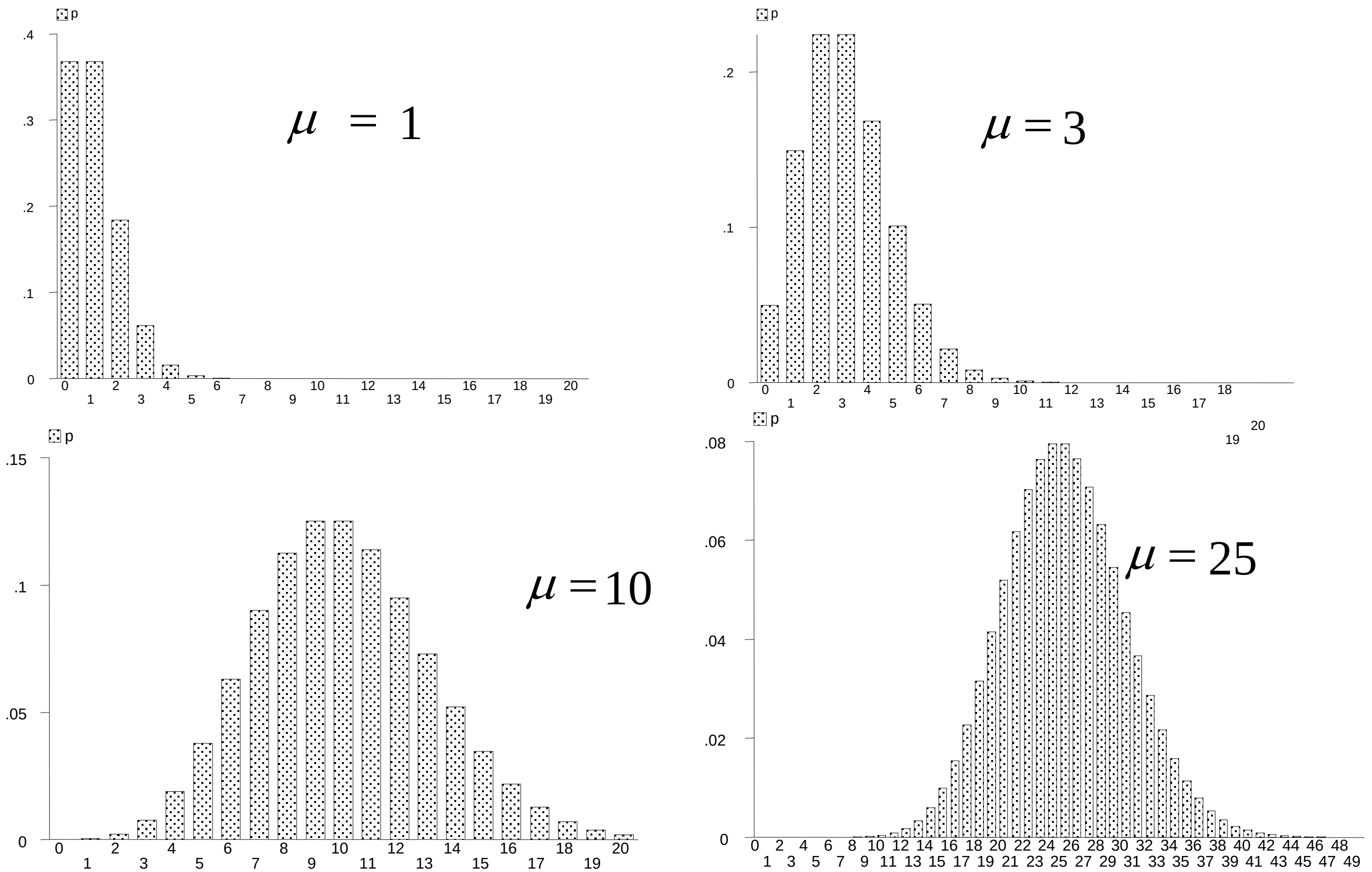


$$\Pr(Y = 3) = \frac{e^{-5}5^{3}}{3!}$$

$$= \frac{(0.00674)(125)}{6}$$

$$= 0.14$$

Poisson Distributions with Different Means



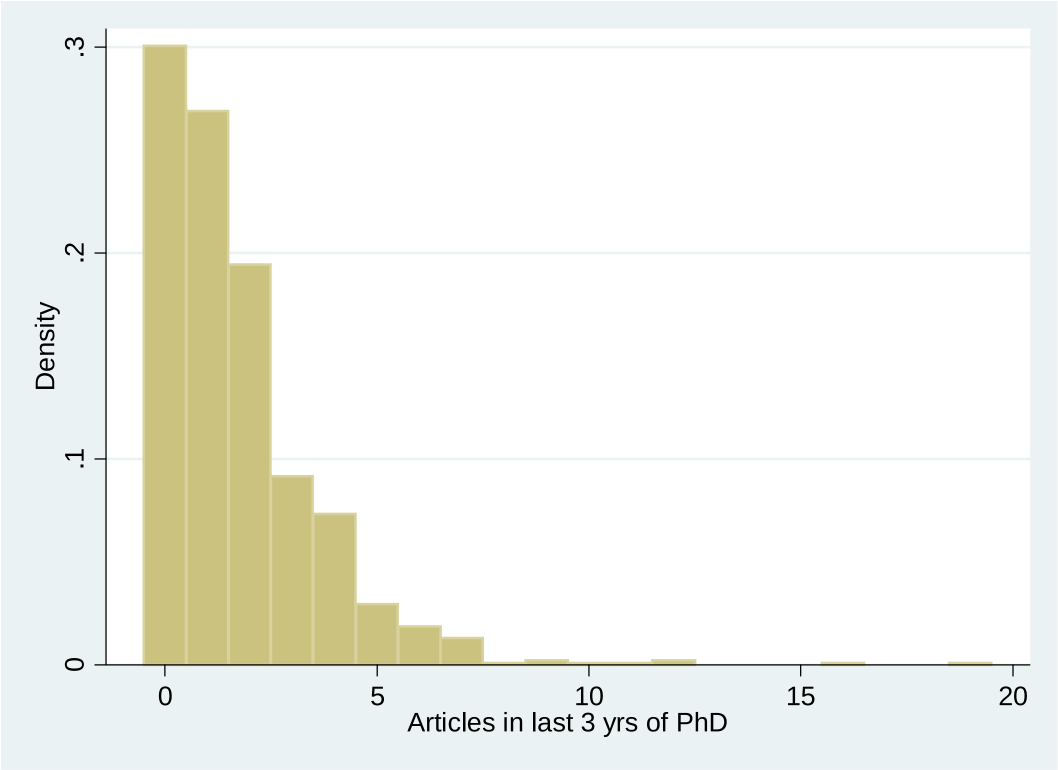
The Poisson converges to normal. If the mean is more than 10, OLS works fine.

Count of Articles Published (couart2.dta)

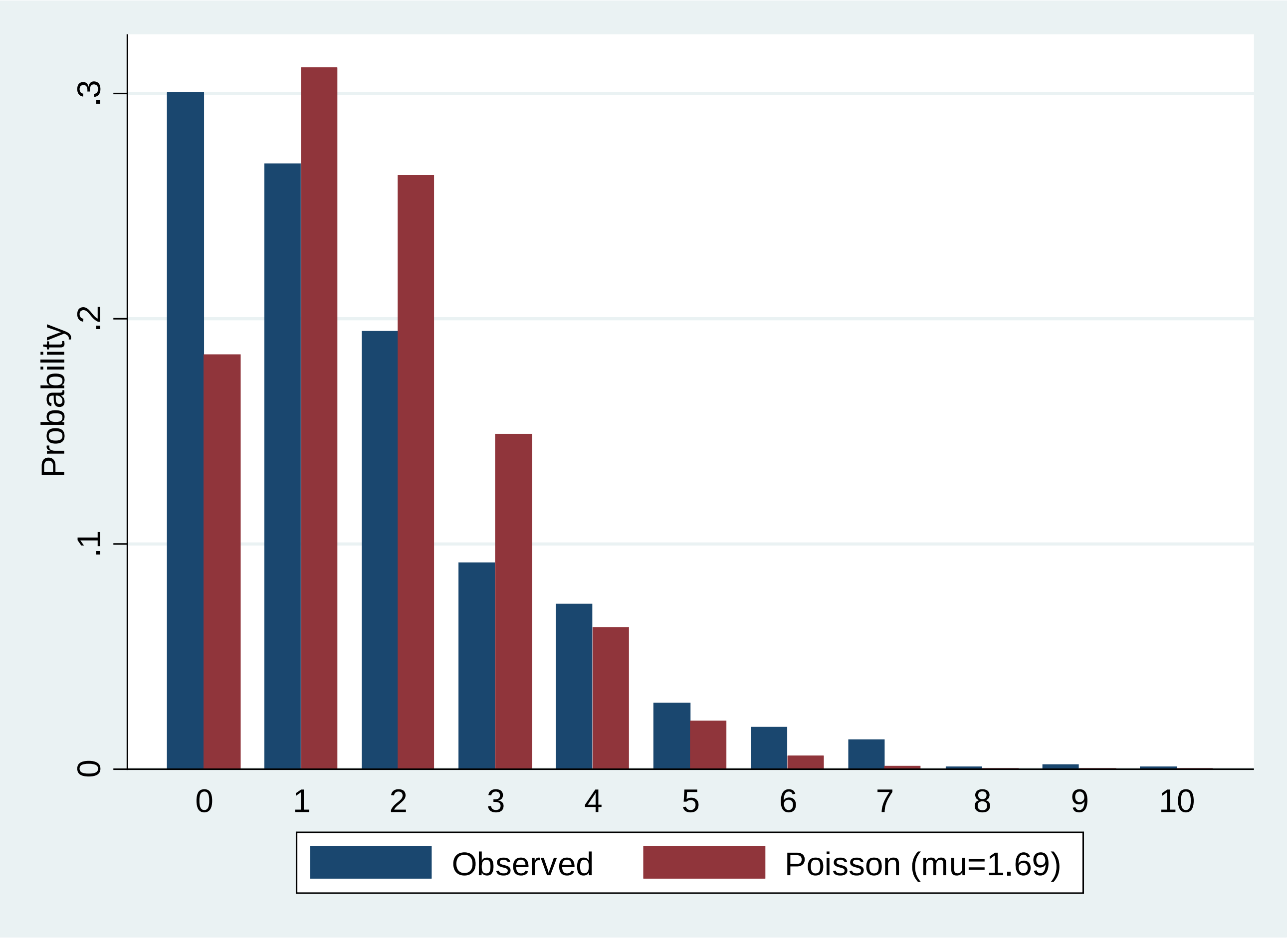
```
. tab art
```

Articles in last 3 yrs of PhD	Freq.	Percent
0	275	30.05
1	246	26.89
2	178	19.45
3	84	9.18
4	67	7.32
5	27	2.95
6	17	1.86
7	12	1.31
8	1	0.11
9	2	0.22
10	1	0.11
11	1	0.11
12	2	0.22
16	1	0.11
19	1	0.11
Total	915	100.00

```
. histogram art, discrete
```



Is it Poisson?



Not exactly. There are both more zeros and more people with 4+ articles than Poisson would predict, i.e. greater variance in articles.

Overdispersion

Variable	Obs	Mean	Std. Dev.	Min	Max
art	915	**1.692896**	**1.926069**	0	19
ment	915	8.767212	9.483915	0	76.99998

$$Y_i = \{0,1,2,3,...,19\} \quad \mu = 1.69 \quad \sigma = 1.93 \quad \sigma^2 = 3.73 \quad \sigma^2 > \mu$$

Overdispersion: the variance exceeds the mean. Perhaps the underlying mean is not the same for each person, and varies as a function of the person's characteristics, e.g. mentor publications.

Why overdispersion? Perhaps not all observations have the same incidence rate. It depends on the independent variables (**X**). If control for relevant **X**, maybe they have the same *conditional* mean. How do Xs affect the mean?

Some Publish More than Others

$\mu_i = \mathbf{x}_i\boldsymbol{\beta}$ is biased, inconsistent, and inefficient.

$\mu_i = f\left(\mathbf{x}_i\right)$ Needs to be positive. Linear does not work.

$\mu_i = e^{\beta_1 + \beta_2 X_{2i} + \ldots + \beta_k X_{ki}} = e^{\mathbf{x}_i\boldsymbol{\beta}}$ works. $0 < e^{\mathbf{x}_i\boldsymbol{\beta}} < \infty$

$$\Pr\left(Y_i \mid \mathbf{x}_i\right) = \frac{e^{-\mu_i}\mu_i^{Y_i}}{Y_i!}, \quad \text{where } \mu_\mathrm{i} = e^{\mathbf{x}_i\boldsymbol{\beta}} \text{ and}$$

$$\mathbf{x}_i\boldsymbol{\beta} = \beta_1 + \beta_2 X_{2i} + \ldots + \beta_k X_{ki}$$

Note that $\mu_\mathrm{i} = e^{\mathbf{x}_i\boldsymbol{\beta}}$ implies $\ln\left(\mu_i\right) = \mathbf{x}_i\boldsymbol{\beta}$, except $\ln\left(0\right)$ is undefined.

Poisson Likelihood Function

Poisson Regression: $\mu_i = E(Y_i \mid \mathbf{x}_i) = e^{\mathbf{x}_i \boldsymbol{\beta}}$

$$\mathcal{L} = \prod_{i=1}^{n} \Pr\left(Y_i \mid \mathbf{x}_i\right) = \prod_{i=1}^{n} \left[\frac{e^{-\mu_i} \mu_i^{Y_i}}{Y_i\,!} \right] = \prod_{i=1}^{n} \left[\frac{e^{-\left(e^{\mathbf{x}_i \boldsymbol{\beta}}\right)} \left(e^{\mathbf{x}_i \boldsymbol{\beta}}\right)^{Y_i}}{Y_i\,!} \right]$$

Maximize $\ln\left(\mathbf{L}\right)$ with respect to $\hat{\boldsymbol{\beta}}$.

Assuming distribution has equidispersion *after controlling for observed heterogeneity* (i.e. the Poisson model fits the conditional distribution), the MLE coefficient estimates are consistent and asymptotically efficient (general properties of ML estimators).

Poisson (null model)

```
. poisson art
Iteration 0:   log likelihood = -1742.5735
Iteration 1:   log likelihood = -1742.5735
Poisson regression                                Number of obs   =        915
                                                  LR chi2(0)      =       0.00
                                                  Prob > chi2     =          .
Log likelihood = -1742.5735                       Pseudo R2       =     0.0000
------------------------------------------------------------------------------
         art |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   .5264408   .0254082    20.72   0.000     .4766416      .57624
------------------------------------------------------------------------------
```

$$\hat{\mu} = e^{\hat{\beta}} = e^{0.526} = 1.69$$

Poisson Regression

```
. poisson art fem mar kid5 phd ment

Iteration 0:   log likelihood = -1651.4574
Iteration 1:   log likelihood = -1651.0567
Iteration 2:   log likelihood = -1651.0563
Iteration 3:   log likelihood = -1651.0563

Poisson regression                                Number of obs   =        915
                                                  LR chi2(5)      =     183.03
                                                  Prob > chi2     =     0.0000
Log likelihood = -1651.0563                       Pseudo R2       =     0.0525

------------------------------------------------------------------------------
         art |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         fem |  -.2245942   .0546138    -4.11   0.000    -.3316352   -.1175532
         mar |   .1552434   .0613747     2.53   0.011     .0349512    .2755356
        kid5 |  -.1848827   .0401272    -4.61   0.000    -.2635305   -.1062349
         phd |   .0128226   .0263972     0.49   0.627     -.038915    .0645601
        ment |   .0255427   .0020061    12.73   0.000     .0216109    .0294746
       _cons |   .3046168   .1029822     2.96   0.003     .1027755    .5064581
------------------------------------------------------------------------------
```

Predicted Values

Unmarried male with no kids, mentor = 9, PhD = 3

$$\mathbf{x}_i\hat{\boldsymbol{\beta}} = 0.305 + 0 + 0 + 0 + 0.0129(3) + 0.0255(9) = 0.573$$

The incidence rate (predicted mean): $\mu_i = e^{\mathbf{x}_i\boldsymbol{\beta}} = e^{0.573} = 1.77$

Probabilities:

$$\Pr(0) = \frac{\left(e^{-1.77}\right)\left(1.77^{0}\right)}{0!} = \frac{(0.170)(1)}{1} = 0.17$$

$$\Pr(1) = \frac{\left(e^{-1.77}\right)\left(1.77^{1}\right)}{1!} = \frac{(0.170)(1.77)}{1} = 0.30$$

$$\Pr(2) = \frac{\left(e^{-1.77}\right)\left(1.77^{2}\right)}{2!} = \frac{(0.170)(3.13)}{2} = 0.27$$

$\Pr(3) = 0.16$, $\Pr(4) = 0.07$, and so on....

Prvalue

```
. prvalue, x(fem=0 kid5=0 mar=0 ment=9 phd=3)

poisson: Predictions for art

Confidence intervals by delta method
                                95% Conf. Interval
  Rate:                1.7735   [ 1.5802,    1.9669]
  Pr(y=0|x):           0.1697   [ 0.1369,    0.2025]
  Pr(y=1|x):           0.3010   [ 0.2756,    0.3264]
  Pr(y=2|x):           0.2669   [ 0.2603,    0.2735]
  Pr(y=3|x):           0.1578   [ 0.1367,    0.1789]
  Pr(y=4|x):           0.0700   [ 0.0530,    0.0870]
  Pr(y=5|x):           0.0248   [ 0.0161,    0.0335]
  Pr(y=6|x):           0.0073   [ 0.0040,    0.0107]
  Pr(y=7|x):           0.0019   [ 0.0008,    0.0029]
  Pr(y=8|x):           0.0004   [ 0.0001,    0.0007]
  Pr(y=9|x):           0.0001   [ 0.0000,    0.0001]

     fem    mar   kid5    phd   ment
x=     0      0      0      3      9
```

Marginal Effect on Mean

For the person described above, the marginal effect of an additional publication by the mentor is

$$\frac{\partial \hat{\mu}_i}{\partial x_k} = \frac{\partial e^{\mathbf{x}_i \hat{\boldsymbol{\beta}}}}{\partial x_k} = e^{\mathbf{x}_i \hat{\boldsymbol{\beta}}} \hat{\beta}_k = \mu_i \hat{\beta}_k$$

$$\frac{\partial \hat{\mu}_i}{\partial ment_i} = (1.77)(0.0255) = 0.0451$$

```
. prchange, x(fem=0 kid5=0 mar=0 ment=9 phd=3)

poisson: Changes in Rate for art
      min->max      0->1     -+1/2    -+sd/2  MargEfct
 fem   -0.3568   -0.3568   -0.3992   -0.1987   -0.3983
 mar    0.2978    0.2978    0.2756    0.1303    0.2753
kid5   -0.7550   -0.2994   -0.3284   -0.2510   -0.3279
 phd    0.0876    0.0220    0.0227    0.0224    0.0227
ment    8.6639    0.0365    0.0453    0.4307    0.0453

exp(xb):   1.7735
             fem        mar      kid5       phd      ment
    x=         0          0         0         3         9
sd(x)=   .498679   .473186    .76488   .984249   9.48392
```

Incidence Rate Ratio (IRR) , aka Factor Change in the Mean

What is the change in the mean (the expected number of publications, or whatever) if X_2 is increased by 1 unit?

$$X_2' = X_2 + 1$$

$$\mu = e^{\beta_1 + \beta_2 X_{2i} + \ldots + \beta_k X_{ki}}$$

$$\mu' = e^{\beta_1 + \beta_2 (X_{2i} + 1) + \ldots + \beta_k X_{ki}}$$

$$= e^{\beta_1 + \beta_2 X_{2i} + \beta_2 + \ldots + \beta_k X_{ki}}$$

The IRR is *constant:*

$$\frac{\mu'}{\mu} = \frac{e^{\beta_1 + \beta_2 X_{2i} + \beta_2 + \ldots + \beta_k X_{ki}}}{e^{\beta_1 + \beta_2 X_{2i} + \ldots + \beta_k X_{ki}}} = e^{\beta_2}$$

Generally:

e^{β_j} when X_j changes by 1 unit, or

$e^{\delta \beta_j}$ when X_j changes by δ

Factor Change: Example

How does having a young child affect articles published during last 3 years of Ph.D? For one child:

$$\hat{\beta}_{kid5} = -0.185 \qquad \frac{\mu'}{\mu} = e^{-0.185} = 0.83$$

The mean number of articles declines 17 percent. Note: does not depend on what the original mean was or the values of the other Xs (like odds ratio in logit). (Also, long way works.)

For two children: $\dfrac{\mu''}{\mu} = e^{(2)(-0.185)} = 0.69$

Could use a standard deviation change in X, if appropriate.

Exposure Time May Vary

- μ is a rate (arrivals per hour, articles per career), etc.
- Observations with longer time period should have greater counts even if the same mean.
- E.g. if rate is 5 per hour, then expected count is 10 over two hours, 15 over three hours, etc.
- If data consist of observations observed for different periods, this “exposure time” must be taken into account.

$$E\left[y_i \mid \mathbf{x}_i, t_i\right] = \mu_i t_i$$

Incorporating Exposure Time

$$\mu_i = e^{\mathbf{x}_i\boldsymbol{\beta}}$$

$$\mu_i t_i = e^{\mathbf{x}_i\boldsymbol{\beta}} t_i$$

$$= e^{\mathbf{x}_i\boldsymbol{\beta}} e^{\ln t_i}$$

$$= e^{\left(\beta_1 + \beta_2 X_{2i} + \ldots + \beta_K X_K + \ln t_i\right)}$$

Exposure time can be included in log form with the coefficient constrained to 1.

poisson options:

exposure(varname_e)
```
include ln(varname_e) in model with coefficient constrained to 1
```

offset(varname_o)
```
include varname_o in model with coefficient constrained to 1
```

Poisson vs. Reg: Ment

```
. poisson art ment
Poisson regression                                Number of obs   =        915
                                                  LR chi2(1)      =     147.86
                                                  Prob > chi2     =     0.0000
Log likelihood = -1668.6432                       Pseudo R2       =     0.0424
------------------------------------------------------------------------------
         art |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ment |   .0260498   .0019175    13.59   0.000     .0222917     .029808
       _cons |   .2599057   .0343609     7.56   0.000     .1925596    .3272518

. reg art ment
      Source |       SS       df       MS              Number of obs =     915
-------------+------------------------------           F(  1,   913) =   94.23
       Model |  317.204914     1  317.204914           Prob > F      =  0.0000
    Residual |  3073.49891   913   3.3663734           R-squared     =  0.0936
-------------+------------------------------           Adj R-squared =  0.0926
       Total |  3390.70383   914   3.7097416           Root MSE      =  1.8348
------------------------------------------------------------------------------
         art |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ment |   .0621168   .0063991     9.71   0.000     .0495581    .0746755
       _cons |   1.148305   .0826231    13.90   0.000     .9861515    1.310458
```

Approximate Log Model

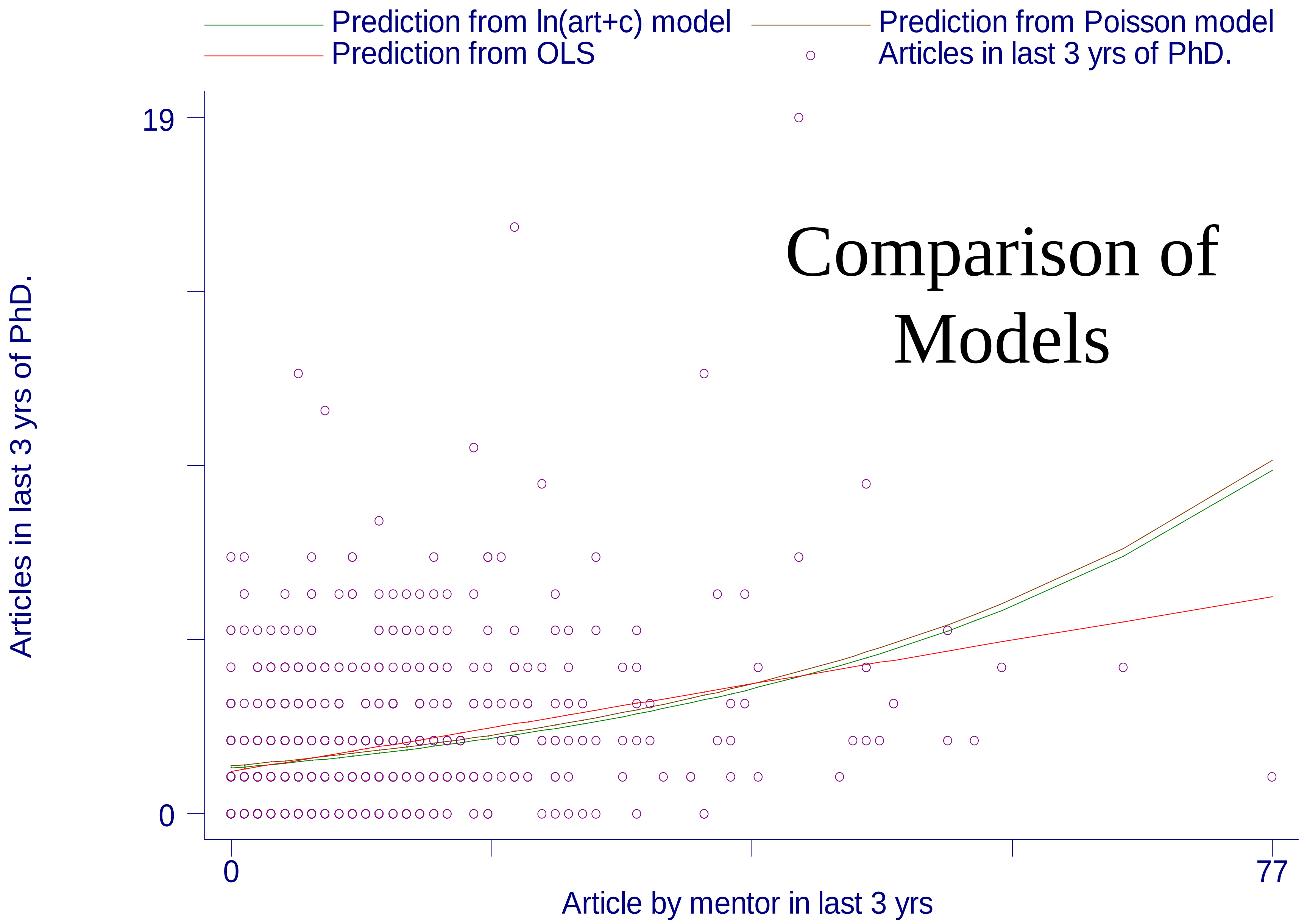
```
                  storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
art             byte   %8.0g                  Articles in last 3 yrs of PhD
lnart           float  %9.0g                  Log of art + .5

. reg lnart ment
      Source |       SS       df       MS              Number of obs =     915
-------------+------------------------------           F(  1,   913) =   84.96
       Model |  57.1044744     1  57.1044744           Prob > F      =  0.0000
    Residual |  613.632624   913  .672105832           R-squared     =  0.0851
-------------+------------------------------           Adj R-squared =  0.0841
       Total |  670.737099   914  .733848029           Root MSE      =  .81982

------------------------------------------------------------------------------
       lnart |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ment |   .0263557   .0028593     9.22   0.000     .0207441    .0319672
       _cons |   .2088502   .0369181     5.66   0.000      .136396    .2813044

. predict lnarthat
(option xb assumed; fitted values)

. gen arthat=exp(lnarthat)
```


Prediction from ln(art+c) model
Prediction from Poisson model
Prediction from OLS
Articles in last 3 yrs of PhD.
Comparison of Models
19
0
Articles in last 3 yrs of PhD.
0
77
Article by mentor in last 3 yrs

Poisson vs. Reg: Kid5

```
. poisson art kid5
Poisson regression                                Number of obs   =        915
                                                  LR chi2(1)      =       4.20
                                                  Prob > chi2     =     0.0405
Log likelihood = -1740.4757                       Pseudo R2       =     0.0012
------------------------------------------------------------------------------
         art |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        kid5 |  -.0697818    .034501    -2.02   0.043    -.1374026   -.002161
       _cons |   .5596001   .0298806    18.73   0.000     .5010353    .6181649

. reg art kid5
------------------------------------------------------------------------------
         art |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        kid5 |   -.113856   .0832528    -1.37   0.172     -.277245    .0495331
       _cons |   1.749264   .0758244    23.07   0.000     1.600454    1.898075
------------------------------------------------------------------------------

. reg lnart kid5
------------------------------------------------------------------------------
       lnart |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        kid5 |  -.0583203   .0370156    -1.58   0.115    -.1309659    .0143252
       _cons |   .4687894   .0337128    13.91   0.000     .4026258    .5349531
```

Observed vs. Unobserved Heterogeneiety

$$Poisson: \quad \mu_i = E\left(Y_i \mid \mathbf{X}_i\right) = e^{\mathbf{X}_i \boldsymbol{\beta}} \quad Var\left(Y_i \mid \mathbf{X}_i\right) = \mu_i$$

Assumes all observations with the same Xs have the same mean. This is not realistic (e.g. left out variables). Also assumes independence: publishing one article does not change the mean (contagion effect). Also unrealistic.

$$Negative\ Binomial: \mu_i = e^{\left(\mathbf{X}_i \boldsymbol{\beta} + \varepsilon_i\right)} \quad Var\left(Y_i \mid \mathbf{x}_i\right) = ??$$

Allows for *unobserved* heterogenity, and also allows for non-independence (contagion). Read results the same as Poisson.

Incorporating Overdispersion

$$\mu_i^{NB} = e^{\mathbf{x}_i\boldsymbol{\beta}+\varepsilon_i} = e^{\mathbf{x}_i\boldsymbol{\beta}}e^{\varepsilon_i} = \mu_i\delta_i$$

Need an identifying assumption. If we assume E(δ_i)=1 (corresponding to E(ε_i)=0 as in OLS, then:

$$E\left[\mu_i^{NB}\right] = E\left[\mu_i\delta_i\right] = \mu_i\left(1\right) = \mu_i$$

The NB has the same expected mean as the Poisson, but the standard errors in Poisson are biased downward, leading to incorrectly large z scores.

If we also assume δ_i has a gamma distribution with a constant variance, then $Y_i|\mathbf{X}_i$ has a negative binomial distribution.

$$Var(Y_i \mid X) = \mu_i + \alpha\mu_i^2$$

Alpha is an overdispersion parameter (from either unobserved heterogenity or “contagion”).

Negative Binomial Regression

nbreg art fem mar kid5 phd ment

```
Negative binomial regression                      Number of obs   =        915
                                                  LR chi2(5)      =      97.96
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -1560.9583                       Pseudo R2       =     0.0304
------------------------------------------------------------------------------
         art |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         fem |  -.2164184   .0726724    -2.98   0.003    -.3588537   -.0739832
         mar |   .1504895   .0821063     1.83   0.067    -.0104359    .3114148
        kid5 |  -.1764152   .0530598    -3.32   0.001    -.2804105     -.07242
         phd |   .0152712   .0360396     0.42   0.672    -.0553652    .0859075
        ment |   .0290823   .0034701     8.38   0.000     .0222811    .0358836
       _cons |    .256144   .1385604     1.85   0.065    -.0154294    .5277174
-------------+----------------------------------------------------------------
    /lnalpha |  -.8173044   .1199372                     -1.052377   -.5822318
-------------+----------------------------------------------------------------
       alpha |   .4416205   .0529667                      .3491069    .5586502
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  180.20 Prob>=chibar2 = 0.000
```

Zero Inflated Models

Often more zeros than expected. If we assume that a *separate process generates at least some of the zeros*, we get the ZIP or ZINB model:

$$\Pr\left(Y_i = 0 \mid \mathbf{x}_i\right) = \Psi_i + \left(1 - \Psi_i\right)\left(poisson\ part\right)$$

$$\Pr\left(Y_i > 0 \mid \mathbf{x}_i\right) = \left(1 - \Psi_i\right)\left(poisson\ part\right)$$

$$e.g.\quad \Psi_i = \Phi\left(\mathbf{z_i}\boldsymbol{\gamma}\right)$$

$$E\left(Y_i \mid \mathbf{x}_i, \mathbf{z}_i\right) = \left[0 \times \Psi_i\right] + \left[\mu_i \times \left(1 - \Psi_i\right)\right]$$

$$= \mu_i - \mu_i \Psi_i$$

```
zip depvar [indepvars] [if] [in] [weight],
               inflate(varlist[, offset(varname)]|_cons) [options]

zinb depvar [indepvars] [if] [in] [weight],
               inflate(varlist[, offset(varname)]|_cons) [options]
```

You need to specify the equation that generates the extra zeros.

ZIP Analysis

```
. zip art fem mar kid5 phd ment, inflate(fem mar kid5 phd ment)

Zero-inflated Poisson regression                  Number of obs     =        915
                                                  Nonzero obs       =        640
                                                  Zero obs          =        275
Inflation model = logit                           LR chi2(5)        =      78.56
Log likelihood  = -1604.773                       Prob > chi2       =     0.0000
------------------------------------------------------------------------------
         art |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
art          |
         fem |  -.2091446   .0634047    -3.30   0.001    -.3334155   -.0848737
         mar |    .103751    .071111     1.46   0.145     -.035624     .243126
        kid5 |  -.1433196   .0474293    -3.02   0.003    -.2362793   -.0503599
         phd |  -.0061662   .0310086    -0.20   0.842     -.066942    .0546096
        ment |   .0180977   .0022948     7.89   0.000     .0135999    .0225955
       _cons |    .640839   .1213072     5.28   0.000     .4030814    .8785967
-------------+----------------------------------------------------------------
inflate      |
         fem |   .1097465   .2800813     0.39   0.695    -.4392028    .6586958
         mar |  -.3540107   .3176103    -1.11   0.265    -.9765155    .2684941
        kid5 |   .2171001    .196481     1.10   0.269    -.1679956    .6021958
         phd |   .0012702   .1452639     0.01   0.993    -.2834418    .2859821
        ment |   -.134111   .0452461    -2.96   0.003    -.2227918   -.0454302
       _cons |  -.5770618   .5093853    -1.13   0.257    -1.575439     .421315
------------------------------------------------------------------------------
```

Zero Truncated Poisson

- Another possibility is *zero truncated* model.
- In general, the goal of modifications is to adjust for lack of fit by a standard Poisson.
- In this case, maybe you only have data on published authors, so that the minimum in the data is 1.

```
ztp depvar [indepvars] [if] [in] [weight] [, options]

ztnb depvar [indepvars] [if] [in] [weight] [, options]
```

Comparison of Poisson, Poisson on Zero Truncated Data, and Zero Truncated Poisson

	poisson	wrong	ztp
fem	-0.225***	-0.162**	-0.229***
	(-4.11)	(-2.98)	(-3.51)
mar	0.155*	0.0682	0.0965
	(2.53)	(1.11)	(1.32)
kid5	-0.185***	-0.0990*	-0.142**
	(-4.61)	(-2.47)	(-2.93)
phd	0.0128	-0.0131	-0.0127
	(0.49)	(-0.49)	(-0.41)
ment	0.0255***	0.0150***	0.0187***
	(12.73)	(7.08)	(8.22)
_cons	0.305**	0.827***	0.671***
	(2.96)	(8.04)	(5.48)
N	915	640	640

ZTNB Example

```
. drop if art==0
(275 observations deleted)

. ztnb art fem mar kid5 phd ment
Zero-truncated negative binomial regression       Number of obs   =        640
                                                  LR chi2(5)      =      44.58
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -1027.3185                       Pseudo R2       =     0.0212
------------------------------------------------------------------------------
         art |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         fem |  -.2446712   .0972181    -2.52   0.012    -.4352153   -.0541272
         mar |   .1034172   .1094297     0.95   0.345    -.1110611    .3178955
        kid5 |  -.1532593   .0722291    -2.12   0.034    -.2948257    -.011693
         phd |  -.0029336   .0480673    -0.06   0.951    -.0971437    .0912766
        ment |   .0237382   .0042868     5.54   0.000     .0153362    .0321402
       _cons |    .355125   .1968307     1.80   0.071    -.0306562    .7409062
-------------+----------------------------------------------------------------
    /lnalpha |  -.6034753   .2249915                     -1.044451   -.1625001
-------------+----------------------------------------------------------------
       alpha |   .5469076   .1230496                      .3518851     .850016
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  105.43 Prob>=chibar2 = 0.000
```