GENERALIZED LINEAR MODELS
Poisson Regression

Γιώτα Τουλούμη
Καθηγήτρια Βιοστατιστικής και Επιδημιολογίας
Εργ. Υγιεινής, Επιδημιολογίας και Ιατρικής Στατιστικής
Ιατρική Σχολή Πανεπιστημίου Αθήνας

gtouloum@med.uoa.gr

The Poisson distribution

Poisson regression is appropriate for

- variables that take non-negative integer values and have highly skewed (i.e., asymmetrical) distributions. For example counts or events over a period of time, like number of customers visiting a bank over a period of time, number of accidents, number of deaths etc
- Rates: events over total prys during which events happened
- Analysis of contingency tables (see Agresti Alan: Categorical data analysis).

The Poisson probability density function is given by

$$P(Y=r)=\frac{\lambda^r e^{-\lambda}}{r!}, r=0,1,2,...$$

The mean and variance of the Poisson distribution is

$$E(Y)=\lambda \text{ and } Var(Y)=\lambda$$

The Poisson as an approximation to the Binomial distribution

Example: Flying-bomb hits in London during World War II

A classic example that shows the derivation of the Poisson distribution as an approximation to the binomial $B(n,\pi)$ when $n \to \infty$ and $\pi \to 0$ but $\lambda = n\pi$ remains fixed is as follows:

The table below lists data from flying bomb hits in south London during WWII. The city was divided into 576 areas of one quarter square kilometers each. There were 537 hits, averaging $\hat{\lambda} = 0.9323$ hits per grid. The data are given below:

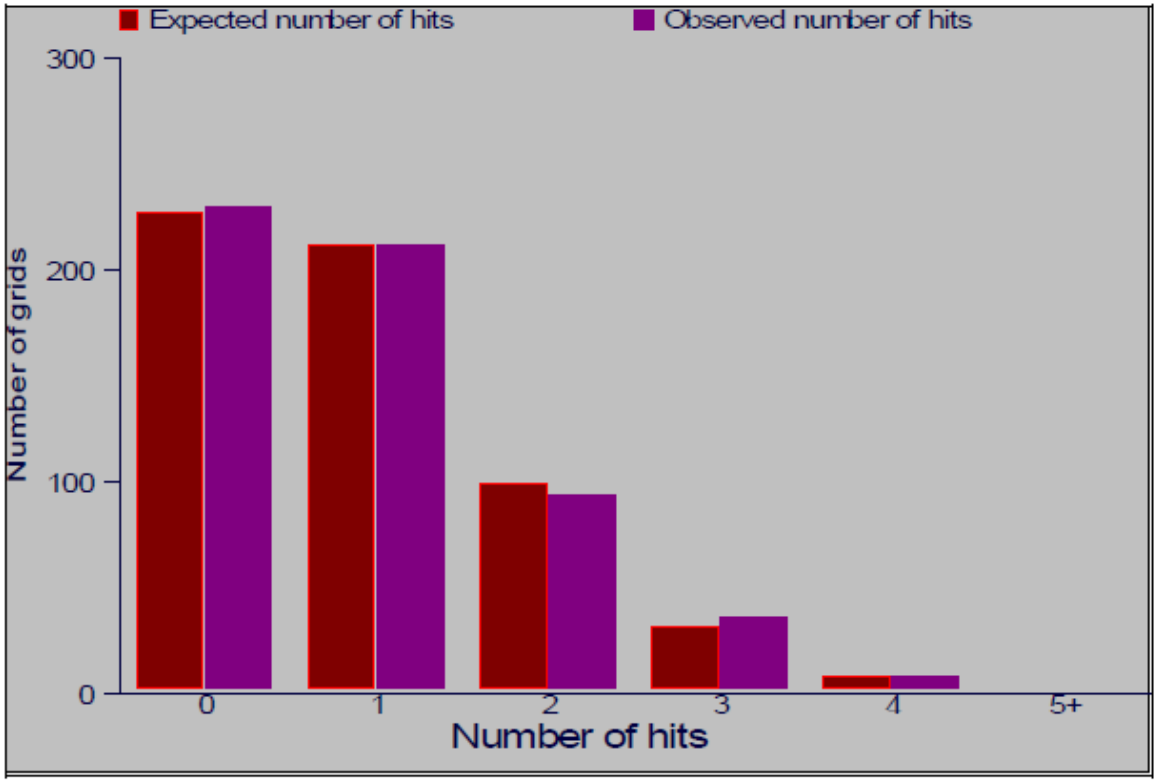
Hits	0	1	2	3	4	5+
Observed	229	211	93	35	7	1
Expected	226.7	211.4	98.6	30.6	7.1	1.6

λ=537/576= 0.9323

Assuming that each particular area had a small chance of being hit but having a large number of attempts leads to a Poisson distribution that approximates well a binomial $B(n, p)$.

Flying bomb hits of London (continued)

The observed and expected frequency distribution is given in the graph below. The agreement is astounding!



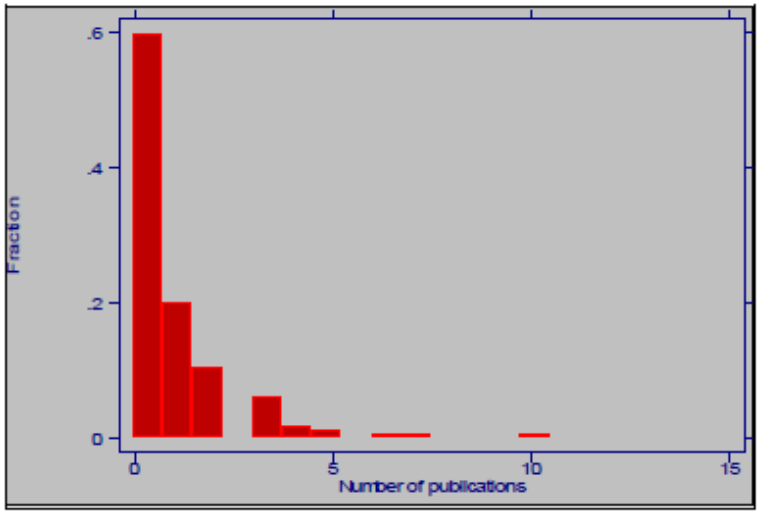
Scientific productivity example (McGinnis, Allison and Long, 1982, Allison, 1999)

An example of a data set that can be analyzed by Poisson methods is as follows: 557 male biochemists received their doctoral degree from 106 American universities in the late 1950s and 1960s.

PDOC	1 if received postdoctoral training, 0 otherwise
AGE	Age in years at completion of Ph.D.
MAR	1 if married, 0 otherwise
DOC	Measure of the prestige of the doctoral institution
UND	Measure of the selectivity of the undergraduate institution
AG	1 if degree is from an agricultural department, 0 otherwise
ARTS	Number of articles published while a graduate student
CITS	Number of citations to published articles
DOCID	ID number of the doctoral institution

Scientific productivity example (continued)

The frequency distribution of the number of publications is given below: This distribution is a good candidate for analysis by a Poisson model in terms of its skewness and few non-zero observations.



The goodness of fit test for a Poisson distribution however, is highly significant (i.e., does not support a Poisson-distributed variable). Notice that you must run a Poisson model before `poisgof`.

```
. quietly poisson arts

. poisgof

        Goodness of fit chi-2 =  1087.821
        Prob > chi2(556)      =    0.0000
```

Analysis with a Poisson GLM

As in any GLM analysis, the expected value of the outcome variable Y, or a function thereof, is associated with a linear combination of the explanatory variables as follows:

$$g[E(Y)] = g(\lambda) = \beta_o + \beta_1 X_1 + \cdots + \beta_p X_p$$

In the case of the Poisson mean, because λ is always positive, the function $g(.)$ is chosen so that the linear predictor $\eta = \beta_o + \beta_1 X_1 + \cdots + \beta_1 X$, that can take any real-number value, gets mapped into the positive real numbers. A good candidate function (link) for the Poisson GLM is the logarithm as follows:

$$\log(\lambda) = \beta_o + \beta_1 X_1 + \cdots + \beta_p X_p = \eta$$

The coefficients $\beta_o, \beta_1, \cdots, \beta_p$ are estimated via maximum-likelihood estimation.

Analysis of the scientific-productivity example

We carry out the Poisson regression using either the `poisson` or `glm` command in STATA. Here we prefer the `glm` command, because it produces the deviance that will be useful in the following.

```
. xi: glm   arts age i.mar doc  und i.ag , nolog fam(poisson)
i.mar                Imar_0-1      (naturally coded; Imar_0 omitted)
i.ag                 Iag_0-1       (naturally coded; Iag_0 omitted)

Residual df  =       551                              No. of obs =       557
Pearson X2   =   1497.36                              Deviance   =  1078.906
Dispersion   =  2.717532                              Dispersion =  1.958087

Poisson distribution, log link
------------------------------------------------------------------------------
        arts |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.0165613   .0101663     -1.629   0.103      -.0364868    .0033642
      Imar_1 |  -.0153611   .1300267     -0.118   0.906      -.2702088    .2394865
         doc |  -.0000399   .0004551     -0.088   0.930      -.0009319    .0008521
         und |   .0723311   .0303235      2.385   0.017       .0128981    .1317641
       Iag_1 |   .0421593    .099889      0.422   0.673      -.1536194     .237938
       _cons |  -.0401209   .3897091     -0.103   0.918      -.8039366    .7236948
------------------------------------------------------------------------------
```

Interpretation of the coefficients

The coefficients $\beta_1,\cdots,\beta_p$ denote the change in $\log(\lambda)$ for each one-unit change in the corresponding explanatory variable. In our example, the only significant variable is UND, the selectivity index of the under graduate institution. So, if two observations i and j have a difference of one unit in explanatory variable X_4 (UND), that is $X_{4i} - X_{4j} = 1$, while all the other explanatory variables are the same, then the difference in $\log(\lambda)$ will be $\log(\lambda_i) - \log(\lambda_j) = \log\left[\frac{\lambda_i}{\lambda_j}\right] = \beta_4$. In other words,

$$\frac{\lambda_i}{\lambda_j} = e^{\beta_k} \Leftrightarrow \frac{\lambda_i}{\lambda_j} - \frac{\lambda_j}{\lambda_j} = e^{\beta 4} - 1 \Leftrightarrow \frac{\lambda_i - \lambda_j}{\lambda_j} = e^{\beta 4} - 1$$

That is, $e^{\beta_4} - 1 = e^{0.0723} - 1 = 0.07498$ (i.e., about 7.5%) is the percent increase in the expected number of publications for each unit increase of the selectivity of the undergraduate institution.

Overdispersion

By the assumptions of the Poisson model, the expected value (mean) of the Poisson distribution is theoretically equal to its variance. Frequently this is not the case and the variance is much higher than the mean. In that situation, we have what is called *overdispersion*.

One way to detect this is by inspection of the `Dispersion` category below the deviance or Pearson chi-square statistics. This is the deviance or Pearson chi-square statistic divided by the number of degrees of freedom. If it is much larger than 1.0, it may indicate the presence of overdispersion.

As a cautionary note one must be aware that the deviance and Pearson chi-square statistics do not approximate a chi-square distribution well in the case of individual data or when the predicted values are small. In this case however, the scaled deviance value of 1.96 and scaled Pearson chi-square of 2.72 point to a potential problem with the model.

Overdispersion (continued)

One way to deal with overdispersion is to divide the chi-square statistic that tests the significance of each variable by the scaled deviance or scaled Pearson chi-square (or equivalently multiply each standard error by the square root of the scaled deviance or scaled Pearson chi-square; Agresti, 1996). The theory of quasi-likelihood (McCullagh and Nelder, 1989) suggests that the latter is better. In the example this is done as follows:

Variable	Estimate	Scaled Pearson X2	Standard error	Adjusted standard error	Adjusted *z*	*p*
AGE	-0.0165613	2.717532	0.0101663	0.0167590	-0.988	0.323
MAR	-0.0153611	2.717532	0.1300267	0.2143482	-0.072	0.943
DOC	-0.0000399	2.717532	0.0004551	0.0007502	-0.053	0.958
UND	0.0723311	2.717532	0.0303235	0.0499881	1.447	0.148
AG	0.0421593	2.717532	0.0998890	0.1646663	0.256	0.798

We see that although the estimates themselves do not change, there are no significant predictors of the number of publications, which indicates that our original results were possibly wrong!

Sqrt(2.717532)=1.6848939

Analysis accounting for overdispersion

The adjustment of the tests and estimates above can be performed automatically by including the option `scale(x2)` (the `x2` in the parenthesis indicates Pearson X2). The results are identical to the above table.

```
. xi: glm   arts age i.mar doc  und i.ag , nolog fam(poisson) scale(x2)
i.mar                Imar_0-1      (naturally coded; Imar_0 omitted)
i.ag                 Iag_0-1       (naturally coded; Iag_0 omitted)

Residual df  =       551                              No. of obs =       557
Pearson X2   =   1497.36                              Deviance   =  1078.906
Dispersion   =  2.717532                              Dispersion =  1.958087

Poisson distribution, log link
------------------------------------------------------------------------------
        arts |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.0165613    .016759    -0.988   0.323      -.0494084    .0162858
      Imar_1 |  -.0153611   .2143482    -0.072   0.943      -.4354759    .4047536
         doc |  -.0000399   .0007502    -0.053   0.958      -.0015103    .0014305
         und |   .0723311   .0499881     1.447   0.148      -.0256439     .170306
       Iag_1 |   .0421593   .1646663     0.256   0.798      -.2805808    .3648994
       _cons |  -.0401209    .642433    -0.062   0.950      -1.299266    1.219025
------------------------------------------------------------------------------
(Standard errors scaled using square root of Pearson X2-based dispersion)
```

Accounting for overdispersion: The Negative Binomial distribution

The Poisson model $\log(\lambda) = \beta_o + \beta_1 X_1 + \cdots + \beta_p X_p$, does not allow for extra variability and is thus susceptible to problems with overdispersion. One way to correct for that while avoiding the inefficient adjustment procedures discussed earlier is by introducing an extra variation term ε.

$$\log(\lambda) = \beta_o + \beta_1 X_1 + \cdots + \beta_p X_p + \varepsilon$$

Assuming that the distribution of *Y conditional* on ε i.e., $F(Y|\varepsilon)$ is Poisson with parameter λ and the distribution of exp(ε) is standard Gamma (Agresti, 1996, p. 74), then the *unconditional* distribution of Y $F(Y)$ is negative binomial.

The negative binomial model is fit in STATA either by the `nbreg` command, or the `glm` command by specifying `family(nbinom)` as the family of distributions. The default is a log link.

Analysis via negative binomial regression

The analysis via negative-binomial regression is produced below:

```
. xi: glm  arts  age i.mar doc und i.ag , family(nbinom) nolog
i.mar                Imar_0-1      (naturally coded; Imar_0 omitted)
i.ag                 Iag_0-1       (naturally coded; Iag_0 omitted)

Residual df  =       551                         No. of obs =       557
Pearson X2   =   805.475                         Deviance   =  602.3391
Dispersion   =  1.461842                         Dispersion =  1.093174

Negative Binomial (k=1) distribution, log link
------------------------------------------------------------------------------
        arts |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.0179106   .0141556    -1.265   0.206      -.0456552    .0098339
      Imar_1 |  -.0082166   .1788316    -0.046   0.963      -.3587201    .3422869
         doc |   .0000457   .0006196     0.074   0.941      -.0011686     .00126
         und |   .0709435   .0411262     1.725   0.085      -.0096623    .1515494
       Iag_1 |   .0463119   .1356871     0.341   0.733      -.2196299    .3122537
       _cons |  -.0272551    .540429    -0.050   0.960      -1.086476    1.031966
------------------------------------------------------------------------------
```

The coefficients are similar to those generated by the poisson regression model, and the dispersion value is a great deal closer to 1.0. The undergraduate selectivity index is significant at the 10% level but not the 5% level in this analysis. No other factors are significant.

$$Var(Y) = \mu + k\mu^2$$

Adjustment for varying time of observation

The Poisson and negative-binomial model assume a fixed (or constant) time of observation, like accidents over a period of time, colds over a season and so on. In the previous example, the data were collected over the same period of time for all observations.

When this is not the case, the varying time of observation must be accounted for by the model.

In the Poisson model we incorporate time into the model as follows:

$$P(Y_i = r) = \frac{(\lambda_i t_i)^r e^{-\lambda_i t_i}}{r!}, \text{ where } r = 0,1,2,...$$

where t_i is the time of observation for subject i, so that the expected number of occurrences is $\lambda_i t_i$.

Models for Rates

Examples:

- Incidence rate of lung cancer in Finish females in 1990
- Mortality rate of men working in the rubber-manufacturing industry

British doctors study

The following data are from a famous cohort study with main aim the investigation of the effect of smoking on coronary heart disease (CHD) among male British doctors.

Agegr	Smokes	Deaths	prys
1: 35-44	1	32	52407
2:45-54	1	104	43248
3:55-64	1	206	28612
4:65-74	1	186	12663
5: 75+	1	102	5317
1: 35-44	0	2	18790
2:45-54	0	12	10673
3:55-64	0	28	5710
4:65-74	0	28	2585
5: 75+	0	31	1462

Smokes: 1 for smokers 0 for non-smokers

Construction of Poisson frequency records

Usually data are collected by individual and therefore are stored in the form:

id	Age (in years)	Smokes	Date entry	Date exit	CHD
101	44	1	1/1/54	3/8/71	1
102	51	1	3/9/58	5/10/69	0

Note: Remember the Lexis diagram (in epidemiology)

To construct from individual data Poisson frequency records as in the example above, you can use the stata commands: stset and stsplit (see manuals)

Exercise: Complete the following table

Group	Person-years of follow-up	CHD deaths	Death rate per 1000 person-years	Rate ratio
Non-Smokers	39,220	101	2,574	1
Smokers	142,247	630	4,429	1.732

Reminder: Crude rate ratio: ignoring age group. In this example=1.72 **INTERPRETATION ?**

Models for rates

Consider events which occur independently in periods of time t_i with rates λ_i. The random variable Y_i represent the number of events in periods of time t_i and have the Poison distributions, with mean $\mu_i=\lambda_i t_i$. The mean can be modelled using a linear predictor of p explanatory variables x_{i1}, x_{i2}, ..., x_{ip} via a suitable link function.

In Poisson nearly always link function: **Log**

- maps positive values of μ to the whole line for the linear predictor
- parameters easily interpreted in terms of **multiplicative** effects on the rates scale
- it is the canonical parameterization for the Poisson distribution

Model: $\ln(\lambda_i) = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}$

or in terms of the mean

$$\ln(\mu_i) - \ln(t_i) = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} \Rightarrow \ln(\mu_i) = \ln(t_i) + \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}$$

t_i with regression coefficient set equal to 1 (through an **offset** in the model)

Analysis of the doctor's data: effect of smoke

xi: poisson deaths i.smokes, e(prys)

i.smokes _Ismokes_0-1 (naturally coded; _Ismokes_0 omitted)

Iteration 0: log likelihood = -480.77391

Iteration 1: log likelihood = -480.52234

Iteration 2: log likelihood = -480.52206

Iteration 3: log likelihood = -480.52206

Poisson regression Number of obs = 10

LR chi2(1) = 29.09

Prob > chi2 = 0.0000

Log likelihood = -480.52206 Pseudo R2 = 0.0294

deaths	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
_Ismokes_1	.5422211	.1071834	5.06	0.000	.3321454	.7522968
_cons	-5.961822	.0995037	-59.92	0.000	-6.156845	-5.766798

prys (exposure)

Analysis of the doctor's data: effect of smoke – Interpretation

β_o=-5.96 : The estimated log-rate for non-smokers

$e^{-5.96}$=0.0026 or rate of CHD for non-smokers: 2.5799 per 1000 prys

β_1=0.542 : the estimated difference in log-rate between non-smokers and smokers

$e^{0.542}$=1.7194 or rate ratio=1.7194 or smokers have 71.94% higher probability of dying from CHD compared to non-smokers.

NOTE: The rate ratio 1.72 is crude, i.e. unadjusted for age (compare it with that found in exercise above).

Doctors'study: Adjusted for age effect of smoke

xi: poisson deaths i.smokes i.agegr, e(prys)

LR chi2(5) = 922.93

Prob > chi2 = 0.0000

Log likelihood = -33.600153 Pseudo R2 = 0.9321

deaths	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
_Ismokes_1	.3545356	.1073741	3.30	0.001	.1440862	.564985
_Iagegr_2	1.484007	.1951034	7.61	0.000	1.101611	1.866403
_Iagegr_3	2.627505	.1837273	14.30	0.000	2.267406	2.987604
_Iagegr_4	3.350493	.1847992	18.13	0.000	2.988293	3.712693
_Iagegr_5	3.700096	.1922195	19.25	0.000	3.323353	4.07684
_cons	-7.919326	.1917618	-41.30	0.000	-8.295172	-7.543479

prys (exposure)

Rate/1000 pys for non-smokers age 35-44: 0.3636 [exp(-7.1917618]

Adjusted Rate ratio: $e^{0.3545356}$=1.4255; 95% CI: ($e^{0.1440862}$, $e^{0.564985}$)= (1.16, 1.76)

Risk, Rate and survival time

1. Risk (or odds) are estimated from studies where each subject is assumed to have been followed for roughly the same length of time (e.g. case-control studies). Logistic regression is typically used.
2. Rates are estimated from studies where each subject cannot be assumed to have been followed for the same length of time (e.g. cohort studies, clinical trials without early withdrawals), but **true rate can be assumed to be constant over reasonably broad bands of time** (Lexis diagram). Poisson regression models are typically used.
3. The basis of the analysis of survival time is time to event, often with censoring. Cox's proportional hazards models and Kaplan-Meier survival curves are typically used (see survival course). It has similarities to Poisson regression, but is based on finer subdivision of time, (Note: a Poisson model for number of events is equivalent to an exponential model for times **between** events).

NOTE: In certain cases (as in the assignment) we can assume that prys or exposed population remains the same, e.g. number of daily total deaths in Athens, the population (and thus prys) is assumed to remain constant.

Number of utterances about prognosis

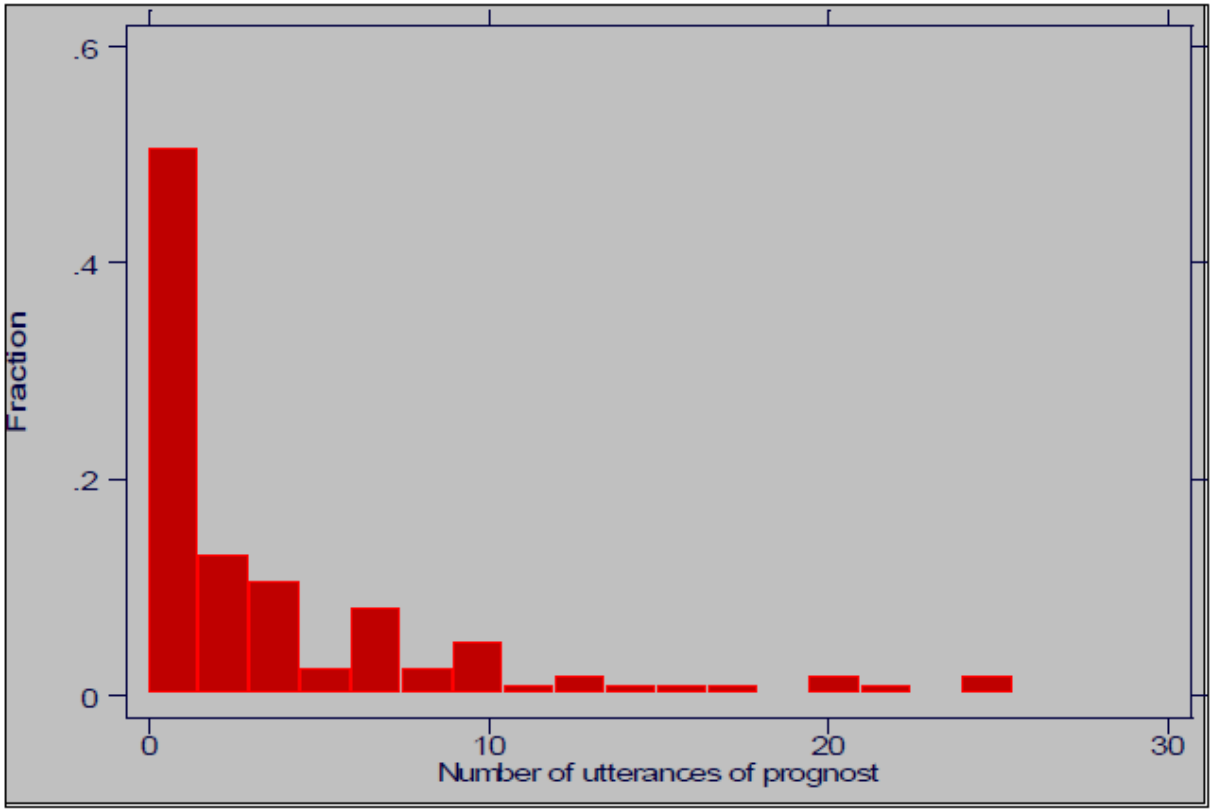
The following data set (Christakis and Levinson, 1998) describes the analysis of the number of utterances concerning prognosis by a doctor during a patient visit. The relevant variables are as follows:

LENGTHPX	Number of utterances regarding prognostic material
PTAGE	Patient age (years)
EZCOMPT	Doctor's rating of how easy it was to communicate with patient (1-5)
MDLIKEPT	Doctor's rating of how much they liked the patient (1-5)
SURGEON	1 if doctor is a surgeon, 0 otherwise
CLAIMS	Number of malpractice claims filed against the doctor
MINUTES	Length of visit in minutes

The problem with these data is that the length of observation (MINUTES) that is, duration of patient visit was not the same for all patients.

Number of utterances about prognosis

The frequency distribution of the LENGTHPX variable is given below:



We see that the data are highly skewed with a substantial proportion of observations at zero.

Offset

The way we incorporate the length of observation (duration of visit) is by adding what is called an "offset" variable to the model. This is done by adding the option `offset(varname)` or `lnoffset(varname)` in the `glm` command. The latter is what we need if the variable has not been transformed to the logarithmic scale already. The results of the analysis for these data are as follows:

```
. xi: glm  lengthpx ptage i.ptsex  ezcompt mdlikept i.surgeon claims, family(po
> isson) lnoffset( minutes) nolog
i.ptsex               Iptsex_0-1   (naturally coded; Iptsex_0 omitted)
i.surgeon             Isurge_0-1   (naturally coded; Isurge_0 omitted)

Residual df   =       114                                 No. of obs =       121
Pearson X2    =  899.6238                                 Deviance   =  682.0299
Dispersion    =  7.891437                                 Dispersion =  5.982718

Poisson distribution, log link, offset ln(minutes)
------------------------------------------------------------------------------
lengthpx |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
   ptage |  -.0014421   .0030592     -0.471   0.637      -.0074381    .0045538
Iptsex_1 |   .5482447   .1048294      5.230   0.000       .3427828    .7537066
 ezcompt |   .1980901   .0760461      2.605   0.009       .0490425    .3471377
mdlikept |  -.0864474   .0743869     -1.162   0.245       -.232243    .0593483
Isurge_1 |   1.343119   .1303694     10.302   0.000       1.087599    1.598638
  claims |   .0519112   .0231909      2.238   0.025       .0064579    .0973644
   _cons |  -3.175498   .3188579     -9.959   0.000      -3.800448   -2.550548
 minutes |  (exposure)
```

Interpretation of the analysis

Almost all variables are significant. It seems that there are 73% more utterances about prognosis when the subject is male ($e^{0.548}-1=0.73$), 22% when the physician thinks the patient is easier to communicate with ($e^{0.198}-1=0.22$), four times more when the physician is a surgeon ($e^{1.343}=3.83$) and 5.3% more for each malpractice claim that has been filed against the doctor ($e^{0.052}-1=0.053$). Notice that there is no coefficient corresponding to `minutes` because this has been constrained to be 1.0.

However, the results of this analysis are questionable, as the scaled Pearson chi-square and scaled deviance statistics are much larger than 1.0.

Thus, significant overdispersion is likely present in these data.

Correcting for overdispersion

To correct for overdispersion, we scale the test statistics corresponding to the coefficients by the scaled Pearson chi-square statistic. Only `surgeon` is significant in predicting prognosis utterances.

```
. xi: glm  lengthpx ptage i.ptsex  ezcompt mdlikept i.surgeon claims, family(po
> isson) lnoffset( minutes) nolog scale(x2)
i.ptsex                Iptsex_0-1    (naturally coded; Iptsex_0 omitted)
i.surgeon              Isurge_0-1    (naturally coded; Isurge_0 omitted)

Residual df  =       114                               No. of obs =       121
Pearson X2   =  899.6238                               Deviance   =  682.0299
Dispersion   =  7.891437                               Dispersion =  5.982718

Poisson distribution, log link, offset ln(minutes)
------------------------------------------------------------------------------
lengthpx |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
   ptage |  -.0014421   .0085938     -0.168   0.867      -.0182857    .0154014
Iptsex_1 |   .5482447   .2944837      1.862   0.063      -.0289327    1.125422
 ezcompt |   .1980901   .2136264      0.927   0.354        -.22061    .6167901
mdlikept |  -.0864474   .2089654     -0.414   0.679      -.4960121    .3231174
Isurge_1 |   1.343119   .3662297      3.667   0.000       .6253216    2.060916
  claims |   .0519112   .0651471      0.797   0.426      -.0757747    .1795971
   _cons |  -3.175498   .8957261     -3.545   0.000      -4.931089   -1.419907
 minutes |  (exposure)
------------------------------------------------------------------------------
(Standard errors scaled using square root of Pearson X2-based dispersion)
```

Correcting for overdispersion (continued)

The previous analysis may be inefficient, so we also undertake a negative binomial regression analysis. The results show that both patient's sex and whether the doctor is a surgeon are significant predictors of the outcome variable.

```
. xi: glm  lengthpx ptage i.ptsex  ezcompt mdlikept i.surgeon claims, family(nb
> inom) lnoffset( minutes) nolog
i.ptsex               Iptsex_0-1    (naturally coded; Iptsex_0 omitted)
i.surgeon             Isurge_0-1    (naturally coded; Isurge_0 omitted)

Residual df  =       114                                 No. of obs =       121
Pearson X2   =  203.1601                                 Deviance   =  197.2956
Dispersion   =  1.782106                                 Dispersion =  1.730663

Negative Binomial (k=1) distribution, log link, offset ln(minutes)
------------------------------------------------------------------------------
lengthpx |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
   ptage |   .0002237   .0069105      0.032   0.974      -.0133208    .0137681
Iptsex_1 |   .5829213   .2197179      2.653   0.008       .1522822     1.01356
 ezcompt |   .1291447   .1487535      0.868   0.385      -.1624067    .4206961
mdlikept |  -.1062208   .1532717     -0.693   0.488      -.4066278    .1941863
Isurge_1 |   1.407069    .246766      5.702   0.000       .9234163    1.890721
  claims |   .0514776   .0543761      0.947   0.344      -.0550975    .1580527
   _cons |  -2.758898   .6843703     -4.031   0.000      -4.100239   -1.417557
 minutes |  (exposure)
```

References

Agresti, A. (1996) *An introduction to Categorical Data Analysis*. New York: John Wiley &Sons.

Levinson W, Roter DL, Mullooly JP, Dull VT, and Frankel RM. (1997). Physician-patient communication. JAMA, **277**, 553-559.

Christakis, NA and Levinson, W. (1998). Casual optimism: Prognostication in routine medical and surgical encounters. Unpuslished manuscript

McCullagh, P. and Nelder, JA. (1989) *Generalized linear models*. Second edition. London: Chapman and Hall

McGinnis, R., Allison, PD and Long, JS (1982). Postdoctoral training in bioscience: Allocation and outcomes. *Social Forces*, **60**, 701-722.