Application of Generalized Linear Mixed Models in Neuroscience

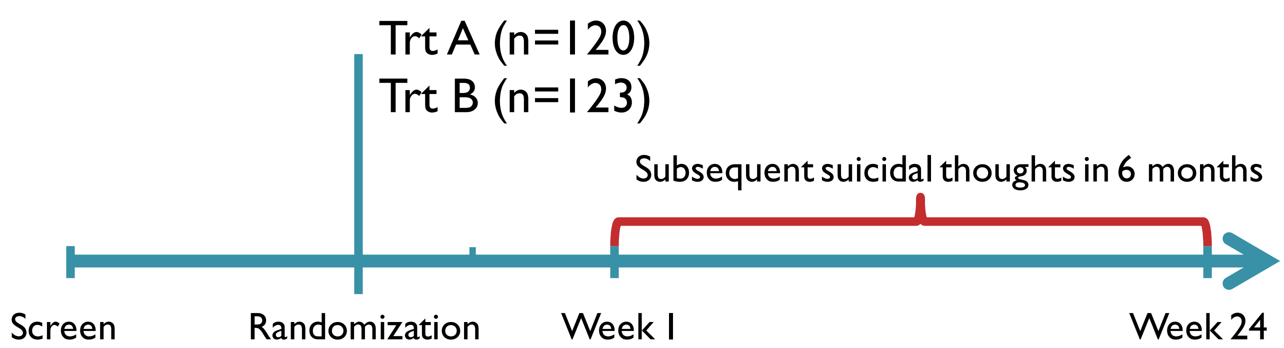
Dr Sourish Saha

Outline

- Motivation Example
- Background Information
- Censored Generalized Poisson
- Implementation in SAS
- Study Design with Generalized Poisson
- Discussions
- References

Depression

- Population & Endpoint
  - Patients with depression
  - # of suicidal thoughts in a week



Study Specified Analysis Method (t-test)

	Trt A	Trt B	P-value
N	120	118	
Mean	2.9	3.3	**0.053**
SD (Var)	1.69 (2.86)	1.52 (2.31)	

- Problems:
  - Data not normally distributed
  - Not considering early drop out
    - Observation periods different
  - Not considering censoring

Censored Poisson Regression

- Poisson distribution: useful for modeling small events

$$P(Y=y)=f(y;\mu)=\frac{e^{-\mu}\mu^{y}}{y!},\qquad y=0,1,2,\ldots,\infty$$

$$E(y)=Var(Y)=\mu$$

- Poisson distribution with censoring

$$P(Y=y)=f(y)=\frac{e^{-\mu}\mu^{y}}{y!},\qquad y=0,1,2,\ldots,c-1$$

$$P(Y=y)=\sum_{y=c}^{\infty}f(y)=\sum_{y=c}^{\infty}\frac{e^{-\mu}\mu^{y}}{y!},\quad y\geq c,\ c\text{ is cenosring point}$$

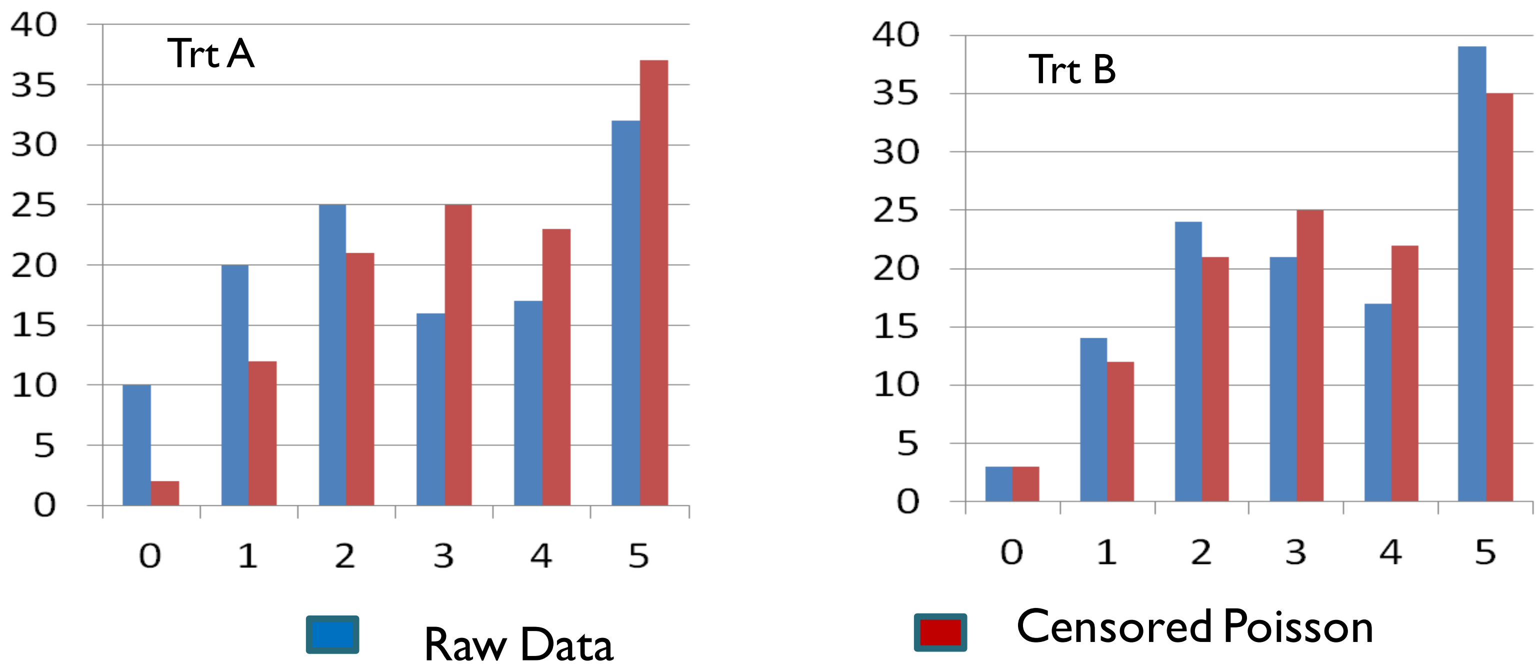
- Censored Poisson regression

$$\log(E(Y\mid X))=\log(t)+\beta_0+\beta_1 x$$

t : exposure time; $\log(t)$: offset; x : trt indicator

Predicted # of Suicidal Thoughts

	Trt A	Trt B	P-value
N	120	118	
Mean	3.6	4.2	**0.017**
SD (Var)	1.9 (3.61)	2.05 (4.20)	



Dispersion with Poisson

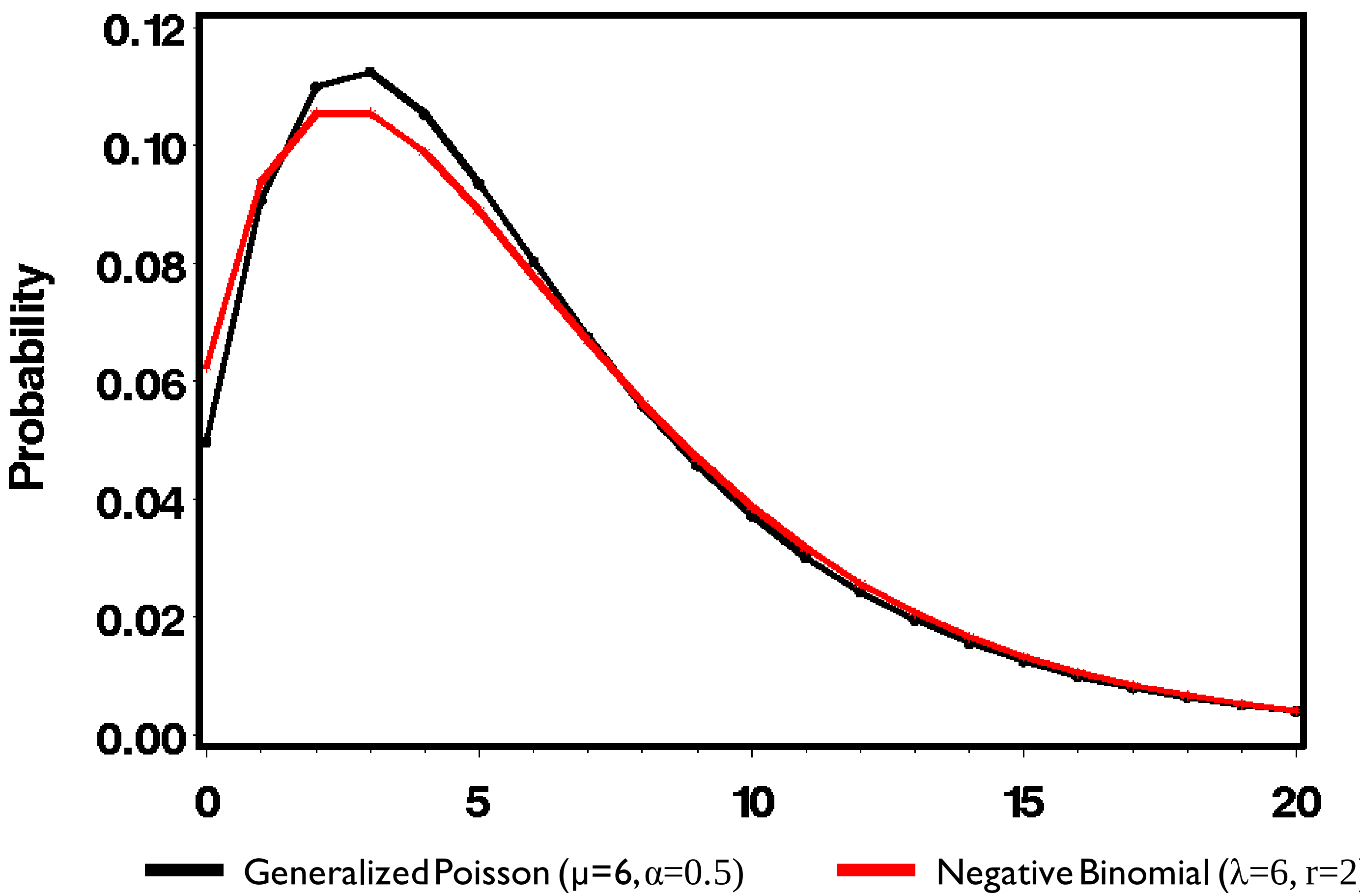
- Dispersion
  - Dispersion measurement: ϕ=var/mean
  - Over-disp.: $\phi > 1$; under-disp.: $\phi < 1$
  - Repeated events are correlated
- Methods to adjust for dispersion
  - Empirical method
    - Multiplying scale parameter ϕ to the variance
  - Parametric method
    - Negative binomial
    - **Generalized poisson**

Generalized Poisson vs. Negative Binomial

- Negative binomial
  - Over-dispersion only
  - More mass at 0
  - Gamma mixing of Poisson

- Generalized Poisson
  - Can be both over- and under-dispersion
  - Heavier tail
  - Poisson mixture (mixing distribution unknown) in over-dispersion case (Joe and Zhu 2005)

Generalized Poisson vs. Negative Binomial (cont.)



Same mean (6) and Variance (24)

Generalized Poisson Distribution

- Joe and Zhu 2005

$$P(Y=y)=f(y;\mu,\alpha)=\mu(1-\alpha)\cdot\left(\mu(1-\alpha)+\alpha y\right)^{y-1}\cdot\exp\left(-\mu(1-\alpha)-\alpha y\right)\cdot\frac{1}{y!},$$

$$y=0,1,2,\ldots,\infty$$

$$E(Y)=\mu,\qquad Var(Y)=\frac{\mu}{(1-\alpha)^2},\quad Dispersion=\frac{1}{(1-\alpha)^2}$$

$$0<\alpha<1:\text{over-disp;}\quad \alpha<0:\text{under-disp;}\quad \alpha=0:\text{reduces to Poisson}$$

- Mahmoud and Alderiny 2010

$$P(Y=y)=f(y;\mu,\alpha)=\frac{\mu}{1+\alpha\mu}\left(\frac{\mu(1+\alpha y)}{1+\alpha\mu}\right)^{y-1}\cdot\exp\left(-\frac{\mu(1+\alpha y)}{1+\alpha\mu}\right)\cdot\frac{1}{y!},$$

$$y=0,1,2,\ldots,\infty$$

$$E(Y)=\mu,\qquad Var(Y)=\frac{\mu}{(1+\alpha\mu)^2},\qquad Dispersion=\frac{1}{(1+\alpha\mu)^2}$$

$$\alpha>0:\text{over-disp.;}\quad \alpha<0:\text{under-disp.;}\quad \alpha=0:\text{reduces to Poisson}$$

Censored Generalized Poisson Regression

- Generalized Poisson with censoring

$$P(Y=y)=[f(y)]^{1-d}\left[1-\sum_{j=0}^{c-1} f(j)\right]^{d},$$

$$d=\begin{cases}1, & \text{if } y>c \\ 0, & o.w.\end{cases}, \quad d=0 \text{ reduces to non-censored case}$$

- Censored generalized Poisson regression

$$\log(E(Y\mid X))=\log(\mu)=\log(t)+\beta_0+\beta_1 x$$

Estimation of Parameters

- Maximum likelihood

$$L(\boldsymbol{\beta},\alpha;y_i)=\prod_{i=1}^{n}\left\{[f(y_i)]^{1-d_i}\left[1-\sum_{j=0}^{c_i-1}f(j)\right]^{d_i}\right\},\quad d_i=\begin{cases}1, & \text{if } y_i>c_i\\ 0, & o.w.\end{cases}$$

$$LL(\boldsymbol{\beta},\alpha;y_i)=\sum_{i=1}^{n}(1-d_i)\log[f(y_i)]+\sum_{i=1}^{n}d_i\log\left[1-\sum_{j=0}^{c_i-1}f(j)\right]$$

- Solving **β**, α simultaneously from likelihood equations by using iterative algorithm

Challenges in Censored Generalized Poisson Regression

- Censoring complicates likelihood function
- Additional dispersion parameter α to estimate
- Solving non-linear likelihood equations simultaneously

Poisson Regression in SAS

- Proc GENMOD
  - Can't handle censoring
  - Doesn't fit generalized Poisson

```
proc genmod data=data;
  class x;
  model y=x/dist=poisson offset=logt;
run;
```

- POICEN macro
  - Can handle censoring
  - Doesn't fit generalized Poisson

Generalized Poisson Regression in SAS

- Proc GLIMMIX
  - Available in SAS 9.2 and above
  - Fits generalized Poisson with/without censoring
- Proc NLMIXED
  - Available in earlier versions of SAS
  - Fits generalized Poisson with/without censoring

Implementation in SAS

Model	Poisson		Generalized Poisson	
Censored?	Yes	No	Yes	No
GENMOD		√		
POICEN macro	√			
GLIMMIX	√	√	√	√
NLMIXED	√	√	√	√

Implementation in Proc GLIMMIX

```
proc glimmix data=data;

  alpha = (1 - 1/exp(_phi_));
  _variance_ = _mu_ / ((1-alpha)*(1-alpha));

  select;
  when (c=0)
    mustar = _mu_*(1-alpha) + alpha*y;
    _logl_ = log(_mu_*(1-alpha))+(y -1)*log(mustar)- mustar - lgamma(y+1);
  when (c=1)
    lc=0;
    do j=0 to (y-1);
    mustar = _mu_*(1-alpha) + alpha*j;
    lc=lc + ((_mu_*(1-alpha))/gamma(j+1))*(mustar **(j -1))*exp(-mustar);
    end;
    _logl_ = log(1-lc);
  end;

    class x;
    model y=x /link=log s offset=logt;

run;
```

Implementation in Proc NLMIXED

```
proc nlmixed data=data;

  alpha = (1 - 1/exp(phi));
  mu = exp(logt + b0 + b1*x);

  select;
  when (c=0)
     mustar = mu*(1-alpha) + alpha*y;
     logl = log(mu*(1-alpha))+(y-1)*log(mustar) - mustar - lgamma(y+1);
  when (c=1)
    lc=0;
    do j=0 to (y-1);
   mustar = mu*(1-alpha) + alpha*j;
   lc=lc + ((mu*(1-alpha))/gamma(j+1))*(mustar**(j -1)) * exp(-mustar);
    end;
    logl = log (1-lc);
  end;

  model y ~ general(logl);

run;
```

Simple Requirement for Input Dataset

- Variables needed
  - X: treatment indicator
  - Y: response variable of interest
  - C: censoring variable
  - T: observation duration on log scale(optional)

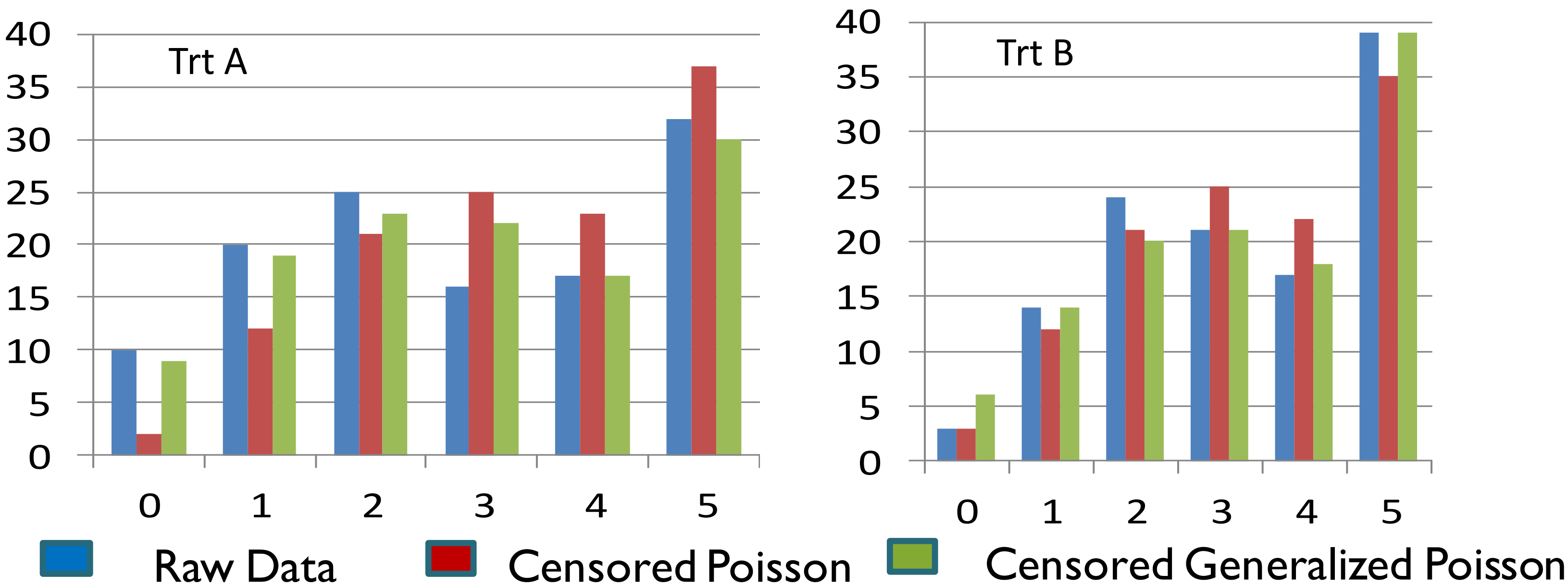
Results Based on Censored Generalized Poisson

Model	**Poisson**		**Generalized Poisson**			
Censored?	Yes	No	Yes		No	
	p-value	*p*-value	Disp.	*p*-value	Disp.	*p*-value
GENMOD	--	0.085				
POICEN macro	0.017	--				
GLIMMIX	0.046	0.086	1.6	0.063	0.8	0.076
NLMIXED	0.046	0.086	1.6	0.063	0.8	0.076

- Dispersion could be masked by censoring (raw data var<mean)!

Predicted # using Censored Generalized Poisson

	Trt A	Trt B	P-value
N	120	118	
Mean	2.9	3.3	**0.063**
SD (Var)	1.63 (2.66)	1.59 (2.53)	



- Censored generalized Poisson fits data best

Discussions

- Generalized Poisson
  - Useful tool for analyzing count data
  - Limitations:
    - Iteration may not converge
      - Change iteration algorithm
      - Change parameter initial values
      - Change convergence criterion
    - Limited software available for study design

References

- Joe, H. and Zhu, R. (2005), Generalized Poisson Distribution: The Property of Mixture of Poisson and Comparison with Negative Binomial Distribution. *Biometrical Journal*, 47, 219–229.
- Hilbe, J. (1995). POICEN: SAS code to estimate censored Poisson regression
- Gu, K. *et. al* (2008). Testing the ratio of two Poisson rates. *Biometrical Journal* 50, 2, 283-298
- Mahmoud, M. and Alderiny, M. (2010). On estimating parameters of censored generalized Poisson regression model. *Applied Mathematical Science*, Vol. 4, no. 13, 623-635
- Ismail, N. and Jemain, A. (2007). Handling overdispersion and negative binomial and generalized Poisson regression models. *Casualty Actuarial Society Forum*, Winter 2007.
- Berk, R. *et. al.* (2007). Overdispersion and Poisson regression.

- Back Up Slides

Differences among SAS Procedures

SAS Procedure	Likelihood Approximation	Optimization Algorithm
GENMOD	Exact likelihood	Ridge-stabilized Newton-Raphson
POICEN macro	Exact likelihood	Trust-region method
GLIMMIX	Method= **RSPL** (Residual subject-specific expansions pseudo-likelihood); MSPL (maximum subject-specific expansions pseudo-likelihood); RMPL (Residual marginal expansions pseudo-likelihood); MMPL (maximum marginal expansions pseudo-likelihood); LAPLACE (Maximum Likelihood with Laplace Approximation); QUAD (Maximum Likelihood with Adaptive Quadrature)	NLOPTIONS Teck= **QUANEW** (quasi-Newton); NEWRAP(Newton-Raphson combining a line-search algorithm (LIS=2) with ridging); CONGRA (conjugate-gradient); DBLDOG (double-dogleg); LEVMAR (Levenberg-Marquardt); NMSIMP (Nelder-Mead simplex); NRRIDG (Newton-Raphson with ridging); TRUREG (trust-region); NONE
NLMIXED	Method= **GAUSS** (adaptive Gauss-Hermite quadrature); FIRO (first-order method of Beal and Sheiner); HARDY(Hardy quadrature based on an adaptive trapezoidal rule); ISAMP (adaptive importance sampling)	Tech= **QUANEW**(Dual Quasi-Newton); CONGRA (conjugate-gradient); DBLDOG (double dogleg); NMSIMP (Nelder-Mead simplex); NEWRAP (Newton-Raphson); NRRIDG (Newton-Raphson with ridging); TRUREG (trust region); NONE

Likelihood Equations

$$\frac{\partial LL(\boldsymbol{\beta},\alpha;y_i)}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n}(1-d_i)\cdot\frac{1}{f(y_i)}\cdot\frac{\partial f(y_i)}{\partial \boldsymbol{\beta}} - \sum_{i=1}^{n}\left(d_i \cdot \frac{1}{1-\sum_{j=0}^{c_i-1} f(j)} \cdot \sum_{j=0}^{c_i-1}\frac{\partial f(j)}{\partial \boldsymbol{\beta}} \right) = 0$$

$$\frac{\partial LL(\boldsymbol{\beta},\alpha;y_i)}{\partial \alpha} = \sum_{i=1}^{n}(1-d_i)\cdot\frac{1}{f(y_i)}\cdot\frac{\partial f(y_i)}{\partial \alpha} - \sum_{i=1}^{n}\left(d_i \cdot \frac{1}{1-\sum_{j=0}^{c_i-1} f(j)} \cdot \sum_{j=0}^{c_i-1}\frac{\partial f(j)}{\partial \alpha} \right) = 0$$

Adjusting for Over-dispersion (1)

- Negative binomial distribution

$$P(y) = \left(\frac{r}{r+\lambda}\right)^r \frac{\Gamma(r+y)}{\Gamma(y+1)\Gamma(r)} \left(\frac{\lambda}{r+\lambda}\right)^y,$$

Γ is the gamma function.

The mean of the negative binomial distribution (like Poisson) is λ, but the variance is

$$\lambda + \frac{\lambda^2}{r} = \lambda\left(1 + \frac{\lambda}{r}\right)$$

where r is called the dispersion parameter.

Adjusting for Over-dispersion (2)

- Censored Negative Binomial distribution:

$$P(y) = \left(\frac{r}{r+\lambda}\right)^{r} \frac{\Gamma(r+y)}{\Gamma(y+1)\Gamma(r)} \left(\frac{\lambda}{r+\lambda}\right)^{y}, y = 0,1,\ldots,c-1$$

$$P(y) = \sum_{y=c}^{\infty} \left(\frac{r}{r+\lambda}\right)^{r} \frac{\Gamma(r+y)}{\Gamma(y+1)\Gamma(r)} \left(\frac{\lambda}{r+\lambda}\right)^{y},$$

$y \geq c$, c is censoring point

Adjusting for Over-dispersion (3)

- If Y is negative binomial distributed then

$$E(Y) = \lambda \text{ and } Var(Y) = \lambda + \frac{\lambda^2}{r} = \lambda(1 + \frac{\lambda}{r})$$

If we define

$$\phi = \frac{V(Y)}{E(Y)} = 1 + \frac{\lambda}{r}$$

then, when r -> ∞, negative binomial becomes Poisson distribution, i.e. the variance of negative binomial is over-dispersed by ϕ compared with a Poisson distribution with mean and variance of λ.

Φ is called an over-dispersion parameter.

Relationship between Poisson & Negative Binomial Distributions

Consider a sequence of negative binomial distributions where the stopping parameter *r* goes to infinity, whereas the probability of success in each trial, *p*, goes to zero in such a way as to keep the mean of the distribution constant.
Denoting this mean *λ*, the parameter *p* will have to be

$$\lambda = r\,\frac{p}{1-p} \quad \Rightarrow \quad p = \frac{\lambda}{r+\lambda}.$$

Under this parameterization the probability mass function will be

$$f(k) = \frac{\Gamma(k+r)}{k!\cdot\Gamma(r)}(1-p)^r p^k = \frac{\lambda^k}{k!}\cdot\frac{\Gamma(r+k)}{\Gamma(r)\;(r+\lambda)^k}\cdot\frac{1}{\left(1+\frac{\lambda}{r}\right)^r}$$

Now if we consider the limit as $r \to \infty$, the second factor will converge to one, and the third to the exponent function:

$$\lim_{r\to\infty} f(k) = \frac{\lambda^k}{k!}\cdot 1\cdot\frac{1}{e^\lambda},$$

which is the mass function of a Poisson-distributed random variable with expected value *λ*.

In other words, the alternatively parameterized negative binomial distribution converges to the Poisson distribution and *r* controls the deviation from the Poisson. This makes the negative binomial distribution suitable as a robust alternative to the Poisson, which approaches the Poisson for large *r*, but which has larger variance than the Poisson for small *r*.

$$\text{Poisson}(\lambda) = \lim_{r\to\infty} \text{NB}\Big(r,\ \frac{\lambda}{\lambda+r}\Big).$$

Poisson-Gamma mixture = Negative Binomial

The negative binomial distribution also arises as a continuous mixture of Poisson distributions where the mixing distribution of the Poisson rate is a gamma distribution.

That is, we can view the negative binomial as a Poisson(λ) distribution, where λ itself is a random variable, distributed according to *Gamma*(r, p/(1 − p)).

Formally, this means that the mass function of the negative binomial distribution can be written as

$$\begin{aligned} f(k) &= \int_0^\infty f_{\text{Poisson}(\lambda)}(k) \cdot f_{\text{Gamma}\left(r, \frac{p}{1-p}\right)}(\lambda) \, \mathrm{d}\lambda \\ &= \int_0^\infty \frac{\lambda^k}{k!} e^{-\lambda} \cdot \lambda^{r-1} \frac{e^{-\lambda(1-p)/p}}{\left(\frac{p}{1-p}\right)^r \Gamma(r)} \, \mathrm{d}\lambda \\ &= \frac{(1-p)^r p^{-r}}{k! \, \Gamma(r)} \int_0^\infty \lambda^{r+k-1} e^{-\lambda/p} \, \mathrm{d}\lambda \\ &= \frac{(1-p)^r p^{-r}}{k! \, \Gamma(r)} \, p^{r+k} \, \Gamma(r+k) \\ &= \frac{\Gamma(r+k)}{k! \, \Gamma(r)} \, (1-p)^r p^k. \end{aligned}$$