GUY CARPENTER

GLM beyond the Exponential Family

GARY G. VENTER

Marsh & McLennan Companies

Idea of GLM

Non-linear model with embedded linear regression

- Each observation is from a distribution whose mean is a non-linear function of a regression mean
- Distribution usually restricted to exponential family

Properties

- Estimation procedure quick – quasi-likelihood
- However not always MLE
  - Less efficient
  - Does not allow comparison of fits
- Distribution defined by variance as function of mean
- Skewness/CV related to variance/mean

Examples

- For variance = c*mean^k, skewness = k*CV
- k = 0, 1, 2, 3 or 1 < k < 2 give distributions:
- Normal, Poisson, Gamma, Inverse Gaussian,

Generalize to GLZ

Definition

- Mean μ is a parameter of the distribution for cell
- Distribution can have any number of other parameters which are constant for all the cells
- $\mu = \eta[\beta z]$ where z is the vector of independent variables (co-variates) for the cell and β is a vector of coefficients constant for all cells
- If mean exists, distribution can be reparam-eterized so mean is a parameter
  - If not, can do generalized median regression

Effects of not restrict-ing to exponential family

- Many distributions possible, like t, Pareto
- Estimation by MLE
  - Possibly more calculation-intensive
- Distributions from exponential family can be re-parameterized to have different relationships of variance and mean
- Usually skewness/cv is not changed by this

Examples Using Frequency, Severity, and Aggregate Loss Distributions

Negative binomial

- mean = rb
- variance = rb(1+b)

1st parameterization

- Replace r with μ/b
- Variance = μ(1+b) is proportional to mean
- $\mu = \eta[\beta z]$ varies by cell due to different z's
- b and β are constant across all cells

2nd parameterization

- Replace b with μ/r
- Variance = $\mu(1+\mu/r)$ is quadratic in mean
- $\mu = \eta[\beta z]$ varies by cell
- r and β are constant across all cells
- In exponential family

Severity Distributions (Defined on x>0 and Continuous)

Inverse Gaussian 1

$$ig_1(x;\mu,\alpha) = \sqrt{\frac{\mu}{2\pi\alpha x^3}}\,e^{\frac{2-x/\mu-\mu/x}{2\alpha}}$$

- EX = μ, variance = $\alpha\mu^2$, so quadratic in mean
- μ is a scale parameter

Inverse Gaussian 2

- More usual (exponential family) is:

$$ig_2(x;\mu,\lambda) = e^{\frac{2-x/\mu-\mu/x}{2\mu/\lambda}}\sqrt{\frac{\lambda}{2\pi x^3}}$$

- EX = μ, variance = μ^3/λ , so cubic in mean

Either way

- Skewness/CV = 3
- $\mu = \eta[\beta z]$ varies by cell due to different z's
- β and either λ or α constant across cells

More Severity Distributions – Gamma

Standard definition

—»

- $F(x; \theta, \alpha) = \Gamma(x/\theta; \alpha)$ with incomplete gamma
- $EX = \alpha\theta$, variance $= \alpha\theta^2$,
- $\theta = \mu/\alpha$, variance $= \mu^2/\alpha$, so quadratic in mean
- In exponential family

Gamma p

—»

- $F(x; \mu, \lambda, p) = \Gamma[x/(\lambda\mu^p); \mu^{1-p}/\lambda]$ has mean μ and variance $\mu^{1+p}\lambda$.
- $p = 1$ is usual gamma
- $p = 0$ gives variance proportional to mean
- $p = -1$ gives constant variance
- Skewness = 2CV in each case
- p can be fit by MLE to get best relationship of variance to mean

Same Thing on Lognormal

Standard definition

- $F(x;\mu,\sigma) = N\left(\frac{\ln(x)-\mu}{\sigma}\right)$
- EX = $e^{\mu+\sigma^2/2}$, variance = $e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right)$

Reparameterize

- $F(x;m,s) = N\left(\frac{\ln\left((x/m)\sqrt{1+s^2/m^2}\right)}{\sqrt{\ln\left(1+s^2/m^2\right)}}\right)$
- $\mu \to \ln\left(\frac{m^2}{\sqrt{s^2+m^2}}\right)$ $\sigma^2 \to \ln\left(\frac{s^2+m^2}{m^2}\right)$
- EX = m, variance = s^2

Add in p

- $F(x;m,s) = N\left(\frac{\ln\left((x/m)\sqrt{1+s^2/m^{p-2}}\right)}{\sqrt{\ln\left(1+s^2/m^{p-2}\right)}}\right)$
- EX = m, variance = s^2m^p

Aggregate Distributions – Poisson-Gamma Case

In general

- Combining a gamma severity in α and θ that has mean $\alpha\theta$ and variance $\alpha\theta^2$ with a Poisson in λ
- Aggregate mean is $\mu = \lambda\alpha\theta$ and variance is $\mu\theta(\alpha+1)$

Tweedie

- In exponential family
- Replace 3 parameters by 3 others: μ, ψ and p:
- $\lambda = \mu^{2-p}/[(2 - p)\psi]$; $\alpha = (2 - p)/(p - 1)$; $\theta = \psi(p - 1)m^{p-1}$
- Then $EZ = \mu$ and variance $= \psi\mu^p$, with $1 < p < 2$
- Not obvious, but this forces the frequency and severity means to move together:
- $\lambda = k(\alpha\theta)^{\alpha}$, with $k = [(2 - p)\psi]^{1/(1-p)}$
- General Poisson-gamma can be used instead, or other special cases constructed

MLE for Poisson-Gamma

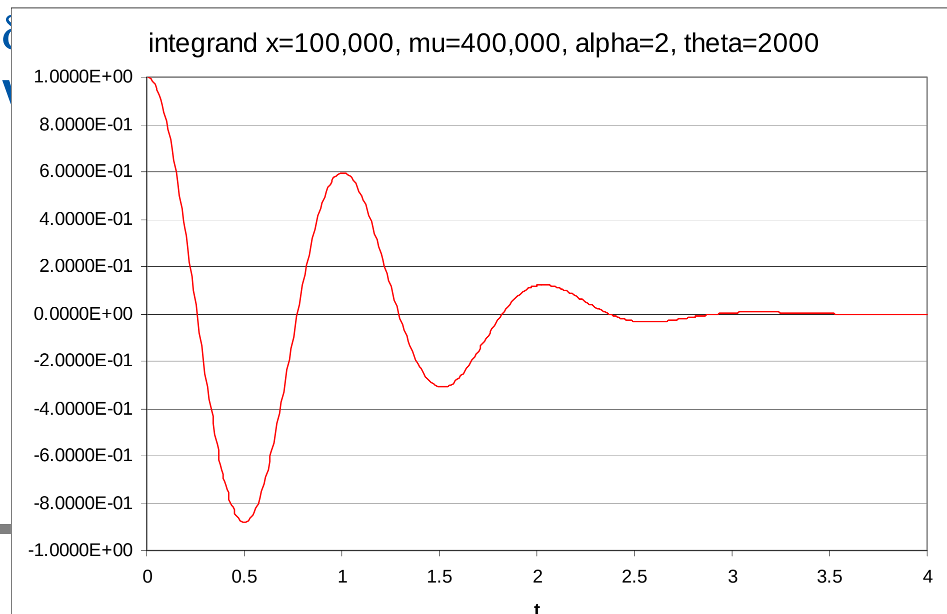
Calculation
—»

- **Often only aggregate observations available**
- **MLE would require aggregate density**
- **Worked out in Mong's 1980 CAS call paper as a numerical integration, inverting characteristic function $\varphi(t) = \exp[-1+\lambda(1 - it\theta)^{-\alpha}]$**

$$f(x) = \frac{1}{\sigma\pi}\int_0^\infty e^{\lambda j(t)} \cos\left[\frac{xt}{\sigma} - \lambda k(t)\right] dt$$

- **where $j(t) = \delta(t)\cos[\rho(t)]-1$; $k(t) = \delta(t)\sin[\rho(t)]$**
- **δ**

Integrand
—»



Other Aggregate Distributions

Poisson – Normal —»

- **Can have negative observations and even mean**
- **Density similar integral to Poisson-gamma**
- **Different ways to make mean a parameter**
- **Does not require constant variance**

Geometric - exponential —»

- **Closed form aggregate density**
- **Point mass at zero – which is often useful**

Poisson – Constant Severity

In exponential family —»

- But only if every aggregate is an integral multiple of the severity
- Common use comparing to chain-ladder often violates this
- Called ODP, but probability = 0 except at kb

Continuous analogue —»

- Zero-modified continuous scaled Poisson, or ZMCSP
- Mack’s 2002 book discusses
- Can use like ODP but also estimate b by MLE
- Point mass at zero is slightly less than ODP’s
- So ODP is actually approximation to ZMCSP

Example – Fitting Long-Haul Trucking Triangle

Models Fit

- Factors for 1st 5 diagonals
- Additive effects for several diagonals
- Additive constant term
- Standard regression assumes normal distribution
- Also fit ZMCSP and gamma p

	lag0	lag1	lag2	lag3	lag4	diag3	4+-10	const	θ,λ,σ	-lnL	AICc/2
ZMCPS	1.618	0.508	0.223	0.103	0.026	-2072	107.1	487.9	306.1	637.8	646.9
Gamma p	1.624	0.504	0.217	0.102	0.027	-1922	132.0	499.8	3,969.0	630.3	642.0
Normal	1.601	0.499	0.211	0.102	0.021	-1832	801.6	527.8	1,387.7	662.2	671.2

Results

- Gamma p best fit, normal worst by AICc
- Parameters very similar across models
- p in gamma p is 0.29 so variance proportional to $\mu^{0.71}$
- Surprising that power is below 1
- Residual plots better for gamma p as well

Example – Fitting Taylor-Ashe Triangle

Models Fit

—»

Model	Moment 37,184	MLE 30,892
Parameter Var	1,103,569,529,544	916,846,252,340
Process Var	718,924,545,072	597,282,959,722
Total Var	1,822,494,074,616	1,514,129,212,061

- 3 accident year factors, 2 lag factors, 1 diagonal
- Fit ODP but model not appropriate as cells not multiple of constant severity
  - Still gives estimates of factors and severity
  - Severity fit by moments method
- Also fit ZMCSP – constant multiple not needed
  - MLE for severity also
  - Fairly different than moments estimate of severity
  - Results in lower estimated variance
- Variance proportional to mean
- Same holds for Poisson-normal
- Gamma p had variance proportional to $\mu^{0.864}$
- All 3 had very similar AICs

Parameter	U_0	U_7	U_a	g_a	g_b	c
Estimate	3,810,000	7,113,775	5,151,180	0.067875	0.173958	0.198533
se 37,184	372,849	698,091	220,508	0.003431	0.005641	0.056896
se 30,892	339,846	636,298	200,989	0.003127	0.005142	0.051860