Loss Cost Modeling: Tweedie vs Quasi-Poisson

BACE Spring 2019

Josh Brady, FCAS, MAAA, PhD

The Cincinnati Insurance Company

CAS Antitrust Notice

- The Casualty Actuarial Society is committed to adhering strictly to the letter and spirit of the antitrust laws. Seminars conducted under the auspices of the CAS are designed solely to provide a forum for the expression of various points of view on topics described in the programs or agendas for such meetings.

- Under no circumstances shall CAS seminars be used as a means for competing companies or firms to reach any understanding – expressed or implied – that restricts competition or in any way impairs the ability of members to exercise independent business judgment regarding matters affecting competition.

- It is the responsibility of all seminar participants to be aware of antitrust regulations, to prevent any written or verbal discussions that appear to violate these laws, and to adhere in every respect to the CAS antitrust compliance policy.

Overview

What is quasi-Poisson?

- Same variance relationship as Poisson
- Defined on all non-negative values instead of just integers

Why use quasi-Poisson in practice?

- Testing in R and SAS shows quasi-Poisson models fit faster
- Predictions are balanced for categorical variables
- Simplifies the offset process

Advantages of Tweedie

- My experience on loss cost data:
  - Tweedie appears to be a more appropriate model based on diagnostics
  - Tweedie has slightly better predictive power
  - Although the data are rarely Tweedie distributed

In practice I almost always use quasi-Poisson over Tweedie for GLM modeling

Framework

- Modeling pure premium directly as opposed to a frequency/severity model
- We desire an interpretable model with multiplicative rating structure
  - Implies GLMs with a log link
- Performance being comparable, we prefer faster fitting models
- Primary goal is predictive power (while maintaining interpretability)
  - How to measure? Gini coefficient, lift, other aggregate diagnostics

Auto Insurance Pure Premium Modeling Example

dataCar*

- This data set is based on one-year vehicle insurance policies taken out in 2004 or 2005.
- 67856 observations
- Frequency ~ 15.5%
- Severity ~ $1900

Fields used in sample model

- claimcst0 - loss
- Exposure
- Pure premium (pp) = claimcst0/exposure
- veh_value in $10,000s
- veh_body
    - Categorical with levels BUS CONVT COUPE HBACK HDTOP MCARA MIBUS PANVN RDSTR SEDAN STNWG TRUCK UTE
- gender
    - categorical with levels F M
- agecat
    - 1 (youngest), 2, 3, 4, 5, 6

Model

Data Adjustments

- veh_body_grp2 = grouped small exposure levels with other levels
- veh_val5 = vehicle value rounded to nearest 0.1 and capped at 5
- Transform agecat to factor to model as categorical variable

Target = Pure Premium (pp) = Claimcst0/exposure

Weight = exposure (EE)

Formula:

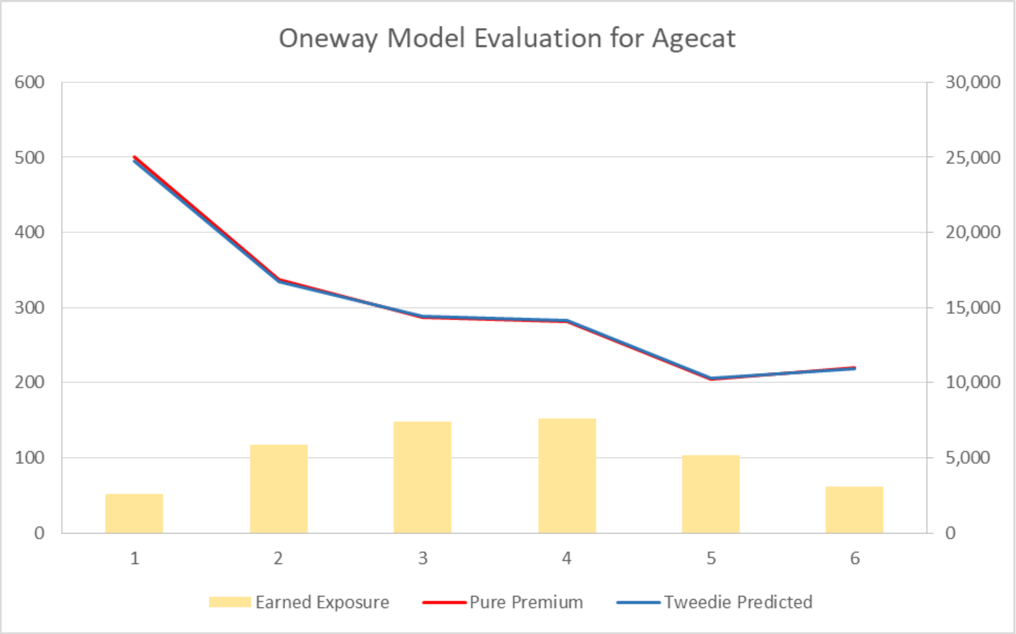
$$pp \sim agecat + gender + veh\_body\_gp2 + veh\_val5$$

*Contained in insuranceData R package https://cran.r-project.org/web/packages/insuranceData/index.html

Initial Motivations

- Fitted GLM using a Tweedie distribution (p=1.5)
- Noticed for categorical/factor covariates that the predicted values did not match the actual pure premiums
- For example on our test dataset the predictions do not match on the grouped age variable (agecat)

agecat	Earned Exposure	Pure Premium	Tweedie Predicted	Difference%
1	2,612	500	495	-1.2%
2	5,892	337	335	-0.5%
3	7,409	288	289	0.3%
4	7,617	282	283	0.6%
5	5,171	205	206	0.5%
6	3,100	221	219	-0.6%
Total	3,100	293	293	0.0%



These mismatches occur even on large datasets with credible data. The cause is due to the model specification.

Why is the Tweedie GLM not balanced for categorical predictors?
First a review of GLM theory…

- A generalized linear model with a distribution in the natural exponential family is parameterized by the following:
  - $E[y] = h(\eta)$, where h is the inverse link and $\eta = \boldsymbol{X}\beta$ is the linear predictor
  - Log likelihood: $$\log f(y|\theta, \phi, w) = \frac{w}{\phi}\left(y\theta - b(\theta)\right) + c\left(y, \frac{\phi}{w}\right)$$
    - θ and $b(\theta)$ are functions of the mean μ
    - ϕ is the dispersion parameter with weight w
- Solve for parameters β by maximizing the log likelihood

$$\frac{\partial}{\partial \beta_j} \sum_i \log f(y_i|\theta_i, \phi, w_i) = \frac{\partial}{\partial \beta_j} \sum_i \left(\frac{w_i}{\phi}\left(y_i\theta_i - b(\theta_i)\right) + c\left(y_i, \frac{\phi}{w_i}\right)\right)$$

$$= 0$$

GLM Review
-Variance Function

- Distributions in the natural exponential family are determined by the variance function $v(\mu)$
- $V[y] = \frac{\phi}{w} v(\mu)$
- Examples of interest
  - Poisson $v(\mu) = \mu$
  - Tweedie $v(\mu) = \mu^p, 1 < p < 2$
  - Gamma $v(\mu) = \mu^2$

GLM Review
-Deviance

- Deviance at observation y with estimated mean μ (actually $\hat{\mu}$)

$$d(y, \mu) = 2\left(\log f(y|\theta_s, \phi, w) - \log f(y|\theta, \phi, w)\right)$$

  - Where θ_s is the saturated model with a parameter at every observation
  - $\theta = \theta(\mu)$ for the current model
- Using the variance function $v(\mu)$, the deviance can be written in an elegant and convenient form

$$d(y, \mu) = \frac{2w}{\phi} \int_{\mu}^{y} \frac{y - t}{v(t)} dt$$

- Maximizing the log likelihood is equivalent to minimizing the deviance
- The total deviance is the sum over all observations

$$D(\mathbf{y}, \boldsymbol{\mu}) = \sum_i \frac{2w_i}{\phi} \int_{\mu_i}^{y_i} \frac{y_i - t}{v(t)} dt$$

GLM Review
-Estimating Equations Simplified

- Using the variance function and definition of deviance the estimating equations take on a simplified form
- Derivation relies on chain rule:

$$d(y_i, \mu_i) = \frac{2w_i}{\phi} \int_{\mu_i}^{y_i} \frac{y_i - t}{v(t)} \mathrm{d}t$$

$$\frac{\partial}{\partial \beta_j} d(y_i, \mu_i) = \frac{\partial d(y_i, \mu_i)}{\partial \mu_i} \frac{\mathrm{d}\mu_i}{\mathrm{d}\eta_i} \frac{\partial \eta_i}{\partial \beta_j}$$
$$= \left(\frac{2w_i}{\phi} \frac{\mu_i - y_i}{v(\mu_i)} \right) h'(\eta_i) X_{ij}$$

(Notice the reversal for y_i and μ_i)

Resulting estimating Equations:

$$\frac{\partial}{\partial \beta_j} D(\mathbf{y}, \boldsymbol{\mu}) = \frac{\partial}{\partial \beta_j} \sum_i d(y_i, \mu_i)$$
$$= \sum_i \frac{2w_i}{\phi} \frac{\mu_i - y_i}{v(\mu_i)} h'(\eta_i) X_{ij}$$
$$= 0$$

Estimating Equations
-Log link and variance power function

Log link

- $\mu = h(\eta) = \exp(\eta)$
- $h'(\eta) = \exp(\eta) = \mu$

Variance power function

- $v(\mu) = \mu^p, 1 \leq p \leq 2$

Apply to derive simplified estimating equations:

$$\begin{aligned}
\frac{\partial}{\partial \beta_j} D(\mathbf{y}, \boldsymbol{\mu}) &= \sum_i \frac{2w_i}{\phi} \frac{\mu_i - y_i}{v(\mu_i)} h'(\eta_i) X_{ij} \\
&= \sum_i \frac{2w_i}{\phi} \frac{\mu_i - y_i}{\mu_i^p} \mu_i X_{ij} \\
&= \sum_i \frac{2w_i}{\phi} (\mu_i - y_i)\, \mu_i^{1-p} X_{ij} \\
&= 0
\end{aligned}$$

Balance Equations

- Separate y_i and μ_i to opposite sides of the equation

$$\sum_i \underbrace{w_i y_i}_{\text{loss}} \mu_i^{1-p} X_{ij} = \sum_i \underbrace{w_i \mu_i}_{\text{[pred loss]}} \mu_i^{1-p} X_{ij}$$

$$\sum_i \text{loss}_i \mu_i^{1-p} X_{ij} = \sum_i [\text{pred loss}]_i \mu_i^{1-p} X_{ij}$$

Bias equations for categorical variables

Sample Design Matrix

Intercept with agecat

Intercept	agecat2	agecat3	agecat4	agecat5	agecat6
1	1	0	0	0	0
1	1	0	0	0	0
1	1	0	0	0	0
1	1	0	0	0	0
1	1	0	0	0	0
1	1	0	0	0	0
1	1	0	0	0	0
1	1	0	0	0	0
1	0	1	0	0	0
1	0	1	0	0	0
1	0	1	0	0	0
1	0	1	0	0	0
1	0	1	0	0	0
1	0	0	1	0	0
1	0	0	1	0	0
1	0	0	1	0	0
1	0	0	1	0	0
1	0	0	1	0	0
1	0	0	1	0	0
1	0	0	1	0	0
1	0	0	1	0	0

Balance equations for a categorical level

Consider the balance equation for categorical level j, say agecat2

$$X_{ij} = \begin{cases} 1, & \text{agecat} = 2 \\ 0, & \text{agecat} \neq 2 \end{cases}$$

Balance equations for level j reduce to

$$\sum_i \text{loss}_i \mu_i^{1-p} X_{ij} = \sum_i [\text{pred loss}]_i \mu_i^{1-p} X_{ij}$$

$$\sum_{\text{agecat}=2} \text{loss}_i \mu_i^{1-p} = \sum_{\text{agecat}=2} [\text{pred loss}]_i \mu_i^{1-p}$$

When would predicted losses equal actual losses? That is,

$$\sum_{\text{agecat}=2} \text{loss}_i = \sum_{\text{agecat}=2} [\text{pred loss}]_i$$

For $p = 1$ the balance equations imply predictions are balanced to losses.

When $p \neq 1$, then in general predictions are not balanced to losses.

Canonical Connection: Are there other cases where predictions are balanced to losses?

Yes, for the canonical link

- A canonical link satisfies $b(\theta) = \theta$ in the log likelihood

$$\log f(y|\theta, \phi, w) = \frac{w}{\phi}\big(y\theta - b(\theta)\big) + c\left(y, \frac{\phi}{w}\right)$$

- Importantly for our formulation the canonical link implies $h'(\eta) = v(\mu)$

- Recall the estimating equations, the canonical link satisfies that predictions are balanced to losses on categorical variables

$$\frac{\partial}{\partial \beta_j} D(\mathbf{y}, \boldsymbol{\mu}) = \sum_i \frac{2w_i}{\phi} \frac{\mu_i - y_i}{v(\mu_i)} h'(\eta_i) X_{ij} = 0$$

- Canonical link for Tweedie:

$$g(\mu) = \frac{\mu^{1-p}}{1-p} \quad \text{or} \quad h(\eta) = g^{-1}(\eta) = \big((1-p)\eta\big)^{\frac{1}{1-p}}$$

- Why not use the canonical link for Tweedie?
    - Not multiplicative and numerical stability issues

Balance Equations -Poisson Distribution

- We saw from the balance equations that $p = 1$ results in predictions being balanced to losses
  - Corresponds to the Poisson distribution
- There is a fundamental issue to using the Poisson distribution for loss cost modeling
  - The Poisson distribution is not defined for non-integers
  - Poisson probability:

$$f(y|\mu) = \frac{e^{\mu}\mu^{y}}{y!}$$

Quasi-Poisson - Finally!

- Recall deviance:

$$d(y, \mu) = \frac{2w}{\phi} \int_{\mu}^{y} \frac{y - t}{v(t)} \mathrm{d}t$$

- Notice that this formula for deviance only requires
  - Variance function $v(\mu)$
  - Link $\mu = h(\eta)$
  - Design matrix and coefficients $\eta = \boldsymbol{X}\beta$
- Nothing here requires a probability distribution
- Quasi-likelihood is defined through the above deviance as (notice the limits are switched)

$$q(y, \mu) = -\frac{1}{2} d(y, \mu) = \frac{w}{\phi} \int_{y}^{\mu} \frac{y - t}{v(t)} \mathrm{d}t$$

- Using the quasi-likelihood we can extend to all non-negative values
- With $v(\mu) = \mu$ the quasi-Poisson deviance is given by

$$d(y, \mu) = \frac{2w}{\phi} \left(y \log \frac{y}{\mu} + \mu - y \right)$$

Balance example on the test data set

agecat	Pure Premium	Total Prediction quasi-Poisson	Total Prediction Tweedie (p=1.5)
1	500.5	500.5	494.6
2	336.9	336.9	335.1
3	287.8	287.8	288.7
4	281.7	281.7	283.4
5	205.3	205.3	206.3
6	220.5	220.5	219.2

veh_body_gp2	Pure Premium	Total Prediction quasi-Poisson	Total Prediction Tweedie (p=1.5)
HBACK	309.3	309.3	308.4
UTE	283.6	283.6	283.7
STNWG	309.4	309.4	304.6
VAN	336.9	336.9	350.2
SEDAN	256.7	256.7	259.1
TRUCK	378.6	378.6	384.4

gender	Pure Premium	Total Prediction quasi-Poisson	Total Prediction Tweedie (p=1.5)
F	273.4	273.4	274.9
M	318.2	318.2	315.9

Within each row green indicates higher values.
Quasi-Poisson total predictions are balanced to actual pure premiums.

Can we extend the Poisson distribution to all non-negative values?

Answer: No

- Pmf for the Poisson distribution with mean μ:
  - $f(k|\mu) = e^{-\mu} \frac{k^{\mu}}{k!}, k \in \{0,1,2, \dots\}$
- Suppose we wanted to extend the Poisson distribution from integers to all non-negative numbers in a way such that the parameter estimates were unchanged
- That is, can we replace $k!$ with a (reasonably nice) function g such that for $y > 0$

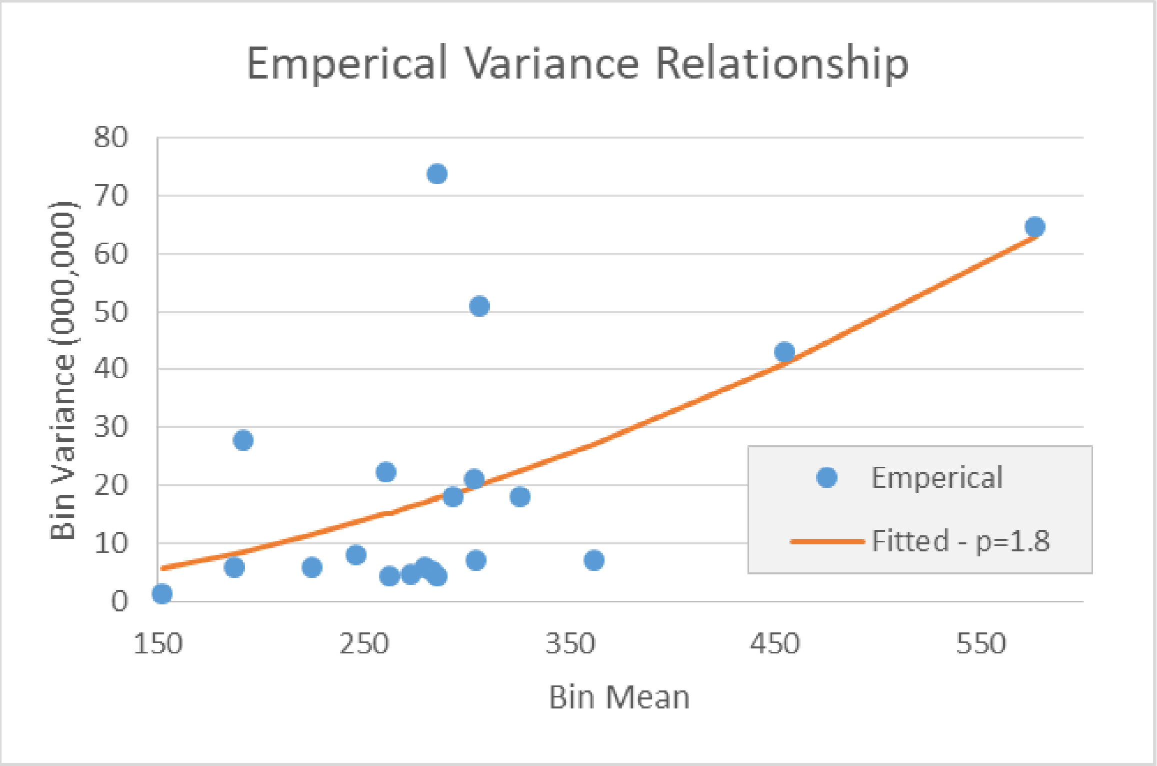
  $f_2(y|\mu) = e^{-\mu} \frac{y^{\mu}}{g(y)}$ is a probability distribution?
- Natural candidate would be $\mathrm{g(y)} = \Gamma(y + 1)$
- Turns out that it is not possible
  - Why?
  - For the curious… Proof relies on showing the moment generating functions are equal on an open domain. Result is to conclude distributions are in fact the same.

Is our data more quasi-Poisson or Tweedie distributed?

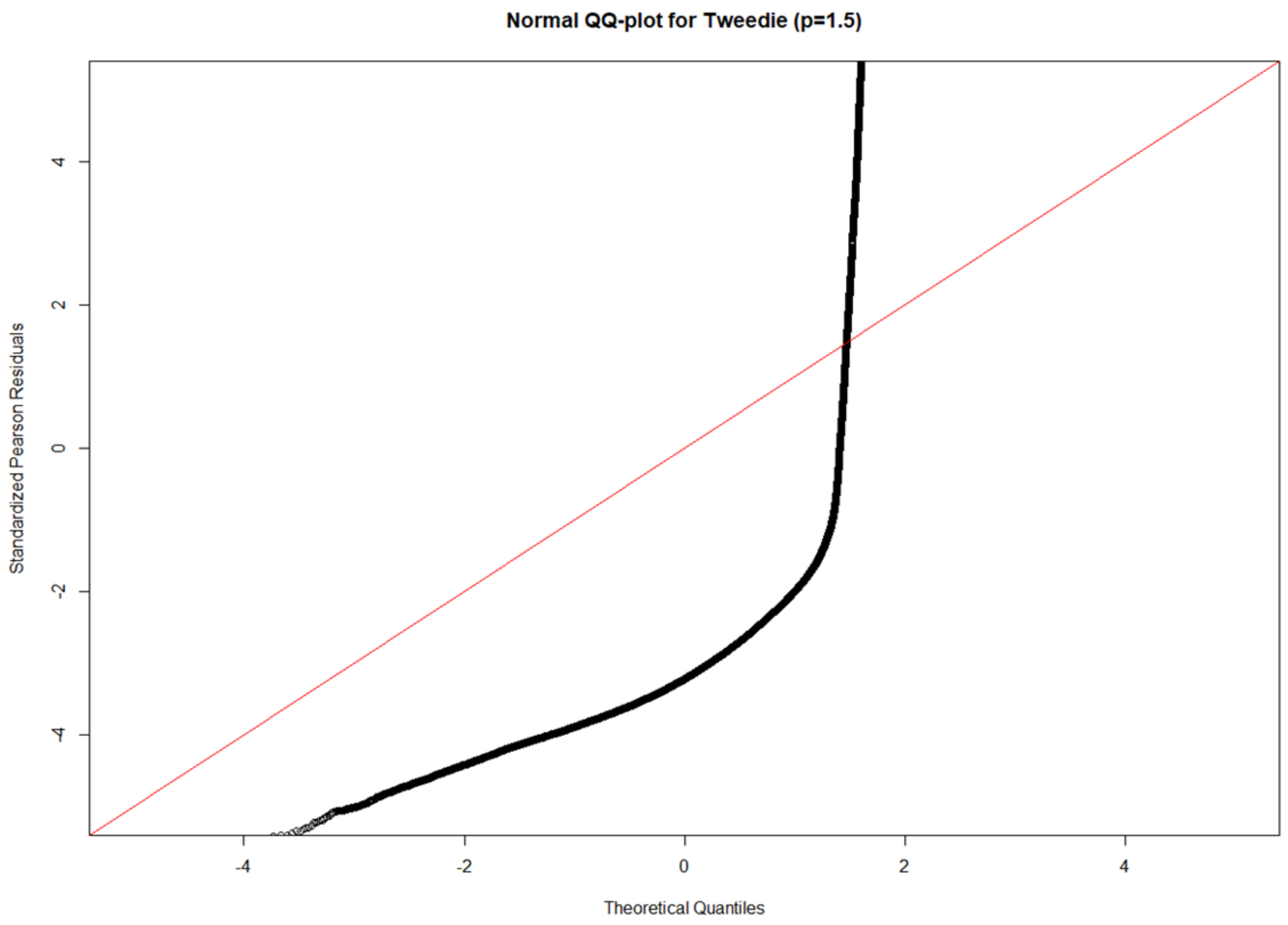
- Our data is almost certainly neither
- Check variance relationship
- Check qq-plot

Examining the variance relationship of the sample data

- Fit Tweedie (p=1.5) model for predicted mean
- Ranked low to high and binned into 20 equal exposure bins
- On each bin calculated empirical mean and variance
- Plot to right shows the relationship
- Fit curve of form $\mathrm{Var}_{\mathrm{bin}} = a\,\mu_{bin}^{p}$
  - a is the intercept
  - p is the fitted power (1.8 in this case)
- Suggests that a Tweedie variance relationship is more appropriate than quasi-Poisson (linear)



QQ-plot



- qq-plots can be misleading for discrete responses
- With varied means and weights a qq-plot is more appropriate (asymptotically)
- qq-plot for quasi-Poisson looks much worse

Computing coefficient standard errors for quasi-likelihood

- Similar to likelihood based GLMs we can calculate the covariance matrix of the coefficients
- The coefficients are distributed (asymptotically) as

$$\hat{\beta} \sim N(\beta, \phi(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1})$$

- Where the diagonal "weight" matrix is defined as

$$W_i = w_i \frac{\left(h'(\eta_i)\right)^2}{v(\mu_i)}$$

How to calculated dispersion?

- We can't rely on an MLE estimation
- We can use the deviance or Pearson estimator
- Pearson (r is the number of parameters)

$$\phi = \frac{1}{n-r} \sum_i \frac{(y_i - \mu_i)^2}{v(\mu_i)}$$

Parameter Comparison

- Dark green represents more extreme values within a row
- %Diff shows %difference in factors between Tweedie (p=1.5) and quasi-Poisson
  - $\%\text{Diff} = \frac{\exp(\beta_{quasi-Poisson}) - \exp(\beta_{Tweedie})}{\exp(\beta_{Tweedie})}$

Coefficients			**Tweedie**		
	Poisson	**p=1.2**	**p=1.5**	**p=1.7**	**%Diff**
(Intercept)	6.085	6.085	6.084	6.082	
agecat1	(base level factor of 0)				
agecat2	-0.384	-0.380	-0.376	-0.373	-0.8%
agecat3	-0.538	-0.531	-0.522	-0.517	-1.6%
agecat4	-0.558	-0.551	-0.540	-0.534	-1.8%
agecat5	-0.879	-0.872	-0.862	-0.857	-1.7%
agecat6	-0.792	-0.790	-0.788	-0.786	-0.5%
genderM	0.163	0.157	0.150	0.146	1.3%
veh_body_gp2HBACK	(base level factor of 0)				
veh_body_gp2SEDAN	-0.148	-0.142	-0.135	-0.130	-1.3%
veh_body_gp2STNWG	-0.108	-0.110	-0.112	-0.113	0.3%
veh_body_gp2TRUCK	0.026	0.039	0.057	0.068	-3.0%
veh_body_gp2UTE	-0.232	-0.226	-0.219	-0.215	-1.3%
veh_body_gp2VAN	0.019	0.038	0.066	0.084	-4.5%
veh_val5	0.072	0.069	0.066	0.064	0.6%

p-value			**Tweedie**	
	Poisson	**p=1.2**	**p=1.5**	**p=1.7**
(Intercept)	0%	0%	0%	0%
agecat1	NA	NA	NA	NA
agecat2	5%	7%	9%	11%
agecat3	1%	1%	2%	2%
agecat4	0%	1%	1%	2%
agecat5	0%	0%	0%	0%
agecat6	0%	0%	0%	0%
genderM	18%	20%	22%	24%
veh_body_gp2HBACK	NA	NA	NA	NA
veh_body_gp2SEDAN	32%	34%	37%	39%
veh_body_gp2STNWG	54%	54%	54%	54%
veh_body_gp2TRUCK	94%	91%	88%	86%
veh_body_gp2UTE	37%	38%	40%	41%
veh_body_gp2VAN	96%	92%	86%	82%
veh_val5	24%	27%	30%	32%

Tweedie log likelihood

Due to the infinite series in the likelihood, computation is computationally intensive.

See Dunn and Smyth (2008) for details.

$$\log f(y|\mu, \phi, p) = \frac{1}{\phi}\left(y\frac{\mu^{1-p}}{1-p} - \frac{\mu^{2-p}}{2-p}\right) + \log a(y, \phi, p)$$

$$a(y, \phi, p) = \begin{cases} \frac{1}{y}\sum_{t=1}^{\infty} \frac{y^{t\alpha}}{(p-1)^{t\alpha}\phi^{t(1+\alpha)}(2-p)^{t}t!\Gamma(t\alpha)}, & y > 0 \\ 1, & y > 0 \end{cases}$$

Where

$$\alpha = \frac{p-2}{p-1}$$

Algorithm Speed – R glm

Model	Relative Time
quasi-Poisson	1
Tweedie (p = 1.01)	1.84
Tweedie (p = 1.2)	1.53
Tweedie (p = 1.5)	1.35
Tweedie (p = 1.7)	2.28

➢ Overall Tweedie appears to be at least 30% slower than quasi-Poisson <u>on the test data set</u>

- R uses the Pearson estimator for dispersion
  - r is the number of parameters

$$\phi = \frac{1}{n-r} \sum_i \frac{(y_i - \mu_i)^2}{v(\mu_i)}$$

- Likelihood is not computed as it is computationally expensive (can use tweedie package to compute)

Algorithm Speed – SAS hpgenselect

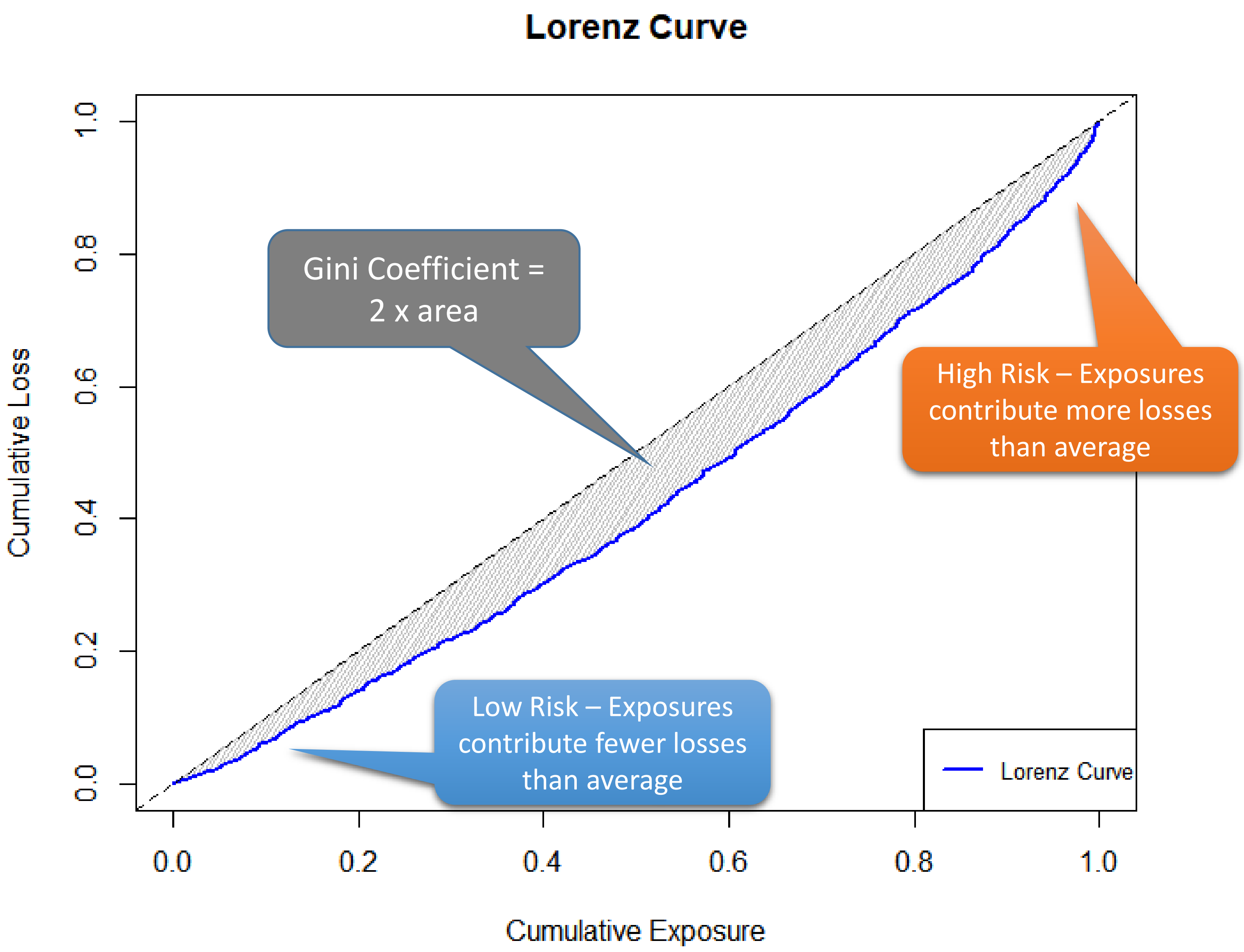
SAS hpgenselect has the option to use maximum likelihood to estimate the dispersion ϕ and power p for the Tweedie distribution

quasi-Tweedie is an option to avoid MLE estimation of ϕ (similar to R)

Model	Relative Time
quasi-Poisson	1
Tweedie (p and dispersion MLE estimated)	61.7
Tweedie with dispersion est. (p=1.5)	21.7
quasi-Tweedie with dispersion est. (p=1.5)	10.0
quasi-Tweedie without dispersion est. (p=1.5)	6.4

Gini Coefficient

- The Gini coefficient is a measure of how well the predictions rank the losses
- Rank low to high by the predictions
- Lorenz curve is the plot of cumulative losses vs cumulative exposure
- The Gini coefficient is twice the area between the 45 degree line and the Lorenz Curve



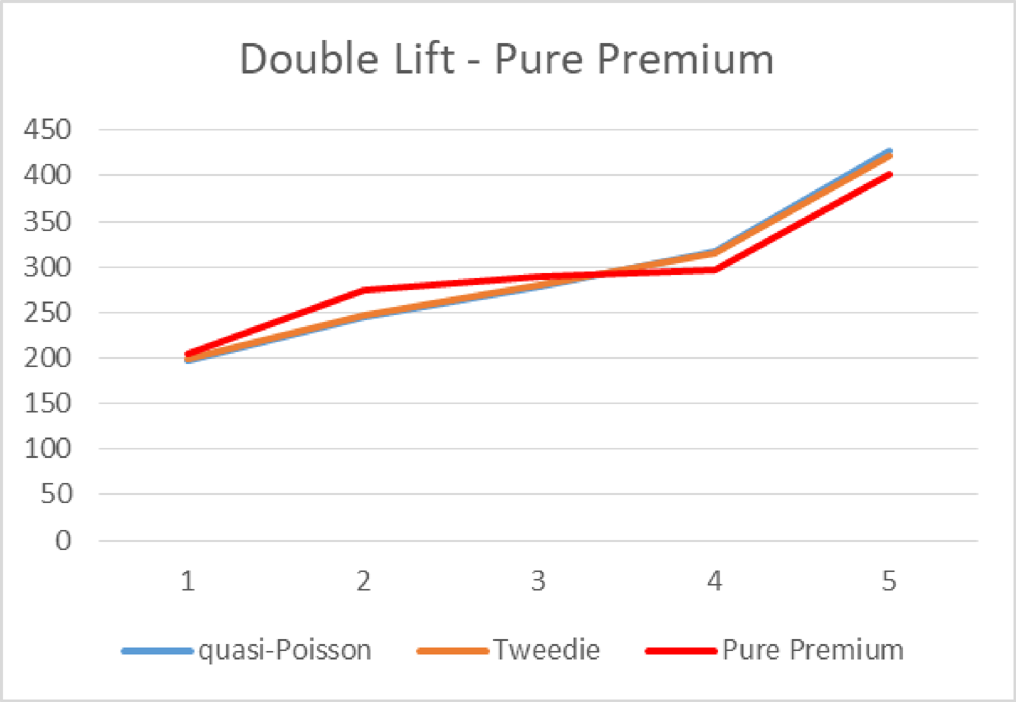
Performance - Gini

- Fit model and evaluate on entire dataset (in-sample testing)
- Cross-Val (10 fold, 5 times)
- Result
  - ➢Tweedie has slightly better in-sample and cross validated performance
  - ➢Quasi-Poisson shows the smallest drop in performance

Model	Full Dataset	Cross-Val	Difference
quasi-Poisson	0.14114	0.12367	0.0175
Tweedie (p = 1.2)	0.14130	0.12374	0.0176
Tweedie (p = 1.5)	0.14173	0.12374	0.0180
Tweedie (p = 1.7)	0.14205	0.12374	0.0183

Performance – Double Lift

- Ranked low to high by Tweedie(p=1.5)/quasi-Poisson
- Group into 5 bins
  - Left chart shows actual vs predicted pure premiums
  - Right chart shows percentage error in each bin



Do many statistical tests extend to quasi-likelihood?

Yes! (Sort of)

- t-statistics
- Log likelihood ratio tests
- Chi-squared tests
- AIC
  - $\text{AIC} = -2\text{loglik} + 2p$
  - These measures allow comparisons between different distributions
  - Unfortunately due to the direct presence of the log likelihood in the formulas one cannot compare AIC on absolute terms
  - Using the deviance we can compute the change in AIC
    - Allows for a step-wise AIC
- Of course we still have traditional methods
  - Consistency testing
  - Residual and prediction plots
  - Cross validation
  - Bootstrap

Many results are only true asymptotically with weaker convergence as compared to likelihood based GLMs

Factor Offset

- Suppose in our model we wanted to apply fixed factors
- These could be in the current model or perhaps previously selected/known factors
- With loss cost data two of the most common ways to apply factors offsets are:
  - Let F denote multiplicative factors to offset
  - Model offset: $\frac{\text{loss}}{\text{EE}} \sim \eta + \log(F)$, weight = EE
  - Exposure offset: $\frac{\text{loss}}{\text{EE} * \text{F}} \sim \eta$, weight = $\text{EE} * \text{F}$
- For the Tweedie model the exposure offset is not equivalent to the model offset

Offset Equivalence

From Shi (2009):

- u_i - offset factor
- e_{ij} - exposure
- c_{ij} - claim counts
- L_{ij} - loss

Model	Distribution	Response Variable	Weight Variable
Frequency	Poisson	Weighted sum of adjusted claim frequency, $(\sum_i c_{ij})/(\sum_i e_{ij}u_i)$	Total number of adjusted exposures, $\sum_i e_{ij}u_i$
Severity	Gamma	Weighted sum of adjusted claim severity, $(\sum_i L_{ij}/u_i)/(\sum_i c_{ij})$	Total number of claims, $\sum_i c_{ij}$
Loss cost	Tweedie(p)	Weighted sum of adjusted loss amount $(\sum_i L_{ij}u_i^{1-p})/(\sum_i e_{ij}u_i^{2-p})$	Total number of adjusted exposures, $\sum_i e_{ij}u_i^{2-p}$

Offset Example – Tweedie

Based on the full model we offset the veh_body_gp2 fitted factors and refit the model.

Full Model

	Coefficients
(Intercept)	6.084
agecat1	base
agecat2	-0.376
agecat3	-0.522
agecat4	-0.540
agecat5	-0.862
agecat6	-0.788
genderM	0.150
veh_body_gp2HBACK	base
veh_body_gp2SEDAN	-0.135
veh_body_gp2STNWG	-0.112
veh_body_gp2TRUCK	0.057
veh_body_gp2UTE	-0.219
veh_body_gp2VAN	0.066
veh_val5	0.066

Model Offset

	Coefficients
(Intercept)	6.084
agecat1	base
agecat2	-0.376
agecat3	-0.522
agecat4	-0.540
agecat5	-0.862
agecat6	-0.787
genderM	0.150
veh_body_gp2HBACK	NA
veh_body_gp2SEDAN	NA
veh_body_gp2STNWG	NA
veh_body_gp2TRUCK	NA
veh_body_gp2UTE	NA
veh_body_gp2VAN	NA
veh_val5	0.066

Exposure Offset

	Coefficients
(Intercept)	6.055
agecat1	base
agecat2	-0.395
agecat3	-0.531
agecat4	-0.564
agecat5	-0.882
agecat6	-0.810
genderM	0.139
veh_body_gp2HBACK	NA
veh_body_gp2SEDAN	NA
veh_body_gp2STNWG	NA
veh_body_gp2TRUCK	NA
veh_body_gp2UTE	NA
veh_body_gp2VAN	NA
veh_val5	0.052

For a quasi-Poisson model an exposure offset is equivalent to a model offset.
Performing an exposure offset then Tweedie modeling may result in a mismatch!

Conclusion

Either quasi-Poisson or Tweedie can be a reasonable choice for modeling loss cost

On the test data set:

- Quasi-Poisson
  - Fits faster
  - Predictions are balanced to losses for categorical variables
  - Exposure offset is equivalent to model offset
- Tweedie
  - Variance structure can be more appropriate
  - Better cross validated performance

References

Dunn, Peter K., and Gordon K. Smyth. "Evaluation of Tweedie exponential dispersion model densities by Fourier inversion." *Statistics and Computing* 18.1 (2008): 73-86.

Jorgensen, Bent. *The theory of dispersion models*. CRC Press, 1997.

McCullagh, Peter. *Generalized linear models*. Routledge, 2018.

Mildenhall, Stephen J. "A systematic relationship between minimum bias and generalized linear models." *Proceedings of the Casualty Actuarial Society*. Vol. 86. No. 164. 1999.

Ohlsson, Esbjörn, and Björn Johansson. *Non-life insurance pricing with generalized linear models*. Vol. 2. Berlin: Springer, 2010.

Shi, Sheng G. "Direct analysis of pre-adjusted loss cost, frequency or severity in tweedie models." *Casualty Actuarial Society E-Forum*. 2010.