On the Importance of Dispersion Modeling for Claims Reserving: Application of the Double GLM Theory

Danaïl Davidov
under the supervision of
Jean-Philippe Boucher

Département de mathématiques
Université du Québec à Montréal

August 2009

Table of Contents

Introduction

Definition

- A loss reserve is a provision for an insurer's liability for claims

Notes

Introduction

Definition

- A loss reserve is a provision for an insurer's liability for claims
- A stochastic model uses random variables in a regression framework

Notes

Introduction

Definition

- A loss reserve is a provision for an insurer's liability for claims
- A stochastic model uses random variables in a regression framework

Notes

- Claims reserves models presented here use GLM theory as introduced in *England, Verrall 2002*

Model Definition

Notations

- $C_{i,j}$ Incremental payments

Model Definition

Notations

- $C_{i,j}$ Incremental payments
- $w_{i,j}$ Exposure

Model Definition

Notations

- $C_{i,j}$ Incremental payments
- $w_{i,j}$ Exposure
- $r_{i,j}$ Incremental number of payments

Model Definition

Notations

- $C_{i,j}$ Incremental payments
- $w_{i,j}$ Exposure
- $r_{i,j}$ Incremental number of payments
- $Y_{i,j}$ Normalized incremental payments

Model Definition

Notations

- $C_{i,j}$ Incremental payments
- $w_{i,j}$ Exposure
- $r_{i,j}$ Incremental number of payments
- $Y_{i,j}$ Normalized incremental payments
- $Y_{i,j} = \frac{C_{i,j}}{w_{i,j}}$

Model Definition

Hypotheses

- $C_{i,j}$ is a compound Poisson-Gamma distribution

Model Definition

Hypotheses

- $C_{i,j}$ is a compound Poisson-Gamma distribution
- Frequency $\sim$ Poisson with mean $\vartheta_{i,j} w_{i,j}$ and variance $\vartheta_{i,j} w_{i,j}$

Model Definition

Hypotheses

- $C_{i,j}$ is a compound Poisson-Gamma distribution
- Frequency $\sim$ Poisson with mean $\vartheta_{i,j} w_{i,j}$ and variance $\vartheta_{i,j} w_{i,j}$
- Severity $\sim$ Gamma with mean $\tau_{i,j}$ and variance $\nu \tau_{i,j}^2$

Model Definition

Hypotheses

- $C_{i,j}$ is a compound Poisson-Gamma distribution
- Frequency $\sim$ Poisson with mean $\vartheta_{i,j} w_{i,j}$ and variance $\vartheta_{i,j} w_{i,j}$
- Severity $\sim$ Gamma with mean $\tau_{i,j}$ and variance $\nu\tau_{i,j}^2$
- Using the following parametrisation

$$\begin{aligned} p &= \frac{\nu+2}{\nu+1} \ , \quad p \in (1,2) \\ \mu &= \vartheta\tau \\ \phi &= \frac{\vartheta^{1-p}\tau^{2-p}}{(2-p)} \end{aligned}$$

Model Definition

Tweedie Model

- $Y_{i,j} \sim Tweedie(\mu_{i,j}, p, \phi, w_{i,j})$

Model Definition

Tweedie Model

- $Y_{i,j} \sim Tweedie(\mu_{i,j}, p, \phi, w_{i,j})$
- $\mu_{i,j} = e^{X\beta}$

Model Definition

Tweedie Model

- $Y_{i,j} \sim Tweedie(\mu_{i,j}, p, \phi, w_{i,j})$
- $\mu_{i,j} = e^{X\beta}$
- $E[Y_{i,j}] = \mu_{i,j}$, $Var[Y_{i,j}] = \frac{\phi}{w_{i,j}} \mu_{i,j}^{p}$

Log-likelihood function

Tweedie Model

$$
\begin{aligned}
l \;=\; & \sum_{i,j} r_{i,j} \log \left(\frac{(w_{i,j}/\phi)^{\nu+1} y_{i,j}^{\nu}}{(p-1)^{\nu}(2-p)} \right) \\
& - \log \left(r_{i,j}! \Gamma \left(r_{i,j}\nu \right) y_{i,j} \right) + \frac{w_{i,j}}{\phi} \left(y_{i,j} \frac{\mu_{i,j}^{1-p}}{1-p} - \frac{\mu_{i,j}^{2-p}}{2-p} \right)
\end{aligned}
$$

Parameters

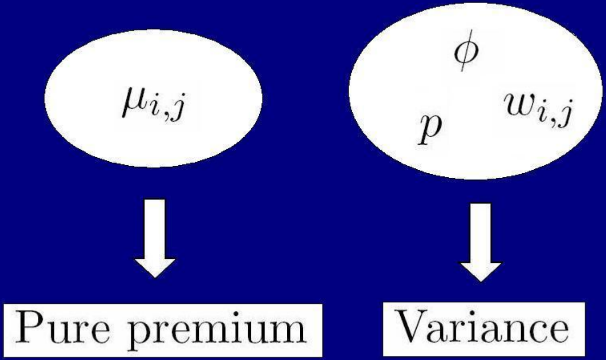

FIGURE: Parameter main influence

Parameter p

Tweedie Model

- Can be estimated only when the number of payments is known

Parameter p

Tweedie Model

- Can be estimated only when the number of payments is known
- Otherwise, it's supposed fixed and known

Parameter p

Tweedie Model

- Can be estimated only when the number of payments is known
- Otherwise, it's supposed fixed and known
- p and ϕ need both to be estimated at the same time

Parameter ϕ

Optimizing ϕ using the likelihood principle

$$\widehat{\phi_p} = \frac{-\sum_{i,j} w_{i,j} \left(y_{i,j} \frac{\mu_{i,j}^{1-p}}{1-p} - \frac{\mu_{i,j}^{2-p}}{2-p} \right)}{(1+\nu) \sum_{i,j} r_{i,j}}$$

Parameter ϕ

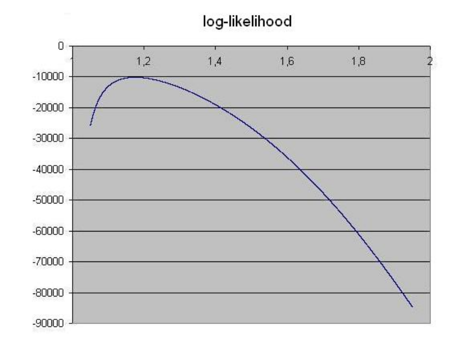

FIGURE: Optimizing p using the likelihood principle for ϕ

Parameter ϕ

Optimizing ϕ using the deviance principle

$$\widehat{\phi_p} = \sum_{i,j} \frac{2}{N-Q} \left(y_{i,j} \frac{y_{i,j}^{1-p} - \mu_{i,j}^{1-p}}{1-p} - \frac{y_{i,j}^{2-p} - \mu_{i,j}^{2-p}}{2-p} \right)$$

Parameter ϕ

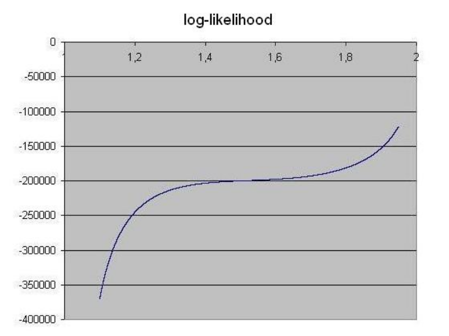

FIGURE: Optimizing p using the deviance principle for ϕ

Parameter $w_{i,j}$

Note

- The exposure has been incorporated in the begining, within the initial hypothesis

Parameter $w_{i,j}$

Note

- The exposure has been incorporated in the begining, within the initial hypothesis
- Different ways of incorporating the exposure within the initial hypothesis would lead to different models

Frequency vs Severity

Compound Poisson Model

$$Y = \sum_{k=1}^{N} X_i$$

	Case 1	Case 2	Case 3
E[N]	10	20	10
Var[N]	10	20	10
E[X]	10	10	20
Var[X]	100	100	400
E[Y]	100	200	200
Var[Y]	2000	4000	8000

TABLE: Mean and variance of total costs for various situations

Frequency vs Severity

Typical situation in a long-tail business

- Decreasing average frequency

Frequency vs Severity

Typical situation in a long-tail business

- Decreasing average frequency
- Increasing average severity

Frequency vs Severity

Typical situation in a long-tail business

- Decreasing average frequency
- Increasing average severity
- Increasing variance in the severity

Impact of the Distribution

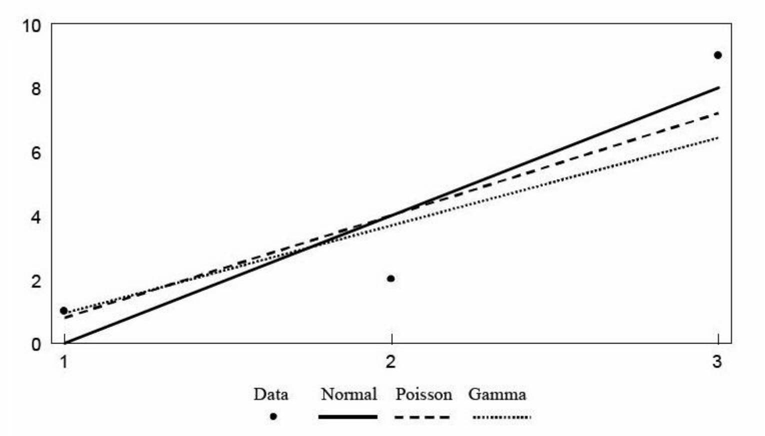

FIGURE: Fitted curve for Normal, Poisson and Gamma models

Dispersion Models

Model Definition

- $Y_{i,j} \sim Tweedie(\mu_{i,j}, p, \phi_{i,j}, w_{i,j})$

Dispersion Models

Model Definition

- $Y_{i,j} \sim Tweedie(\mu_{i,j}, p, \phi_{i,j}, w_{i,j})$
- $\mu_{i,j} = e^{X\beta}$

Dispersion Models

Model Definition

- $Y_{i,j} \sim Tweedie(\mu_{i,j}, p, \phi_{i,j}, w_{i,j})$
- $\mu_{i,j} = e^{X\beta}$
- $E[Y_{i,j}] = \mu_{i,j}\ ,\ \ Var[Y_{i,j}] = \phi_{i,j}\mu_{i,j}^{p}$

Dispersion Models

Model Definition

- $Y_{i,j} \sim Tweedie(\mu_{i,j}, p, \phi_{i,j}, w_{i,j})$
- $\mu_{i,j} = e^{X\beta}$
- $E[Y_{i,j}] = \mu_{i,j}$, $Var[Y_{i,j}] = \phi_{i,j}\mu_{i,j}^{p}$
- $\phi_{i,j} = e^{V\gamma}$

Dispersion Models

Log-likelihood

$$l = \sum_{i,j} r_{i,j} \log \left(\frac{(w_{i,j}/\phi_{i,j})^{\nu+1} y_{i,j}^{\nu}}{(p-1)^{\nu}(2-p)} \right) - \log \left(r_{i,j}! \Gamma \left(r_{i,j}\nu \right) y_{i,j} \right)$$
$$+ \frac{w_{i,j}}{\phi_{i,j}} \left(y_{i,j} \frac{\mu_{i,j}^{1-p}}{1-p} - \frac{\mu_{i,j}^{2-p}}{2-p} \right)$$

Dispersion Models

Notes

- The p parameter is optimized at the same time as all other parameters using implicitly the likelihood principle

Dispersion Models

Notes

- The p parameter is optimized at the same time as all other parameters using implicitly the likelihood principle
- Accident years do not have a significant impact on the dispersion parameter

Dispersion Models

Notes

- The p parameter is optimized at the same time as all other parameters using implicitly the likelihood principle
- Accident years do not have a significant impact on the dispersion parameter
- Due to lack of data in the last column, the model was build so that the last two columns have the same dispersion parameter

Dispersion Models

Notes

- The p parameter is optimized at the same time as all other parameters using implicitly the likelihood principle
- Accident years do not have a significant impact on the dispersion parameter
- Due to lack of data in the last column, the model was build so that the last two columns have the same dispersion parameter
- Possibility to incorporate trends in the dispersion parameter by using the Hoerl's curve parametrisation

Double GLMs

Algorithm

1. Start with the initial exposure and find the normalized incremental payments

Double GLMs

Algorithm

1. Start with the initial exposure and find the normalized incremental payments
2. Find the deviance between fitted and observed values

Double GLMs

Algorithm

1. Start with the initial exposure and find the normalized incremental payments
2. Find the deviance between fitted and observed values
3. Model the deviance

Double GLMs

Algorithm

1. Start with the initial exposure and find the normalized incremental payments
2. Find the deviance between fitted and observed values
3. Model the deviance
4. Establish the new exposure and start over

Double GLMs

Component	Mean modeling	Variance modeling
Target variable	Y	d
Mean	μ	ϕ
Variance	$\phi V(\mu)$	$2\phi^2$
Link	$\eta = g(\mu)$	$\eta_d = g_d(\phi)$
Linear predictor	$\eta = \mathbf{X}\beta$	$\eta_d = \mathbf{V}\gamma$
Deviance	$2\int_\lambda^y \frac{y-\mu}{V(\mu)} d\mu$	$2(-\log(d/\phi) + (d-\phi)/\phi$
Exposure	w/ϕ	1

FIGURE: Inter-relationship between the two sub-models

Iterative Weighted Least Squares

Algorithm additional specifications

- Analogous to Fisher's weighted scoring method for optimization

Iterative Weighted Least Squares

Algorithm additional specifications

- Analogous to Fisher's weighted scoring method for optimization
- IWLS implicitly uses the deviance principle for estimating ϕ

Restricted Maximum Likelihood

Notes

- Maximum likelihood estimators are biased downwards when the number of estimators is large compared to the number of observations

Restricted Maximum Likelihood

Notes

- Maximum likelihood estimators are biased downwards when the number of estimators is large compared to the number of observations
- REML produces estimators which are approximately and sometimes exactly unbiased

Restricted Maximum Likelihood

Notes

- Maximum likelihood estimators are biased downwards when the number of estimators is large compared to the number of observations
- REML produces estimators which are approximately and sometimes exactly unbiased
- Approximately maximizes the penalized log-likelihood

$$l*_p(y;\gamma;p) = l(y;\beta_\gamma;\gamma;p) + \frac{1}{2}\log\left|X^TWX\right|$$

Optimizing p

Algorithm

1. Suppose p fixed and known

Optimizing p

Algorithm

1. Suppose p fixed and known
2. Evaluate all the other parameters using the DGLM

Optimizing p

Algorithm

1. Suppose p fixed and known
2. Evaluate all the other parameters using the DGLM
3. Evaluate the penalized log-likelihood

Optimizing p

Algorithm

1. Suppose p fixed and known
2. Evaluate all the other parameters using the DGLM
3. Evaluate the penalized log-likelihood
4. Start over with different values for p and compare the penalized log-likelihood

Reserve Variability

Mean Square Error of Prediction

- Dispersion Models : overdispersion included implicitly in the parameter covariance matrix

Reserve Variability

Mean Square Error of Prediction

- Dispersion Models : overdispersion included implicitly in the parameter covariance matrix
- GLMs : overdispersion is included manually in the parameter covariance matrix

Data Analyzed

Incremental Payments

AY	1	2	3	4	5	6	7	8	9	10	11
1	17 841 110	7 442 433	895 413	407 744	207 130	61 569	15 978	24 924	1 236	15 643	321
2	19 519 117	6 656 520	941 458	155 395	69 458	37 769	53 832	111 391	42 263	25 833	
3	19 991 172	6 327 483	1 100 177	279 649	162 654	70 000	56 878	9 881	19 656		
4	19 305 646	5 889 791	793 020	309 042	145 921	97 465	27 523	61 920			
5	18 291 478	5 793 282	689 444	288 626	345 524	110 585	115 843				
6	18 832 520	5 741 214	581 798	248 563	106 875	94 212					
7	17 152 710	5 908 286	524 806	230 456	346 904						
8	16 615 059	5 111 177	553 277	252 877							
9	16 835 453	5 001 897	489 356								

Swiss Motor Industry (Wütrich 2003)

Data Analyzed

Exposure

w_i
112 953
110 364
105 400
102 067
99 124
101 460
94 753
92 326
89 545

Total Volume per Accident Year i

Data Analyzed

Number of Payments

AY	1	2	3	4	5	6	7	8	9	10	11
1	6 229	3 500	425	134	51	24	13	12	6	4	1
2	6 395	3 342	402	108	31	14	12	5	6	5	
3	6 406	2 940	401	98	42	18	5	3	3		
4	6 148	2 898	301	92	41	23	12	10			
5	5 952	2 699	304	94	49	22	7				
6	5 924	2 692	300	91	32	23					
7	5 545	2 754	292	77	35						
8	5 520	2 459	267	81							
9	5 390	2 224	223								

Optimizing p

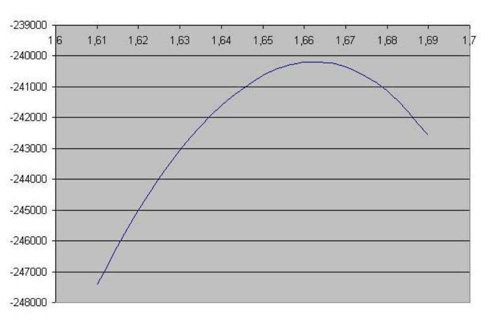

FIGURE: Restricted log-likelihood for various p in a DGLM

Analysis of Results

		GLM 1 parameter $p = 1.1741$		Dispersion 9 parameters $p = 1.8111$		DGLM 9 parameters $p = 1.6625$	
AY	$w_{i,j}$	R_i	$MSEP^{1/2}$	R_i	$MSEP^{1/2}$	R_i	$MSEP^{1/2}$
1	112 953	-	-	-	-	-	-
2	110 364	326	2 638	324	736	291	2 809
3	105 400	21 565	26 804	21 352	27 769	19 586	32 040
4	102 067	40 716	35 556	40 185	34 658	37 191	41 717
5	99 124	89 298	53 297	87 224	54 103	83 592	68 788
6	101 460	138 335	66 052	138 203	62 649	131 053	80 811
7	94 753	204 262	80 906	202 469	69 443	206 128	91 283
8	92 326	360 484	111 999	359 148	82 510	343 778	137 516
9	89 545	597 056	150 003	596 119	93 407	560 207	151 951
Total	-	1 452 042	271 886	1 445 024	234 818	1 381 827	334 269

Analysis of Results

	GLM $p = 1.1741$			Dispersion 9 parameters $p = 1.8111$			DGLM 9 parameters $p = 1.6625$		
AY	Estimation	Process	$EQMP^{1/2}$	Estimation	Process	$EQMP^{1/2}$	Estimation	Process	$EQMP^{1/2}$
1	-	-	-	-	-	-	-	-	-
2	1 869	1 861	2 638	546	493	736	2 728	668	2 809
3	15 601	21 795	26 804	16 984	21 970	27 769	23 141	22 161	32 040
4	19 144	29 962	35 556	20 001	28 304	34 658	28 436	30 524	41 717
5	25 976	46 538	53 297	28 127	46 217	54 103	39 518	56 304	68 788
6	30 564	58 556	66 052	32 881	53 327	62 649	44 424	67 505	80 811
7	35 230	72 833	80 906	34 783	60 103	69 443	48 145	77 554	91 283
8	45 664	102 268	111 999	40 846	71 690	82 510	56 744	125 263	137 516
9	61 307	136 903	150 003	47 079	80 675	93 407	61 431	138 980	151 951
Total	180 126	203 658	271 886	183 345	146 711	234 818	248 122	223 989	334 269

Analysis of Results

# Obs.	Principle	1	2	3	4	5	6	7	8	9	10	11
0	Deviance	3 921	3 921	3 921	3 921	3 921	3 921	3 921	3 921	3 921	3 921	3 921
0	1 parameter	199	199	199	199	199	199	199	199	199	199	199
0	9 parameters	202	169	259	384	716	648	1 003	1 269	901	904	904
7	Deviance	1 857	1 857	1 857	1 857	1 857	1 857	1 857	1 857	1 857	1 857	1 857
7	1 parameter	198	198	198	198	198	198	198	198	198	198	198
7	9 parameters	202	166	250	384	637	648	800	964	898	903	903
14	Deviance	1 117	1 117	1 117	1 117	1 117	1 117	1 117	1 117	1 117	1 117	1 117
14	1 parameter	198	198	198	198	198	198	198	198	198	198	198
14	9 parameters	204	166	250	385	637	648	922	969	880	904	904

FIGURE: Adjusting the most deviant observations has a bigger influence on the deviance principle

Analysis of Results

Cell		Principle	
i	j	Maximum likelihood	Deviance
1	1	14,19	149,56
2	1	15,28	0,77
3	1	15,58	0,02
4	1	15,14	8,32
5	1	14,49	3,31
6	1	14,84	30,41
7	1	13,73	4,93
8	1	13,38	11,30
9	1	13,52	31,57
1	2	6,90	267,52
2	2	6,28	11,70
3	2	6,03	10,56
⋮	⋮	⋮	⋮
Total		199,28	3 921,00

FIGURE: Contribution of each cell to the dispersion. $p = 1.1741$, $w_{i,j} \equiv 1$

Conclusion

Further Discussion

- Lack of observations and abundance of parameters is a hostile environment for DGLMs

Conclusion

Further Discussion

- Lack of observations and abundance of parameters is a hostile environment for DGLMs
- The deviance cannot be estimated in the last column for a DGLM and is hence ignored when estimating ϕ

Conclusion

Further Discussion

- Lack of observations and abundance of parameters is a hostile environment for DGLMs
- The deviance cannot be estimated in the last column for a DGLM and is hence ignored when estimating ϕ
- The algorithm does not converge for p fixed when there are more than 7 parameters

Conclusion

Further Discussion

- Lack of observations and abundance of parameters is a hostile environment for DGLMs
- The deviance cannot be estimated in the last column for a DGLM and is hence ignored when estimating ϕ
- The algorithm does not converge for *p* fixed when there are more than 7 parameters
- The algorithm does converge for *p* fixed for one parameter, but *p* cannot be optimized

Conclusion

Further Discussion

- Lack of observations and abundance of parameters is a hostile environment for DGLMs
- The deviance cannot be estimated in the last column for a DGLM and is hence ignored when estimating ϕ
- The algorithm does not converge for *p* fixed when there are more than 7 parameters
- The algorithm does converge for *p* fixed for one parameter, but *p* cannot be optimized
- The "bounds" of convergence need to be explored further

References

England, PD and Verrall, RJ (2002).
Stochastic Claims Reserving in General Insurance.
Institute of Actuaries and Faculty of Actuaries

Smyth, G. and Jorgensen, B. (2002).
Fitting Tweedie's Compound Poisson Model to Insurance Claims Data : Dispersion Modelling.
ASTIN Bulletin, 1, 143–157

Wüthrich, MV (2003).
Claims Reserving Using Tweedie's Compound Poisson Model.
ASTIN Bulletin, 331–346

The end