

CLAIMS RESERVING USING TWEEDIE'S COMPOUND POISSON MODEL

BY

MARIO V. WÜTHRICH

ABSTRACT

We consider the problem of claims reserving and estimating run-off triangles. We generalize the gamma cell distributions model which leads to Tweedie's compound Poisson model. Choosing a suitable parametrization, we estimate the parameters of our model within the framework of generalized linear models (see Jørgensen-de Souza [2] and Smyth-Jørgensen [8]). We show that these methods lead to reasonable estimates of the outstanding loss liabilities.

KEYWORDS

Claims Reserving, Run-off Triangles, IBNR, Compound Poisson Model, Exponential Family, GLM, MSEP.

INTRODUCTION

Claims reserving and IBNR estimates are classical problems in insurance mathematics. Recently Jørgensen-de Souza [2] and Smyth-Jørgensen [8] have fitted Tweedie's compound Poisson model to insurance claims data for tarification. Using the connection between tarification and claims reserving analysis (see Mack [3]), we translate the fitting procedure to our run-off problem. Our model should be viewed within the context of stochastic methods for claims reserving. For excellent overviews on this topic we refer to England-Verrall [1] and Taylor [9].

The starting point of this work was the gamma cell distributions model presented in Section 7.5 of Taylor [9]. The gamma cell distributions model assumes that every cell of the run-off triangle consists of r_{ij} independent payments which are gamma distributed with mean τ_{ij} and shape parameter γ . These assumptions enable the calculation of convoluted distributions of incremental payments. Unfortunately, this model does not allow one to estimate e.g. the mean square error of prediction (MSEP), since one has not enough information. We assume that the number of payments r_{ij} are realisations of random variables R_{ij} , i.e. the number of payments R_{ij} and the size of the individual payments $X_{ij}^{(k)}$ are both modelled stochastically. This can be done assuming that

R_{ij} is Poisson distributed. These assumptions lead to *Tweedie's compound Poisson model* (see e.g. Jørgensen-de Souza [2]). Choosing a clever parametrization for Tweedie's compound Poisson model, we see that the model belongs to the exponential dispersion family with variance function $V(\mu) = \mu^p$, $p \in (1, 2)$, and dispersion ϕ . It is then straightforward to use generalized linear model (GLM) methods for parameter estimations. A significant first step into that direction has been done by Wright [11].

In this work we study a version of Tweedie's compound Poisson model with constant dispersion ϕ (see Subsection 4.1). This model should be viewed within the context of the over-dispersed Poisson model (see Renshaw-Verrall [6] or England-Verrall [1], Section 2.3) and the Gamma model (see Mack [3] and England-Verrall [1], Section 3.3): The over-dispersed Poisson model and the Gamma model correspond to the two extreme cases $p = 1$ and $p = 2$, resp. Our extension closes continuously the gap between these two models, since $p \in (1, 2)$. To estimate p we additionally use the information r_{ij} which is not used in the parameter estimations for $p = 1$ and $p = 2$. Though we have one additional parameter, we obtain in general better estimates since we also use more information and have more degrees of freedom.

Moreover, our parametrization is such that the variance parameters p and ϕ are orthogonal to the mean parameter. This leads to a) efficient parameter estimations (fast convergence), b) good estimates of MSEP.

At the end of this article we demonstrate the method using motor insurance datas. Our results are compared to several different classical methods. Of course, in practice it would not be wise to trust in just these methods. It should be pointed out that the methods presented here are all payment based. Usually it is also interesting to compare payment based results to results which rely on total claims incurred datas (for an overview we refer to Taylor [9] and the references therein).

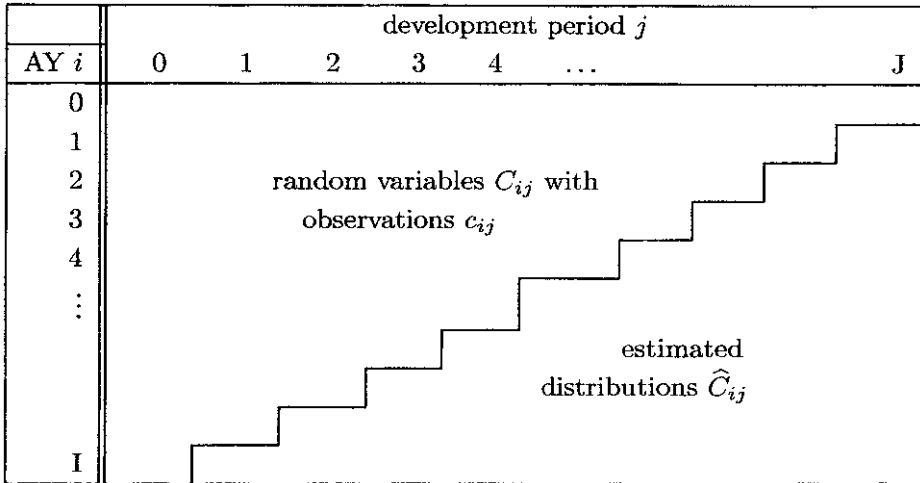
In the next section we define the model. In Section 3 we recall the definition of Tweedie's compound Poisson model. In Section 4 we apply Tweedie's compound Poisson model to our run-off problem. In Section 5 we give an estimation procedure for the mean square error of prediction (MSEP). Finally, in Section 6 we give the examples.

2. DEFINITION OF THE MODEL

We use the following (well-known) structure for the run-off patterns: the accident years are denoted by $i \leq I$ and the development periods are denoted by $j \leq J$. We are interested in the random variables C_{ij} . C_{ij} denote the incremental payments for claims with origin in accident year i during development period j . Usually one has observations c_{ij} of C_{ij} for $i + j \leq I$ and one tries to complete (estimate) the triangle for $i + j > I$. The following illustration may be helpful.

Definition of the model:

1. The number of payments R_{ij} are independent and Poisson distributed with parameter $\lambda_{ij} w_i > 0$. The weight $w_i > 0$ is an appropriate measure for the volume.



2. The individual payments $X_{ij}^{(k)}$ are independent and gamma distributed with mean $\tau_{ij} > 0$ and shape parameter $\gamma > 0$.
3. R_{ij} and $X_{mn}^{(k)}$ are independent for all indices. We define the incremental payments paid in cell (i, j) as follows

$$C_{ij} = 1_{\{R_{ij} > 0\}} \cdot \sum_{k=1}^{R_{ij}} X_{ij}^{(k)} \quad \text{and} \quad Y_{ij} = C_{ij} / W_i. \tag{2.1}$$

Remarks:

- There are several different possibilities to choose appropriate weights w_i , e.g. the number of policies or the total number of claims incurred, etc. If one chooses the total number of claims incurred one needs first to estimate the number of IBNyR cases (cases incurred but not yet reported).
- Sometimes it is also convenient to define R_{ij} as the number of claims with origin in i which have at least one payment in period j .
- Y_{ij} denotes the normalized incremental payments in cell (i, j) .
- One easily sees that conditionally, given R_{ij} , the random variable C_{ij} is gamma distributed with mean $R_{ij}\tau_{ij}$ and shape parameter $R_{ij}\gamma$ (for $R_{ij} > 0$).

3. TWEEDIE'S COMPOUND POISSON MODEL

In this section we formulate our model in a reparametrized version, this has already been done in the tariffication problems of [2] and [8]. Therefore we try to keep this section as short as possible and give the main calculations in Appendix A.

For the moment we skip the indices i and j . The distribution Y (for given weight w) is parametrized by the three parameters λ , τ and γ . We now choose new parameters μ , ϕ and p such that the density of Y can be written as, $y \geq 0$, (see (A.2) below and formula (12) in [2])

$$f_Y(y; \mu, \phi / w, p) = c(y; \phi / w, p) \exp \left\{ \frac{w}{\phi} \left(y \frac{\mu^{1-p}}{1-p} - \frac{\mu^{2-p}}{2-p} \right) \right\}, \quad (3.1)$$

where $c(y; \phi / w, p)$ is given in Appendix A and

$$p = (\gamma + 2) / (\gamma + 1) \in (1, 2), \quad (3.2)$$

$$\mu = \lambda \cdot \tau, \quad (3.3)$$

$$\phi = \lambda^{1-p} \tau^{2-p} / (2-p). \quad (3.4)$$

If we set $\theta = \mu^{1-p} / (1-p)$ we see that the density of Y can be written as (see also [2], formula (12))

$$f_Y(y; \mu, \phi / w, p) = c(y; \phi / w, p) \exp \left\{ \frac{w}{\phi} (y\theta - \kappa_p(\theta)) \right\}, \quad (3.5)$$

with $\kappa_p(\theta) = \frac{1}{2-p} ((1-p)\theta)^{\frac{2-p}{1-p}}$.

Hence, the distribution of Y belongs to the exponential dispersion family with parameters μ , ϕ and $p \in (1, 2)$ (see e.g. McCullagh-Nelder [5], Section 2.2.2). We write for $p \in (1, 2)$

$$Y \sim \text{ED}^{(p)}(\mu, \phi / w). \quad (3.6)$$

For $Y \sim \text{ED}^{(p)}(\mu, \phi / w)$ we have (see [2] Section 2.2)

$$E[Y] = \kappa'_p(\theta) = \mu, \quad (3.7)$$

$$\text{Var}(Y) = \frac{\phi}{w} \cdot V(\mu) = \frac{\phi}{w} \mu^p. \quad (3.8)$$

ϕ is the so-called dispersion parameter and $V(\cdot)$ the variance function with $p \in (1, 2)$. For our claims reserving problem we consider the following situation:

Constant dispersion ϕ (see Subsection 4.1): $p \in (1, 2)$ and Y_{ij} are independent with

$$Y_{ij} \sim \text{ED}^{(p)}(\mu_{ij}, \phi / w_i) \Rightarrow E[Y_{ij}] = \mu_{ij} \text{ and } \text{Var}(Y_{ij}) = \frac{\phi}{w_i} \mu_{ij}^p. \quad (3.9)$$

Interpretation and Remarks:

- Tweedie [10] seems to be the first one to study the compound Poisson model with gamma severities from the point of view of exponential dispersion models. For this reason this model is known as Tweedie's compound Poisson model in the literature, see e.g. [8].
- $p = (\gamma + 2) / (\gamma + 1)$ is a function of γ (shape parameter of the single payments distributions $X_{ij}^{(k)}$). Hence the shape parameter γ determines the behaviour of the variance function $V(\mu) = \mu^p$. Furthermore we have chosen a parametrization (μ, ϕ, p) such that μ is *orthogonal* to (ϕ, p) in the sense that the Fisher information matrix is zero in the off-diagonal (see e.g. [2], page 76, or [8]). I.e. our parametrization focuses attention to variance parameters (ϕ, p) and a mean parameter μ which are orthogonal to each other. This orthogonality has many advantages to alternative parametrizations. E.g. we have efficient algorithms for parameter estimations which typically rapidly converge (see Smyth [7]). Moreover the estimated standard errors of μ , which are of most interest, do not require adjustments by the standard errors of the variance parameters, since these are orthogonal.
- Our model closes continuously the gap between the over-dispersed Poisson Model (see Renshaw-Verrall [6] or England-Verrall [1], Section 2.3) where we have a linear variance function ($p = 1$):

$$\text{Var}(Y_{ij}) = \phi / w_i \cdot \mu_{ij}, \quad (3.10)$$

and the Gamma model (see Mack [3] and England-Verrall [1], Section 3.3) where

$$\text{Var}(Y_{ij}) = \phi / w_i \cdot \mu_{ij}^2. \quad (3.11)$$

In our case p is estimated from the data using additionally the information r_{ij} (see (4.6)). The information r_{ij} is not used in the boundary cases $p = 1$ and $p = 2$.

- Naturally in our model we have $p \in (1, 2)$, since $\gamma > 0$. We estimate p from the data, so theoretically the estimated p could lie outside the interval $[1, 2]$ which would mean that none of our models fits to the problem (e.g. $p = 0$ implies normality, $p = 3$ implies the inverse Gaussian model). In all our claims reserving examples we have observed that the estimated p was lying strictly within $(1, 2)$.

4. APPLICATION OF TWEEDIE'S MODEL TO CLAIMS RESERVING

4.1. Constant dispersion parameter model

We assume that all the Y_{ij} are independent with $Y_{ij} \sim \text{ED}^{(p)}(\mu_{ij}, \phi / w_i)$, i.e. Y_{ij} belongs to the exponential dispersion family with $p \in (1, 2)$, and

$$E[Y_{ij}] = \mu_{ij} \quad \text{and} \quad \text{Var}(Y_{ij}) = \frac{\phi}{w_i} \mu_{ij}^p. \tag{4.1}$$

We use the notation $\boldsymbol{\mu} = (\mu_{00}, \dots, \mu_{IJ})'$. Given the observations $\{(r_{ij}, y_{ij}), i + j \leq I, \sum_{i,j} r_{ij} > 0\}$, the log-likelihood function for the parameters $(\boldsymbol{\mu}, \phi, p)$ is given by (see Appendix A and [2], Section 3)

$$L(\boldsymbol{\mu}, p, \phi) = \sum_{i,j} \left[r_{ij} \log \left\{ \frac{(w_i/\phi)^{\gamma+1} y_{ij}^\gamma}{(p-1)^\gamma (2-p)} \right\} - \log(r_{ij}! \Gamma(r_{ij}\gamma) y_{ij}) + \frac{w_i}{\phi} \left\{ y_{ij} \frac{\mu_{ij}^{1-p}}{1-p} - \frac{\mu_{ij}^{2-p}}{2-p} \right\} \right]. \tag{4.2}$$

Formula (4.2) immediately shows that given p the observations $y_{ij} = c_{ij}/w_i$ are sufficient for MLE estimation of μ_{ij} (one does not need r_{ij}). Moreover, for constant ϕ , the dispersion parameter has no influence on the estimation of $\boldsymbol{\mu}$.

Next we assume a multiplicative model (often called chain-ladder type structure): i.e. there exist parameters $\alpha(i)$ and $f(j)$ such that for all $i \leq I$ and $j \leq J$

$$\mu_{ij} = \alpha(i) \cdot f(j). \tag{4.3}$$

After suitable normalization, α can be interpreted as the expected ultimate claim in accident year i and f is the proportion paid in period j . It is now straightforward to choose the logarithmic link function

$$\eta_{ij} = \log(\mu_{ij}) = \mathbf{x}_{ij} \boldsymbol{\beta}, \tag{4.4}$$

where $\boldsymbol{\beta} = (\log \alpha(0), \dots, \log \alpha(I), \log f(0), \dots, \log f(J))'$ and $X = (\mathbf{x}_{00}, \dots, \mathbf{x}_{IJ})$ is the appropriate design matrix.

Parameter estimation:

- a) **For p known.** We deal with a generalized linear model (GLM) of the form (4.1)-(4.4). Hence we can use standard software packages for the estimation of $\boldsymbol{\mu}$.
- b) **For p unknown.** Usually p and γ , resp., are unknown. Henceforth we study the profile likelihood for γ (here we closely follow [2] Section 3.2): For $\boldsymbol{\mu}$ and p given, the MLE of ϕ is given by (see (4.2))

$$\hat{\phi}_p = \frac{-\sum_{i,j} w_i \left(y_{ij} \frac{\mu_{ij}^{1-p}}{1-p} - \frac{\mu_{ij}^{2-p}}{2-p} \right)}{(1+\gamma) \sum_{i,j} r_{ij}}. \tag{4.5}$$

From this we obtain the profile likelihood for p and γ , resp., ($\sum_{i,j} r_{ij} > 0$) as

$$\begin{aligned}
 L_{\mu}(p) = L(\boldsymbol{\mu}, p, \hat{\phi}_p) &= (1 + \gamma) \sum_{i,j} r_{ij} \left(\log \frac{w_i}{\hat{\phi}_p} - 1 \right) \\
 &+ \sum_{i,j} r_{ij} \log \left(\frac{1}{2-p} \left(\frac{y_{ij}}{p-1} \right)^{\gamma} \right) - \sum_{i,j} \log \Gamma(r_{ij} \gamma).
 \end{aligned}
 \tag{4.6}$$

Given $\boldsymbol{\mu}$, the parameter p is estimated maximizing (4.6).

- c) **Finally we combine a) and b).** The main advantage of our parametrization is (as already mentioned above) the orthogonality of $\boldsymbol{\mu}$ and (ϕ, p) . $\boldsymbol{\mu}$ can be estimated as if (ϕ, p) were known and vice versa. Alternating the updating procedures for $\boldsymbol{\mu}$ and (ϕ, p) leads to an efficient algorithm: Set initial value $p^{(0)}$ and estimate $\boldsymbol{\mu}^{(1)}$ via a). Then estimate $p^{(1)}$ from $\boldsymbol{\mu}^{(1)}$ via (4.6), and iterate this procedure. We have seen that typically one obtains very fast convergence of $(\boldsymbol{\mu}^{(k)}, p^{(k)})$ to some limit (for our examples below we needed only 4 iterations).

4.2. Dispersion modelling

So far we have always assumed that ϕ is constant over all cells (i, j) . If we consider the definitions (3.3) and (3.4) we see that every factor which increases λ increases the mean μ and decreases the dispersion ϕ because $p \in (1, 2)$. Increasing the average payment size τ increases both the mean and the dispersion. Changing λ and τ such that $\lambda^{1-p} \tau^{2-p}$ remains constant has only an effect on the mean μ . Hence it is necessary to model both the mean and the dispersion in order to get a fine structure, i.e. model μ_{ij} and ϕ_{ij} for each cell (i, j) individually and estimate p . Such a model has been studied in the context of tarification by Smyth-Jørgensen [8].

We do not further follow these ideas here since we have seen that in our situation such models are over-parametrized. Modelling the dispersion parameters while also trying to optimize the power of the variance function allows too many degrees of freedom: e.g. if we apply the dispersion modelling model to the data given in Example 6.1 one sees that p is blown up when allowing the dispersion parameters to be modelled too. It is even possible that there is no unique solution when modelling ϕ_{ij} and p at the same time (in all our examples we have observed rather slow convergence even when choosing “meaningful” initial values which indicates this problematic).

5. MEAN SQUARE ERROR OF PREDICTION

To estimate the mean square error of prediction (MSEP) we proceed as in England-Verrall [1]. Assume that the incremental payments C_{ij} are independent,

and \widehat{C}_{ij} are unbiased estimators depending only on the past (and hence are independent from C_{ij}). Assume $\widehat{\eta}_{ij}$ is the GLM estimate for $\eta_{ij} = \log \mu_{ij}$, then (see e.g. [1], (7.6)-(7.7))

$$\begin{aligned} \text{MSE}_{C_{ij}}(\widehat{C}_{ij}) &= E\left[\left(C_{ij} - \widehat{C}_{ij}\right)^2\right] = \text{Var}(C_{ij}) + \text{Var}(\widehat{C}_{ij}) \\ &\approx \phi w_i \cdot \mu_{ij}^p + (w_i \mu_{ij})^2 \text{Var}(\widehat{\eta}_{ij}). \end{aligned} \quad (5.1)$$

The last term is usually available from standard statistical software packages, all the other parameters have been estimated before. The first term in (5.1) denotes the process error, the last term the estimation error.

The estimation of the MSE for several cells (i, j) is more complicated since we obtain correlations from the estimates. We define Δ to be the unknown triangle in our run-off pattern. Define the total outstanding payments

$$C = \sum_{(i,j) \in \Delta} C_{ij} \quad \text{and} \quad \widehat{C} = \sum_{(i,j) \in \Delta} \widehat{C}_{ij}. \quad (5.2)$$

Then

$$\begin{aligned} \text{MSE}_C(\widehat{C}) &= E\left[(C - \widehat{C})^2\right] \approx \sum_{(i,j) \in \Delta} \phi_{ij} w_i \cdot \mu_{ij}^p + \sum_{(i,j) \in \Delta} (w_i \mu_{ij})^2 \text{Var}(\widehat{\eta}_{ij}) \\ &\quad + \sum_{\substack{(i_1, j_1), (i_2, j_2) \in \Delta, \\ (i_1, j_1) \neq (i_2, j_2)}} w_{i_1} \mu_{i_1 j_1} w_{i_2} \mu_{i_2 j_2} \text{Cov}(\widehat{\eta}_{i_1 j_1}, \widehat{\eta}_{i_2 j_2}). \end{aligned}$$

The evaluation of the last term needs some care: Usually one obtains a covariance matrix for the estimated GLM parameters $\log \alpha(i)$ and $\log f(j)$. This covariance matrix needs to be transformed into a covariance matrix for η with the help of the design matrices.

6. EXAMPLE

Example 6.1.

We consider Swiss Motor Insurance datas. We consider 9 accident years over a time horizon of 11 years. Since we want to analyze the different methods rather mechanically, this small part of the truth is already sufficient for drawing conclusions.

TABLE 6.2
OBSERVATIONS FOR THE NORMALIZED INCREMENTAL PAYMENTS $Y_{ij} = C_{ij}/w_i$.

y_{ij}	Development period j										
AY i	0	1	2	3	4	5	6	7	8	9	10
0	157.95	65.89	7.93	3.61	1.83	0.55	0.14	0.22	0.01	0.14	0.00
1	176.86	60.31	8.53	1.41	0.63	0.34	0.49	1.01	0.38	0.23	
2	189.67	60.03	10.44	2.65	1.54	0.66	0.54	0.09	0.19		
3	189.15	57.71	7.77	3.03	1.43	0.95	0.27	0.61			
4	184.53	58.44	6.96	2.91	3.46	1.12	1.17				
5	185.62	56.59	5.73	2.45	1.05	0.93					
6	181.03	62.35	5.54	2.43	3.66						
7	179.96	55.36	5.99	2.74							
8	188.01	55.86	5.46								

TABLE 6.3
NUMBER OF PAYMENTS R_{ij} AND VOLUME w_i .

r_{ij}	Development period j											
AY i	0	1	2	3	4	5	6	7	8	9	10	w_i
0	6'229	3'500	425	134	51	24	13	12	6	4	1	112'953
1	6'395	3'342	402	108	31	14	12	5	6	5		110'364
2	6'406	2'940	401	98	42	18	5	3	3			105'400
3	6'148	2'898	301	92	41	23	12	10				102'067
4	5'952	2'699	304	94	49	22	7					99'124
5	5'924	2'692	300	91	32	23						101'460
6	5'545	2'754	292	77	35							94'753
7	5'520	2'459	267	81								92'326
8	5'390	2'224	223									89'545

Remark: As weights w_i we take the number of reported claims (the number of IBNyR claims with reporting delay of more than two years is almost zero for this kind of business).

a) Tweedie's compound Poisson model with constant dispersion.

We assume that Y_{ij} are independent with $Y_{ij} \sim ED^{(p)}(\mu_{ij}, \phi/w_i)$ (see (4.1)). Define the total outstanding payments C as in (5.2). If we start with initial value $p^{(0)} = 1.5 \in (1, 2)$ and then proceed the estimation iteration as in Subsection 4.1, we observe that already after 4 iterations we have sufficiently converged to equilibrium (for the choice of p one should also have a look at Figure 1):

TABLE 6.4
ESTIMATION OF p .

Iteration k	0	1	2	3	4
$p^{(k)}$	1.5000	1.1743	1.1741	1.1741	1.1741
Outstanding payments $\hat{C}^{(k)}$		1'431'266	1'451'288	1'451'300	1'451'299

For $p = 1.1741$ the GLM output is as follows: Dispersion $\hat{\phi} = 29'281$ and parameter estimates:

TABLE 6.5
PARAMETERS α AND f FOR $p = 1.1741$.

j	0	1	2	3	4	5	6	7	8	9	10
$\widehat{\log \alpha(j)}$	-5.862	-5.825	-5.762	-5.782	-5.777	-5.819	-5.792	-5.837	-5.809		
$\widehat{\log f(j)}$	11.01	9.90	7.79	6.78	6.45	5.51	5.13	5.08	4.17	4.16	0.00

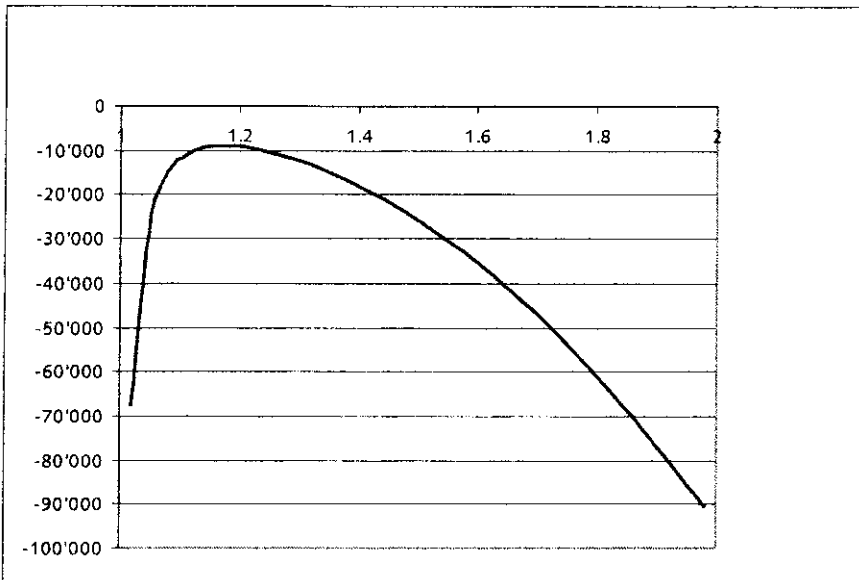


Figure 1: Profile likelihood function $L_{\mu}(p)$ (see (4.6)).

Altogether this leads to the following result:

TABLE 6.6
ESTIMATED OUTSTANDING PAYMENTS FROM TWEEDIE'S COMPOUND POISSON MODEL.

Tweedie constant $\phi = 29'281$ and $p = 1.1741$					
AY i	Outst. payments	MSEP ^{1/2}	in %	Estimation error	Process error
1	326	2'636	808.9%	1'867	1'860
2	21'565	26'773	124.2%	15'584	21'770
3	40'716	35'515	87.2%	19'122	29'927
4	89'278	53'227	59.6%	25'940	46'479
5	138'338	65'977	47.7%	30'529	58'489
6	204'269	80'815	39.6%	35'191	72'751
7	360'117	111'797	31.0%	45'584	102'082
8	596'690	149'775	25.1%	61'212	136'695
Total	1'451'299	271'503	18.7%	179'890	203'355

The results in Table 6.6 show that there is considerable uncertainty in the reserve estimates, especially in the old years where the outstanding payments are small. This comes from the fact that we have only little information to estimate $f(j)$ for large j and it turns out that the parameter estimation error lives on the same scale as the process error. For young accident years we have on the one hand a lot of information to estimate $f(j)$ for small j and on the other hand $f(j)$ for j large has a rather small influence on the overall outstanding payments estimate for young accident years in our example. Therefore the relative prediction error is smaller for young accident years

b) Over-dispersed Poisson and Gamma Model.

We first compare our result to the two boundary cases $p = 1$ and $p = 2$. These models are described in Renshaw-Verrall [6] or England-Verrall [1], Section 2.3 (over-dispersed Poisson model) and Mack [3] or England-Verrall [1], Section 3.3 (Gamma model). We also refer to (3.10)-(3.11). We obtain the following results:

TABLE 6.7
ESTIMATED OUTSTANDING PAYMENTS FROM THE OVER-DISPersed POISSON MODEL.

Over-dispersed Poisson model with $\phi = 36'642$ and $p = 1$					
AY i	Outst. payments	MSEP ^{1/2}	in %	Estimation error	Process error
1	330	4'947	1500.7%	3'520	3'475
2	21'663	34'776	160.5%	20'386	28'174
3	41'007	46'070	112.3%	24'896	38'763
4	88'546	65'229	73.7%	31'786	56'961
5	140'150	80'795	57.6%	37'316	71'662
6	204'157	95'755	46.9%	41'089	86'491
7	362'724	125'433	34.6%	49'421	115'286
8	602'784	161'023	26.7%	61'978	148'618
Total	1'461'360	371'208	21.7%	216'965	231'403

TABLE 6.8
ESTIMATED OUTSTANDING PAYMENTS FROM THE GAMMA MODEL.

Gamma model with $\phi = 29.956$ and $p = 2$					
AY i	Outst. payments	MSEP ^{1/2}	in %	Estimation error	Process error
1	447	346	77.3%	255	233
2	20'248	13'602	67.2%	8'527	10'597
3	40'073	20'127	50.2%	13'178	15'213
4	122'899	56'984	46.4%	37'465	42'936
5	121'740	50'091	41.1%	35'106	35'730
6	221'524	91'174	41.2%	66'731	62'126
7	331'115	147'730	44.6%	107'386	101'451
8	527'988	250'816	47.5%	194'155	158'784
Total	1'386'034	336'842	24.3%	265'771	206'950

Conclusions: It is not very surprising that the over-dispersed Poisson model gives a better fit than the Gamma model (especially for young accident years we have a huge estimation error term in the Gamma model, see Table 6.8). Tweedie’s compound Poisson model converges to the over-dispersed Poisson model for $p \rightarrow 1$ and to the Gamma model for $p \rightarrow 2$. For our data set $p = 1.1741$ is close to 1, hence we expect that Tweedie’s compound Poisson results are close to the over-dispersed Poisson results. Indeed, this is the case (see Tables 6.6 and 6.7). Moreover we observe that the estimation error term is essentially smaller in Tweedie’s model than in the over-dispersed Poisson model. Two main reasons for this fact are 1) For the parameter estimations in Table 6.6 we additionally use the information coming from the number of payments r_{ij} (which is used for the estimation of p). 2) In our model, the variance parameters (ϕ, p) are orthogonal to μ , hence their uncertainties have no influence on the parameter error term coming from $\text{Var}(\hat{\mu})$.

c) Mack’s model and log-normal model.

A classical non-parametric model is the so-called chain-ladder method where we apply Mack’s formulas (see Mack [4]) for the MSEP estimation. We apply the chain-ladder method to the cumulative payments

$$D_{ij} = \sum_{k=0}^j C_{ik} = w_i \sum_{k=0}^j Y_{ik}. \tag{6.1}$$

We choose the chain-ladder factors and the estimated standard errors as follows (for the definition of $f(j)$ and $\sigma_j^2 = \alpha_j^2$ we refer to Mack [4], formulas (3) and (5)). Of course there is insufficient information for the estimation of σ_{10} . Since it is not our intention to give good strategies for estimating ultimates (this would go beyond the scope of this paper) we have just chosen a value which looks meaningful.

TABLE 6.9
CHAIN-LADDER PARAMETERS IN MACK'S MODEL.

<i>j</i>	1	2	3	4	5	6	7	8	9	10
<i>f(j)</i>	1.3277	1.0301	1.0107	1.0076	1.0030	1.0020	1.0019	1.0008	1.0008	1.0000
σ_j	157.28	34.16	14.17	23.31	5.70	7.78	8.67	3.89	3.00	0.50

This leads to the following result:

TABLE 6.10
ESTIMATED OUTSTANDING PAYMENTS FROM MACK'S MODEL.

Chain-ladder estimates					
AY <i>i</i>	Outst. payments	MSEP ^{1/2}	in %	Estimation error	Process error
1	330	3'740	1134.6%	2'661	2'627
2	21'663	19'903	91.9%	11'704	16'099
3	41'007	30'090	73.4%	15'954	25'512
4	88'546	57'012	64.4%	26'295	50'585
5	140'150	71'511	51.0%	31'476	64'212
6	204'157	75'522	37.0%	31'746	68'526
7	362'724	138'915	38.3%	49'300	129'872
8	602'784	156'413	25.9%	54'293	146'688
Total	1'461'360	286'752	19.6%	177'616	225'120

A look at the results shows that Tweedie's compound Poisson model is close to the chain-ladder estimates. For the outstanding payments this is not surprising since for $p = 1.1741$, we expect that Tweedie's estimate for the outstanding payments is close to the Poisson estimate (which is identical with the chain-ladder estimate). For the error terms it is more surprising that they are so similar. The reason for this similarity is not so clear because we have estimated a different number of parameters with a different number of observations. Furthermore, MSEP is obtained in completely different ways (see also discussion in [1], Section 7.6).

Another well-known model is the so-called parametric chain-ladder method, which is based on the log-normal distribution (see Taylor [9], Section 7.3). We assume that

$$\log(D_{i,j+1}/D_{i,j}) \sim \mathcal{N}(\xi_j, \sigma_j^2) \text{ and are independent.} \tag{6.2}$$

This model is different from the one usually used in claims reserving, which would apply to incremental data (see e.g. [1], Section 3.2). We have chosen the model from Taylor [9] because it is very easy to handle.

Living in a “normal” world we estimate the parameters as in Taylor [9], formulas (7.11)-(7.13): i.e. since we assume that the parameters only depend on the development period, we take the unweighted averages to estimate ξ_j and the canonical variance estimate for σ_j^2 . This implies:

TABLE 6.11
PARAMETER ESTIMATES IN THE LOG-NORMAL MODEL.

<i>j</i>	1	2	3	4	5	6	7	8	9	10
$\xi(j)$	0.2832	0.0293	0.0106	0.0077	0.0030	0.0020	0.0019	0.0008	0.0008	0.0000
σ_j	0.0274	0.0067	0.0027	0.0046	0.0011	0.0015	0.0016	0.0007	0.0004	0.0001

The prediction errors are estimated according to Taylor [9], formulas (7.29)-(7-35). This leads to the following result:

TABLE 6.12
ESTIMATED OUTSTANDING PAYMENTS FROM THE LOG-NORMAL MODEL.

Log-normal model					
AY <i>i</i>	Outst. payments	MSEP ^{1/2}	in %	Estimation error	Process error
1	330	3'905	1183.7%	2'761	2'761
2	21'603	14'297	66.2%	8'412	11'561
3	40'814	26'680	65.4%	13'991	22'717
4	88'535	53'940	60.9%	25'130	47'728
5	140'739	69'027	49.0%	30'676	61'836
6	205'396	71'506	34.8%	31'043	64'416
7	367'545	131'216	35.7%	49'386	121'568
8	608'277	147'156	24.2%	54'163	136'826
Total	1'473'238	271'252	18.4%	170'789	210'733

The log-normal model gives estimates for the outstanding payments which are close to the chain-ladder estimates, and hence are close to Tweedie’s estimates. We have very often observed this similarity. One remarkable difference between Tweedie’s MSEP estimates and log-normal MSEP estimates is, that the Tweedie model gives more weight to the uncertainties for high development periods where one has only a few observations. This may come from the fact that for the chain-ladder model we consider cumulative data. This cumulation has already some smoothing effect.

CONCLUSIONS

Of course, we start the actuarial analysis of our claims reserving problem by the chain-ladder method. The chain-ladder reserve can very easily be calculated.

But we believe that it is also worth to perform Tweedie's compound Poisson method. Using the additional information r_{ij} one obtains an estimate for the variance function $V(\mu) = \mu^p$. If p is close to 1, Tweedie's compound Poisson method supports that the chain-ladder estimate. Whereas for p different from 1 it is questionable to believe in the chain-ladder reserve, since Tweedie's model tells us that we should rather consider a different model (e.g. the Gamma model for p close to 2).

A. REPARAMETRIZATION

We closely follow [2]. We skip the indices i, j . The joint density of (Y, R) is

$$\begin{aligned}
 f_{R,Y}(r, y; \lambda, \tau, \gamma) dy &= P[y < Y < y + dy | R = r] \cdot P[R = r] \\
 &= \frac{(\gamma/\tau)^{r\gamma} (yw)^{\gamma y - 1}}{\Gamma(r\gamma)} \exp\left\{-\frac{\gamma}{\tau} yw\right\} \cdot \frac{(w\lambda)^r}{r!} \exp\{-w\lambda\} w dy \\
 &= \left\{ \lambda (\gamma/\tau)^\gamma y^\gamma w^{\gamma+1} \right\}^r \frac{1}{r! \Gamma(r\gamma) y} \exp\left\{ \frac{w}{\phi} \left(-y \frac{\gamma\phi}{\tau} - \lambda\phi \right) \right\} dy \tag{A.1} \\
 &= \left\{ \frac{(w/\phi)^{\gamma+1} y^\gamma}{(p-1)^\gamma (2-p)} \right\}^r \frac{1}{r! \Gamma(r\gamma) y} \exp\left\{ \frac{w}{\phi} \left(y \frac{\mu^{1-p}}{1-p} - \frac{\mu^{2-p}}{2-p} \right) \right\} dy.
 \end{aligned}$$

Hence the density of Y can be obtained summing over all possible values of R :

$$\begin{aligned}
 f_Y(y; \mu, \phi/w, p) &= \sum_r f_{R,Y}(r, y; \lambda, \tau, \gamma) \\
 &= \left(\sum_r \left\{ \frac{(w/\phi)^{\gamma+1} y^\gamma}{(p-1)^\gamma (2-p)} \right\}^r \frac{1}{r! \Gamma(r\gamma) y} \right) \exp\left\{ \frac{w}{\phi} \left(y \frac{\mu^{1-p}}{1-p} - \frac{\mu^{2-p}}{2-p} \right) \right\} \tag{A.2} \\
 &= c(y; \phi/w, p) \cdot \exp\left\{ \frac{w}{\phi} (y \cdot \theta - \kappa_p(\theta)) \right\}.
 \end{aligned}$$

This proves that Y belongs to the exponential dispersion family $ED^{(p)}(\mu, \phi/w)$.

ACKNOWLEDGEMENTS

The author thanks the anonymous referees for their remarks which have substantially improved this manuscript, especially concerning Subsection 4.2 and the examples section.

REFERENCES

- [1] ENGLAND, P.D., and VERRALL, R.J. (2002) Stochastic claims reserving in general insurance, Institute of Actuaries and Faculty of Actuaries, <http://www.actuaries.org.uk/files/pdf/sessional/sm0201.pdf>
- [2] JØRGENSEN, B., and DE SOUZA, M.C.P. (1994) Fitting Tweedie's compound Poisson model to insurance claims data, *Scand. Actuarial J.*, 69-93.
- [3] MACK, T. (1991) A simple parametric model for rating automobile insurance or estimating IBNR claims reserves, *Astin Bulletin* **21**, 93-109.
- [4] MACK, T. (1997) Measuring the variability of chain ladder reserve estimates, *Claims Reserving Manual* **2**, D6.1-D6.65.
- [5] McCULLAGH, P., and NELDER, J.A. (1989) Generalized linear models, 2nd edition, Chapman and Hall.
- [6] RENSHAW, A.E., and VERRALL, R.J. (1998) A stochastic model underlying the chain-ladder technique, *B.A.J.* **4**, 903-923.
- [7] SMYTH, G.K. (1996) Partitioned algorithms for maximum likelihood and other nonlinear estimation, *Statistics and Computing* **6**, 201-216.
- [8] SMYTH, G.K., and JØRGENSEN, B. (2002), Fitting Tweedie's compound Poisson model to insurance claims data: dispersion modelling, *Astin Bulletin* **32**, 143-157.
- [9] TAYLOR, G. (2000), Loss reserving: an actuarial perspective, Kluwer Academic Publishers.
- [10] TWEEDIE, M.C.K. (1984), An index which distinguishes between some important exponential families. In *Statistics: Applications and new directions. Proceeding of the Indian Statistical Golden Jubilee International Conference*, Eds. J.K. Ghosh and J. Roy, 579-604, Indian Statistical Institute, Calcutta.
- [11] WRIGHT, T.S. (1990), A stochastic method for claims reserving in general insurance, *J.I.A.* **117**, 677-731.

MARIO V. WÜTHRICH
Winterthur Insurance
Römerstrasse 17, P.O. Box 357,
CH-8401 Winterthur
Switzerland
E-mail: mario.wuethrich@winterthur.ch