Model risk in claims reserving within Tweedie's compound Poisson models

Dr Pavel V. Shevchenko
Principal Research Scientist, Team leader
CSIRO Division of Mathematical and Information Sciences
Quantitative Risk Management group, Sydney, Australia
E-mail: Pavel.Shevchenko@csiro.au  www.cmis.csiro.au/Pavel.Shevchenko

Gareth Peters (UNSW/CSIRO), Pavel Shevchenko (CSIRO), and Mario Wüthrich (ETH). Model risk in claims reserving within Tweedie's compound Poisson models, preprint 2008.

ASTIN Colloquium, Manchester, UK, July 2008
Commonwealth Scientific and Industrial Research Organization of Australia (CSIRO) – national research agency formed in 1926. Approx 6500 staff (Divisions: Industrial Physics, Minerals, Mathematical & Information Sciences, Marine and Atmospheric Research, etc.) www.csiro.au

Division of Mathematical and Information Sciences (CMIS) (over 100 researchers): Decision Technology, Biotechnology and Health Informatics, Environmental Informatics www.cmis.csiro.au

Claims Reserving (non-life insurance), solvency requirements, claims development triangle (real data)

<table>
<thead>
<tr>
<th>accident year $i$</th>
<th>development years $j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0 1 2 3 4 5 6 7 8 9 10</td>
</tr>
<tr>
<td>1</td>
<td>$D_I = {Y_{i,j}; \quad i + j \leq I}$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td></td>
</tr>
<tr>
<td>$i$</td>
<td></td>
</tr>
<tr>
<td>$\vdots$</td>
<td></td>
</tr>
<tr>
<td>$I-1$</td>
<td></td>
</tr>
<tr>
<td>$I$</td>
<td></td>
</tr>
</tbody>
</table>

- **Observed claims payments**: $Y_{i,j} \in D_I$
- **Outstanding claims payment**: $R = \sum_{i=1}^{I} R_i = \sum_{i+j>I} Y_{i,j}$

- **Development triangle $D_I^c$**: $\{Y_{i,j}; \quad i + j > I, \quad i \leq I\}$
Content

- Tweedie’s compound Poisson family to model annual claims
- Process Uncertainty, Parameter Estimation Error, Model uncertainty
- Variable selection
- Maximum likelihood and Bayesian estimation
- MCMC (random walk Metropolis-Hastings within Gibbs)
- Analysis/Conclusions
\( \hat{R} \) - predictor for \( R \) and estimator for \( E [ R \mid D_I ] \)

\[
R = \sum_{i=1}^{I} R_i = \sum_{i+j>I} Y_{i,j} \quad E [ R \mid D_I ] = \sum_{i=1}^{I} E [ R_i \mid D_I ]
\]

\[
msep_{R|D_I} \left( \hat{R} \right) = E \left[ \left( R - \hat{R} \right)^2 \right| D_I ] \quad \text{Mean Square Error of Prediction}
\]

\[
msep_{R|D_I} \left( \hat{R} \right) = \text{Var} \left( R \mid D_I \right) + \left( E [ R \mid D_I ] - \hat{R} \right)^2
\]

\[
= \text{process variance} + \text{estimation error}
\]

\( \hat{R} = E [ R \mid D_I ] \quad \text{“best estimate” of reserve} \)

**Bayesian context – variance decomposition**

\[
\text{Var} \left( R \mid D_I \right) = E \left[ \text{Var} \left( R \mid \theta, D_I \right) \right| D_I ] + \text{Var} \left( E [ R \mid \theta, D_I ] \right| D_I ]
\]

\[
= \text{average process variance} + \text{parameter estimation error}.
\]

\( \theta \) is model parameter vector modelled as random variable
Tweedie’s compound Poisson model

\[ Y_{i,j} \text{ are independent for } i, j \in \{0, \ldots, I\} \]

\[ Y_{i,j} = 1\{N_{i,j}>0\} \sum_{k=1}^{N_{i,j}} X_{i,j}^{(k)} \]

\[ N_{i,j} \text{ and } X_{i,j}^{(k)} \text{ are independent for all } k \]

\[ N_{i,j} \text{ is Poisson distributed with parameter } \lambda_{i,j} \]

\[ X_{i,j}^{(k)} \text{ are independent gamma severities with mean } \tau_{i,j} > 0 \text{ and shape parameter } \gamma > 0 \]
Tweedie’s compound Poisson: exponential dispersion family representation

\[ P [Y_{i,j} = 0] = P [N_{i,j} = 0] = \exp \left\{ -\phi_{i,j}^{-1} \kappa_p(\theta_{i,j}) \right\} \]

\[ f_{\theta_{i,j}}(y; \phi_{i,j}, p) = c(y; \phi_{i,j}, p) \exp \left\{ \frac{y \theta_{i,j} - \kappa_p(\theta_{i,j})}{\phi_{i,j}} \right\} \quad \text{for} \quad y > 0 \]

where \( \theta_{i,j} < 0, \phi_{i,j} > 0, \quad \kappa_p(\theta) \overset{\text{def.}}{=} \frac{1}{2 - p} [(1 - p)\theta]^{\gamma} \)

\[ c(y; \phi, p) = \sum_{r \geq 1} \left( \frac{(1/\phi)^{\gamma+1} y^\gamma}{(p - 1)\gamma(2 - p)} \right)^r \frac{1}{r! \Gamma(r\gamma) y} \]

\[ p = p(\gamma) = \frac{\gamma + 2}{\gamma + 1} \quad \in \quad (1, 2) \quad \phi_{i,j} = \frac{\lambda_{i,j}^{1-p} \tau_{i,j}^{2-p}}{2 - p} > 0 \]

\[ \theta_{i,j} = \left( \frac{1}{1-p} \right) (\mu_{i,j})^{(1-p)} < 0, \quad \mu_{i,j} = \lambda_{i,j} \tau_{i,j} > 0 \]
Tweedie’s compound Poisson model: final representation

\[
P[Y_{i,j} = 0] = P[N_{i,j} = 0] = \exp \left\{ -\frac{\phi_{i,j}^{-1}}{2-p} \frac{\mu_{i,j}^{2-p}}{2} \right\}
\]

\[
f_{\mu_{i,j}}(y; \phi_{i,j}, p) = c(y; \phi_{i,j}, p) \exp \left\{ \phi_{i,j}^{-1} \left[ y \frac{\mu_{i,j}^{1-p}}{1-p} - \frac{\mu_{i,j}^{2-p}}{2-p} \right] \right\} \quad \text{for } y > 0
\]

\[
E[Y_{i,j}] = \frac{\partial}{\partial \theta_{i,j}} \kappa_p(\theta_{i,j}) = \kappa'_p(\theta_{i,j}) = [(1-p)\theta_{i,j}]^{1/(1-p)} = \mu_{i,j}
\]

\[
\text{Var}(Y_{i,j}) = \phi_{i,j} \kappa''_p(\theta_{i,j}) = \phi_{i,j} \mu_{i,j}^p
\]

\[p \in (1, 2), \text{ typically fixed by the modeller} \quad \text{Model Risk}
\]

\[p \to 1, \text{ overdispersed Poisson model}
\]

\[p \to 2, \text{ gamma model}\]
Parameter estimation: estimate $\mu_{i,j}$, $p$ and $\phi_{i,j}$ using $\mathcal{D}_I$

Model assumptions: multiplicative model $\mu_{i,j} = \alpha_i \beta_j$, $\alpha_0 = 1$, $\phi_{i,j} = \phi$ and $\alpha_i > 0$, $\beta_j > 0$

<table>
<thead>
<tr>
<th>accident year $i$</th>
<th>development pattern $\beta = (\beta_0, \ldots, \beta_I)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

observed claims payments $Y_{i,j} \in \mathcal{D}_I$

$\mathcal{D}_I = \{Y_{i,j}; \ i + j \leq I\}$

outstanding claims payment

$$R = \sum_{i=1}^{I} R_i = \sum_{i+j>I}^{i} Y_{i,j}.$$  

$\mathcal{D}_I^c = \{Y_{i,j}; \ i + j > I, i \leq I\}$
Likelihood function

\[ \theta = (p, \phi, \alpha, \beta) \]

\[ L_{D_1}(\theta) = \prod_{i+j \leq I} c(Y_{i,j}; \phi, p) \exp \left\{ \phi^{-1} \left[ Y_{i,j} \frac{(\alpha_i \beta_j)^{1-p}}{1-p} - \frac{(\alpha_i \beta_j)^{2-p}}{2-p} \right] \right\} \]

\[ c(y; \phi, p) = \sum_{r \geq 1} \left( \frac{(1/\phi)^{\gamma+1}y^{\gamma}}{(p-1)^{\gamma}(2-p)} \right)^r \frac{1}{r! \Gamma(r \gamma)y} = \frac{1}{y} \sum_{r \geq 1} W_r \]

\[ \log W_r = r \log z - \log \Gamma(1 + r) - \log \Gamma(\gamma r), \quad z = \frac{(1/\phi)^{\gamma+1}y^{\gamma}}{(p-1)^{\gamma}(2-p)} \]

\[ W_r \text{ is unimodal in } r, \quad R_0 = R_0(\phi, p) = \frac{y^{2-p}}{(2-p)\phi} \]

find \( R_L < R_0 < R_U \) such that

\[ W_{R_L} \leq e^{-37}W_{R_0} \text{ (or } R_L = 1) \text{ and } W_{R_U} \leq e^{-37}W_{R_0} \]

\[ c(y; \phi, p) \approx \tilde{c}(y; \phi, p) = \frac{1}{y} \sum_{r=R_L}^{R_U} W_r \]
Maximum likelihood estimation
maximizing $L_{D_I}(\theta)$ in $\theta = (p, \phi, \alpha, \beta)$ leads to

$\hat{\theta}^{\text{MLE}} = (\hat{p}^{\text{MLE}}, \hat{\phi}^{\text{MLE}}, \hat{\alpha}^{\text{MLE}}, \hat{\beta}^{\text{MLE}})$

typically MLE is done for fixed $p$ (expert choice)

$\hat{R}^{\text{MLE}} = \sum_{i+j>I} \hat{\alpha}_i^{\text{MLE}} \hat{\beta}_j^{\text{MLE}}$ the best estimate reserves for $R$

$\text{cov} \left( \hat{\theta}_i^{\text{MLE}}, \hat{\theta}_j^{\text{MLE}} \right) = (I^{-1})_{i,j}$, $(I)_{i,j} = -\left. \frac{\partial^2 \ln L_{D_I}(\theta)}{\partial \theta_i \partial \theta_j} \right|_{\theta=\hat{\theta}^{\text{MLE}}}$

$\hat{\beta}_i^{\text{MLE}} = Y_{0,i}$, $\text{cov}(\hat{\beta}_i^{\text{MLE}}, \hat{\theta}_i^{\text{MLE}}) = 0$, $\hat{\theta}_i^{\text{MLE}} \neq \beta_i^{\text{MLE}}$
Maximum likelihood: process and estimation errors

\[
\hat{R}^{\text{MLE}} = \sum_{i+j>1} \hat{\alpha}_i^{\text{MLE}} \hat{\beta}_j^{\text{MLE}}
\]

\[
\text{stdev} \left( \hat{R}^{\text{MLE}} \right) = \sqrt{\text{Var} \left( \hat{R}^{\text{MLE}} \right)} \quad \text{parameter estimation error}
\]

\[
\hat{\text{Var}} \left( \hat{R}^{\text{MLE}} \right) = \sum_{i_1+j_1>1} \sum_{i_2+j_2>1} \hat{\alpha}_{i_1}^{\text{MLE}} \hat{\alpha}_{i_2}^{\text{MLE}} \text{cov} \left( \hat{\beta}_{j_1}^{\text{MLE}}, \hat{\beta}_{j_2}^{\text{MLE}} \right)
\]

\[+ \sum_{i_1+j_1>1} \sum_{i_2+j_2>1} \hat{\beta}_{j_1}^{\text{MLE}} \hat{\beta}_{j_2}^{\text{MLE}} \text{cov} \left( \hat{\alpha}_{i_1}^{\text{MLE}}, \hat{\alpha}_{i_2}^{\text{MLE}} \right)
\]

\[+ 2 \sum_{i_1+j_1>1} \sum_{i_2+j_2>1} \hat{\alpha}_{i_1}^{\text{MLE}} \hat{\beta}_{j_2}^{\text{MLE}} \text{cov} \left( \hat{\alpha}_{i_2}^{\text{MLE}}, \hat{\beta}_{j_1}^{\text{MLE}} \right)
\]

\[
\hat{\text{Var}} \left( R \right) = \sum_{i+j>1} \left( \hat{\alpha}_i^{\text{MLE}} \hat{\beta}_j^{\text{MLE}} \right) \hat{\rho}^{\text{MLE}} \hat{\phi}^{\text{MLE}} \quad \text{process variance}
\]

\[
\hat{\text{mse}}_{\phi,\rho} \left( \hat{R}^{\text{MLE}} \right) = \hat{\text{Var}} \left( R \right) + \hat{\text{Var}} \left( \hat{R}^{\text{MLE}} \right)
\]

\[= \text{MLE process variance} + \text{MLE estimation error}\]
Bayesian inference

\[ \theta = (p, \phi, \alpha, \beta) \] are treated as random

\[ \pi(\theta \mid \mathcal{D}_I) \propto L_{\mathcal{D}_I}(\theta) \pi(\theta) \] Markov Chain Monte Carlo (MCMC)

\[ \text{MAP} : \quad \hat{\theta}^{\text{MAP}} = \arg \max_{\theta} \left[ \pi(\theta \mid \mathcal{D}_I) \right], \]

\[ \text{MMSE} : \quad \hat{\theta}^{\text{MMSE}} = E[\theta \mid \mathcal{D}_I]. \]

if the prior \( \pi(\theta) \) is constant and the parameter range includes the MLE, then the MAP of the posterior is the same as the MLE.

Gaussian approximation for \( \pi(\theta \mid \mathcal{D}_I) \)

\[ \ln \pi(\theta \mid \mathcal{D}_I) \approx \ln \pi(\hat{\theta}^{\text{MAP}} \mid \mathcal{D}_I) \]

\[ + \frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln \pi(\theta \mid \mathcal{D}_I) \bigg|_{\theta = \hat{\theta}^{\text{MAP}}} \left( \theta_i - \hat{\theta}_i^{\text{MAP}} \right) \left( \theta_j - \hat{\theta}_j^{\text{MAP}} \right) \]
Bayesian inference estimates

\[ E \left[ R \mid D_I \right] = \sum_{i+j>I} E \left[ \alpha_i \beta_j \mid D_I \right] \]

\[ \tilde{R} = E \left[ R \mid \theta \right] = \sum_{i+j>I} \alpha_i \beta_j \]

the best consistent estimate of reserves (ER)

\[ \hat{R}^B = E \left[ \tilde{R} \mid D_I \right] = \sum_{i+j>I} E \left[ \alpha_i \beta_j \mid D_I \right] = E \left[ R \mid D_I \right] \]

\[ \text{mse}_{\text{p}}_{R \mid D_I} \left( \hat{R}^B \right) = E \left[ \left( R - \hat{R}^B \right)^2 \left| D_I \right. \right] = \text{Var} \left( R \mid D_I \right) \]

\[ \text{Var} \left( R \mid D_I \right) = \text{Var} \left( \sum_{i+j>I} Y_{i,j} \left| D_I \right. \right) \]

\[ = \sum_{i+j>I} E \left[ (\alpha_i \beta_j)^p \phi \left| D_I \right. \right] + \text{Var} \left( \tilde{R} \left| D_I \right. \right) \]

= average process variance (PV) + parameter estimation error (EE)

**Note, model error is incorporated via averaging over values of** \( p \)
Random Walk Metropolis Hastings (RW-MH) within Gibbs

1. Initialize randomly or deterministically for $t = 0$
   the parameter vector $\theta^t=0$ to the maximum likelihood estimates.
2. For $t = 1, \ldots, T$
   a) Set $\theta^t = \theta^{t-1}$
   b) For $i = 1, \ldots, 2I + 3$
      Sample proposal $\theta^*_i$ from Gaussian truncated density
      \[ f_N^T(\theta_i^*, \theta_i^t, \sigma_{RW_i}) = \frac{f_N(\theta_i^*, \theta_i^t, \sigma_{RW_i})}{F_N(b_i; \theta_i^t, \sigma_{RW_i}) - F_N(a_i; \theta_i^t, \sigma_{RW_i})} \]
      to obtain $\theta^* = (\theta_1^t, \ldots, \theta_{i-1}^t, \theta_i^*, \theta_{i+1}^t, \ldots)$
      Accept proposal with acceptance probability
      \[ \alpha(\theta^t, \theta^*) = \min \left\{ 1, \frac{\pi(\theta^* \mid \mathcal{D}_I) f_N^T(\theta_i^t; \theta_i^*, \sigma_{RW_i})}{\pi(\theta^t \mid \mathcal{D}_I) f_N^T(\theta_i^*; \theta_i^t, \sigma_{RW_i})} \right\} \]
      Note: normalization constant in posterior is not needed; optimal acceptance rate is 0.234
The prior domains
\( p \in (1.1, 1.95), \ \phi \in (0.01, 100), \ \alpha_i \in (0.01, 100) \) and \( \beta_j \in (0.01, 10^4) \)

<table>
<thead>
<tr>
<th></th>
<th>MLE</th>
<th>MLE stdv</th>
<th>Bayesian posterior</th>
<th>( \sigma_{RW} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>MMSE</td>
<td>stddev</td>
</tr>
<tr>
<td>( p )</td>
<td>1.259</td>
<td>0.149</td>
<td>1.332 (0.007)</td>
<td>0.143 (0.004)</td>
</tr>
<tr>
<td>( \phi )</td>
<td>0.351</td>
<td>0.201</td>
<td>0.533 (0.013)</td>
<td>0.289 (0.005)</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.918</td>
<td>0.056</td>
<td>0.901 (0.004)</td>
<td>0.074 (0.001)</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>0.946</td>
<td>0.051</td>
<td>0.946 (0.003)</td>
<td>0.073 (0.001)</td>
</tr>
<tr>
<td>( \alpha_3 )</td>
<td>0.861</td>
<td>0.048</td>
<td>0.861 (0.003)</td>
<td>0.068 (0.001)</td>
</tr>
<tr>
<td>( \alpha_4 )</td>
<td>0.891</td>
<td>0.049</td>
<td>0.902 (0.003)</td>
<td>0.072 (0.002)</td>
</tr>
<tr>
<td>( \alpha_5 )</td>
<td>0.879</td>
<td>0.051</td>
<td>0.876 (0.003)</td>
<td>0.070 (0.001)</td>
</tr>
<tr>
<td>( \alpha_6 )</td>
<td>0.842</td>
<td>0.048</td>
<td>0.843 (0.002)</td>
<td>0.069 (0.001)</td>
</tr>
<tr>
<td>( \alpha_7 )</td>
<td>0.762</td>
<td>0.046</td>
<td>0.762 (0.003)</td>
<td>0.066 (0.001)</td>
</tr>
<tr>
<td>( \alpha_8 )</td>
<td>0.773</td>
<td>0.047</td>
<td>0.765 (0.003)</td>
<td>0.067 (0.001)</td>
</tr>
<tr>
<td>( \alpha_9 )</td>
<td>0.848</td>
<td>0.059</td>
<td>0.856 (0.003)</td>
<td>0.090 (0.002)</td>
</tr>
</tbody>
</table>

Table 3: MLE and Bayesian estimators.
Table 3: MLE and Bayesian estimators.

<table>
<thead>
<tr>
<th></th>
<th>MLE</th>
<th>MLE stdev</th>
<th>Bayesian posterior</th>
<th>σRW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>MMSE</td>
<td>stdev</td>
</tr>
<tr>
<td>β₀</td>
<td>669.1</td>
<td>27.7</td>
<td>672.7 (2.1)</td>
<td>39.7 (0.7)</td>
</tr>
<tr>
<td>β₁</td>
<td>329.0</td>
<td>14.4</td>
<td>331.1 (1.0)</td>
<td>20.6 (0.4)</td>
</tr>
<tr>
<td>β₂</td>
<td>77.43</td>
<td>4.38</td>
<td>78.06 (0.24)</td>
<td>6.10 (0.06)</td>
</tr>
<tr>
<td>β₃</td>
<td>24.59</td>
<td>1.96</td>
<td>24.95 (0.08)</td>
<td>2.64 (0.03)</td>
</tr>
<tr>
<td>β₄</td>
<td>16.28</td>
<td>1.55</td>
<td>16.65 (0.05)</td>
<td>2.09 (0.03)</td>
</tr>
<tr>
<td>β₅</td>
<td>7.773</td>
<td>1.028</td>
<td>8.068 (0.024)</td>
<td>1.356 (0.020)</td>
</tr>
<tr>
<td>β₆</td>
<td>5.776</td>
<td>0.937</td>
<td>6.115 (0.022)</td>
<td>1.261 (0.016)</td>
</tr>
<tr>
<td>β₇</td>
<td>1.219</td>
<td>0.396</td>
<td>1.494 (0.006)</td>
<td>0.609 (0.013)</td>
</tr>
<tr>
<td>β₈</td>
<td>1.188</td>
<td>0.476</td>
<td>1.622 (0.008)</td>
<td>0.802 (0.016)</td>
</tr>
<tr>
<td>β₉</td>
<td>1.581</td>
<td>0.790</td>
<td>2.439 (0.021)</td>
<td>1.496 (0.026)</td>
</tr>
</tbody>
</table>
Figure 3: Predicted distribution of reserves, $\tilde{R} = \sum_{i+j>l} \alpha_i \beta_j$. 
Full predictive distribution

\[ f(R | D_I) = \int g(R | \theta) \pi(\theta | D_I) d\theta \]

Figure 5: Distribution of total outstanding claims payment \( R = \sum_{i+j>1} Y_{i,j} \), accounting for all process, estimation and model uncertainties.
Variable selection models

\[ \text{development pattern } \beta = (\beta_0, \ldots, \beta_I) \]

<table>
<thead>
<tr>
<th>accident year ( i )</th>
<th>development years ( j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
</tr>
<tr>
<td>( i )</td>
<td>\vdots</td>
</tr>
<tr>
<td>( I - 1 )</td>
<td>\vdots</td>
</tr>
<tr>
<td>( I )</td>
<td>\vdots</td>
</tr>
</tbody>
</table>

observed claims payments \( Y_{i,j} \in \mathcal{D}_I \)

\[ \mathcal{D}_I = \{Y_{i,j}; \ i + j \leq I\} \]

outstanding claims payment

\[ R = \sum_{i=1}^{I} R_i = \sum_{i+j > I} Y_{i,j}. \]

\[ \mathcal{D}^c_I = \{Y_{i,j}; \ i + j > I, \ i \leq I\} \]
Variable selection models

- $M_0 : \theta_{[0]} = \left( p, \phi, \tilde{\alpha}_0 = \alpha_0, ..., \tilde{\alpha}_I = \alpha_I, \tilde{\beta}_0 = \beta_0, ..., \tilde{\beta}_I = \beta_I \right)$

- $M_1 : \theta_{[1]} = \left( p, \phi, \tilde{\beta}_0 \right)$ with $\left( \tilde{\beta}_0 = \beta_0 = ... = \beta_I \right)$, $(\alpha_0 = ... = \alpha_I = 1)$

- $M_2 : \theta_{[2]} = \left( p, \phi, \tilde{\alpha}_1, \tilde{\beta}_0, \tilde{\beta}_1 \right)$ with $\left( \alpha_0 = ... = \alpha_4 = 1 \right)$, $\left( \tilde{\alpha}_1 = \alpha_5 = ... = \alpha_I \right)$, $\left( \tilde{\beta}_0 = \beta_0 = ... = \beta_4 \right)$, $\left( \tilde{\beta}_1 = \beta_5 = ... = \beta_I \right)$.

- $M_6 : \theta_{[6]} = \left( p, \phi, \alpha_0, \tilde{\alpha}_1, \beta_0, \beta_1, ..., \beta_I \right)$ with $\left( \tilde{\alpha}_1 = \alpha_1 = ... = \alpha_I \right)$
Variable selection models

the joint posterior distribution \( \pi(M_k, \theta_{[k]} \mid D_I) \), \( \theta_{[k]} = (\tilde{\theta}_1, \tilde{\theta}_2, \ldots, \tilde{\theta}_{N_{[k]}}) \)
a prior distribution \( \pi(M_k) \) for the model
a prior for the parameters conditional on the model \( \pi(\theta_{[k]} \mid M_k) \)

\[
\pi(M_k \mid D_I) = \int \pi(M_k, \theta_{[k]} \mid D_I) d\theta_{[k]} = \int \pi(M_k \mid \theta_{[k]}, D_I) \pi(\theta_{[k]} \mid D_I) d\theta_{[k]} 
\approx \frac{1}{T - T_b} \sum_{j=T_b+1}^{T} \frac{L_{D_I}(M_k, \theta_{j,[k]})}{\sum_{m=0}^{K} L_{D_I}(M_m, \theta_{j,[m]})}
\]

<table>
<thead>
<tr>
<th></th>
<th>( M_0 )</th>
<th>( M_1 )</th>
<th>( M_2 )</th>
<th>( M_3 )</th>
<th>( M_4 )</th>
<th>( M_5 )</th>
<th>( M_6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \pi(M_k \mid D_I) )</td>
<td>0.71</td>
<td>4.19E-54</td>
<td>3.04E-43</td>
<td>1.03E-28</td>
<td>6.71E-20</td>
<td>2.17E-21</td>
<td>0.29</td>
</tr>
<tr>
<td>DIC</td>
<td>399</td>
<td>649</td>
<td>600</td>
<td>535</td>
<td>498</td>
<td>507</td>
<td>398</td>
</tr>
<tr>
<td>LHR p-value</td>
<td>1</td>
<td>2.76E-50</td>
<td>1.67E-40</td>
<td>3.53E-28</td>
<td>5.78E-21</td>
<td>3.03E-23</td>
<td>0.043</td>
</tr>
</tbody>
</table>

Table 4: Posterior model probabilities \( \pi(M_k \mid D_I) \), Deviance Information Criterion (DIC) and Likelihood Ratio (LHR) p-values for variable selection models \( M_0, \ldots, M_6 \)
## Claims reserves

<table>
<thead>
<tr>
<th>Model Averaging</th>
<th>Model Selection for $p$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimated Reserves</strong></td>
<td>$ER = \hat{R}^B = E[\hat{R}</td>
</tr>
<tr>
<td><strong>Process Variance</strong></td>
<td>$PV = E \left[ \sum \phi (\alpha_i \beta_j)^p</td>
</tr>
<tr>
<td><strong>Estimation Error</strong></td>
<td>$EE = \text{Var}(\hat{R}</td>
</tr>
</tbody>
</table>

Table 5: Quantities used for analysis of the claims reserving problem

\[
ER = E[ER_p|\mathcal{D}_I],
\]

\[
PV = E[PV_p|\mathcal{D}_I],
\]

\[
EE = E[EE_p|\mathcal{D}_I] + \text{Var}(ER_p|\mathcal{D}_I)
\]
<table>
<thead>
<tr>
<th>Statistic</th>
<th>Bayesian Estimate</th>
<th>MLE Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ER$</td>
<td>624.1 (0.7)</td>
<td>602.630</td>
</tr>
<tr>
<td>$\sqrt{PV}$</td>
<td>37.3 (0.2)</td>
<td>25.937</td>
</tr>
<tr>
<td>$\sqrt{EE}$</td>
<td>44.8 (0.5)</td>
<td>28.336</td>
</tr>
<tr>
<td>$\sqrt{MSEP}$</td>
<td>58.3 (0.5)</td>
<td>38.414</td>
</tr>
</tbody>
</table>

Table 6: Model averaged estimates

<table>
<thead>
<tr>
<th>$VaR_q$</th>
<th>$R$</th>
<th>$\tilde{R}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$VaR_{75%}$</td>
<td>659.8 (0.9)</td>
<td>650.6 (1.0)</td>
</tr>
<tr>
<td>$VaR_{90%}$</td>
<td>698.4 (1.2)</td>
<td>680.4 (1.3)</td>
</tr>
<tr>
<td>$VaR_{95%}$</td>
<td>724.0 (1.5)</td>
<td>701.7 (1.6)</td>
</tr>
</tbody>
</table>

Table 7: Bayesian model averaged estimates of Value at Risk
Figure 6: Estimates of quantities from Table 5 conditional on $p$. 
Results: average over $p$

- Claims reserve MLE, $\hat{R}^\text{MLE}$, is less than Bayesian estimate $\hat{R}^\text{B}$
- $\sqrt{EE}$ and $\sqrt{PV}$ are of the same magnitude, 6-7% of total reserve
- MLEs for $\sqrt{EE}$ and $\sqrt{PV}$ are less than Bayesian estimates, $\approx 30\%$
- The difference between $\hat{R}^\text{MLE}$ and $\hat{R}^\text{B}$ is of the order of magnitude as $\sqrt{EE}$ and $\sqrt{PV}$

Results: conditioning on $p$

- MLE of $ER_p$ is almost constant
- Bayesian estimates for $ER_p$ changes as a function of $p$.
- Bayesian estimators for $\sqrt{PV}_p$ and $\sqrt{EE}_p$ increase as $p$ increases
  The MLEs for $PV_p$ and $EE_p$ are significantly less than Bayesian estimators.
Conclusions

- Development of a Bayesian model for claims reserving under Tweedie’s compound Poisson model covering range between Poisson and Gamma models
- Quantification process, estimation and model uncertainties and variable selection using MCMC (random walk Metropolis-Hastings within Gibbs)
- MLEs are materially different from Bayesian estimators – posterior distributions are different from Gaussian.

Future work: variable selection problem – Reversible Jump MCMC; considering model parameter $p$ outside the (1, 2) range; dynamic model.
Thank you