A photograph of a city skyline, likely New York City, viewed from a boat. The water is in the foreground, and the city buildings are in the background. A large, dark, curved object, possibly a boat's railing or part of the steering wheel, is in the immediate foreground, slightly out of focus. An orange rectangular block is visible on the left side of the image.

An Application of Bayesian Analysis in Forecasting Insurance Loss Payments

Yanwei (Wayne) Zhang, Statistical Research, CNA Insurance Company





Highlights:

- Bayesian methodology and actuarial science
- Case study in loss reserving
- Questions
- Appendix: Bayesian analysis in Excel



Part I

Bayesian methodology and actuarial science



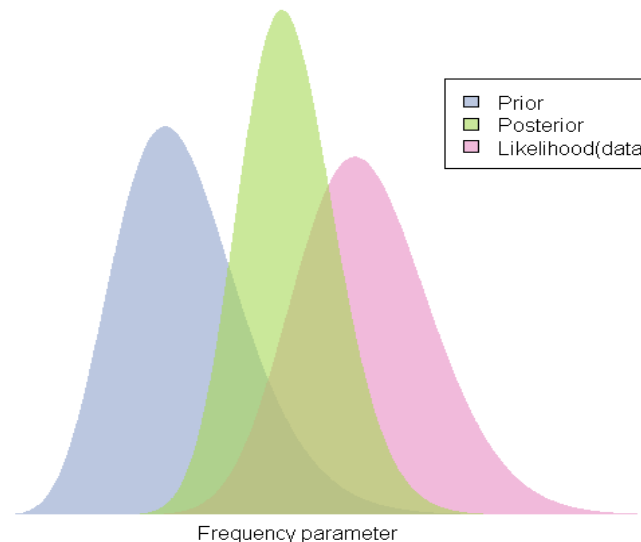
Bayesian methodology

- Fundamental question is

Given data and a specified model, what is the distribution of the parameters?

- With the posterior distribution of the parameters, the distribution of any quantities of interest can be obtained
- The key is the Bayes' theorem:

Posterior distribution is proportional to data distribution * prior distribution





Application to the actuarial field

- Most of you are Bayesian!
 - Bornhuetter-Ferguson type reserving to regulate data or account for information not in data with prior knowledge of the average loss ratio
 - Credibility. Bühlmann and Gisler (2005) said
“Credibility theory belongs mathematically to the area of Bayesian statistics [and it] is motivated by questions arising in insurance practice.”
- So, when you are talking about these, you are thinking in a Bayesian world
- But.....Few of you are doing Bayesian analysis!
- Now, there is an opportunity to be a real Bayesian!

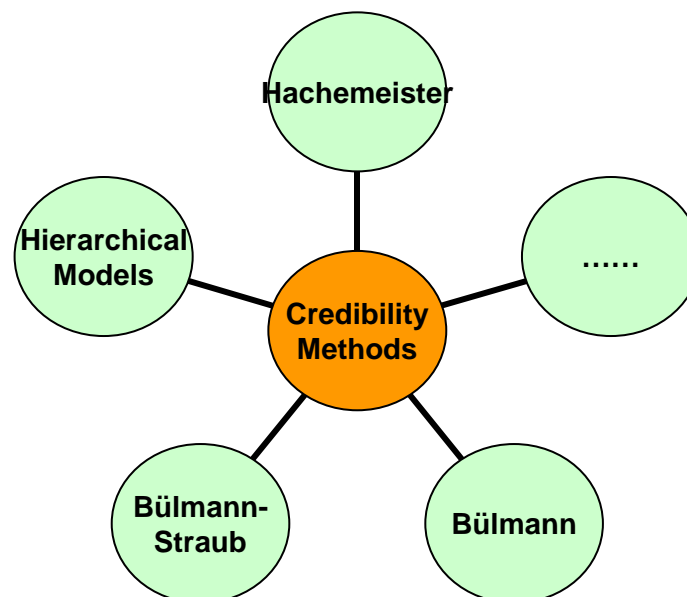


More on credibility

- Credibility theory refers to

“any procedure that uses information (‘borrows strength’) from samples from different, but related, populations.” -- Klugman (1987)

- We should not retain the word just for the actuarial credibility formulas
- These formulas are only a subset of all credibility methods

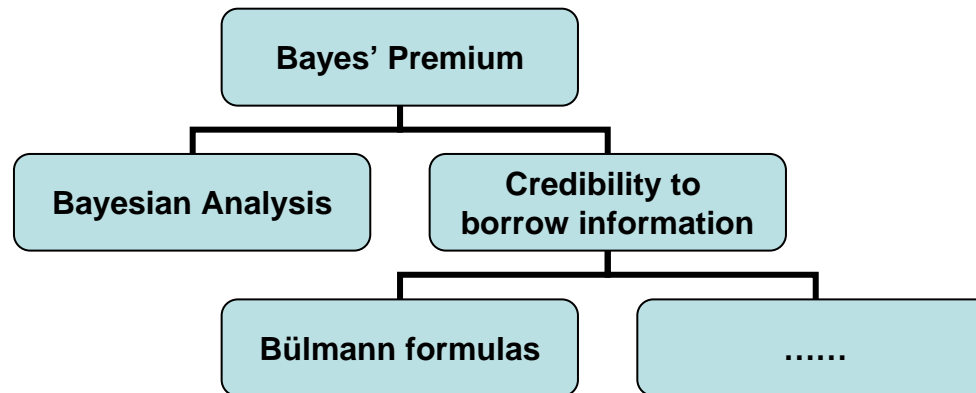




More on credibility

- We should recall that Bayesian analysis is where actuarial credibility theory started

Given a group of policyholders with some common risk factor and past claims experience, what is the Bayes' premium to be charged for each policyholder?



- These formulas are only linear approximations to overcome computational difficulties:
 - No closed form except for some simple models and distributions
 - Hard to estimate the population parameters



Credibility example

- Now, there is no reason to linearly approximate Bayesian methods as advances in statistical computation in the past several decades have enabled more complex and realistic models to be constructed and estimated
- Consider the following example in Workers' Comp (see Scollnik 2001):

| Year | Group 1 | | Group 2 | | Group 3 | |
|------|---------|----------|---------|----------|---------|----------|
| | Payroll | # Claims | Payroll | # Claims | Payroll | # Claims |
| 1 | 280 | 9 | 260 | 6 | | |
| 2 | 320 | 7 | 275 | 4 | 145 | 8 |
| 3 | 265 | 6 | 240 | 2 | 120 | 3 |
| 4 | 340 | 13 | 265 | 8 | 105 | 4 |

- Question: What's the expected count for year 5, given the observed claim history?
- Let's do the Bayesian analysis and then compare it with other estimates



Visualization of the hierarchies

- Intuitively, we assume claim count to be a Poisson distribution

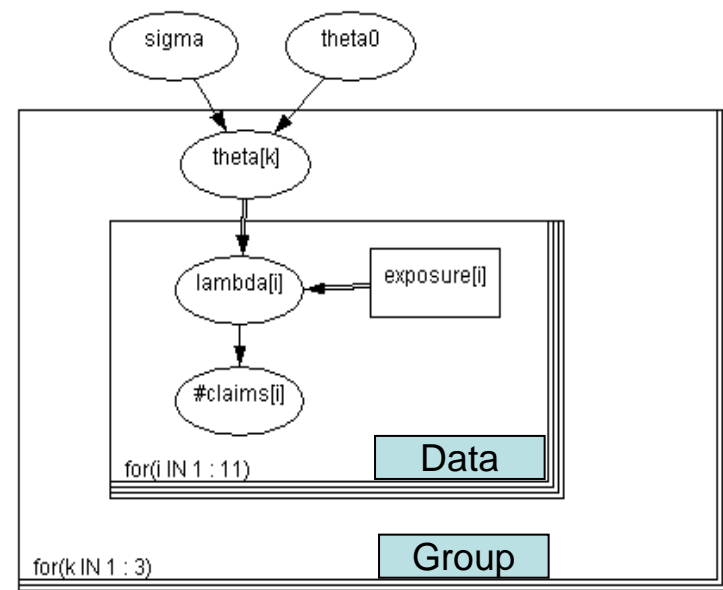
$$\#claims_{ik} \sim Pois(exposure_{ik} \times \theta_k)$$

- Credibility view assumes that each group has a different claim rate per exposure θ_k , but each θ_k arises from the same distribution, say

$$\log \theta_k \sim N(\log \theta_0, \sigma^2)$$

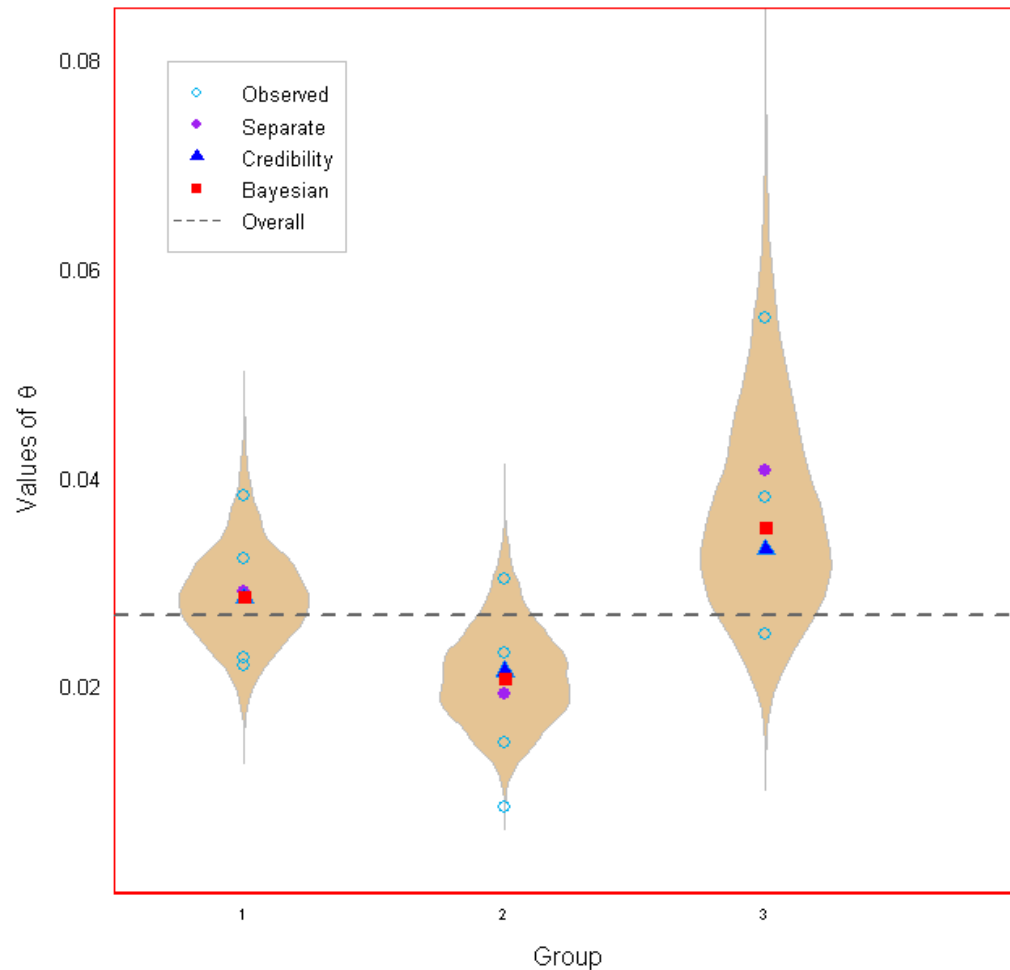
- If θ_0 is estimated using all the data, so will each θ_k . Thus, the estimation of one group will borrow information from other groups, and will be pooled toward the overall mean
- Assign non-informative (flat) priors so that σ and θ_0 are estimated from the data, e.g.

$$\sigma \sim U(0,100); \theta_0 \sim N(0,100^2)$$





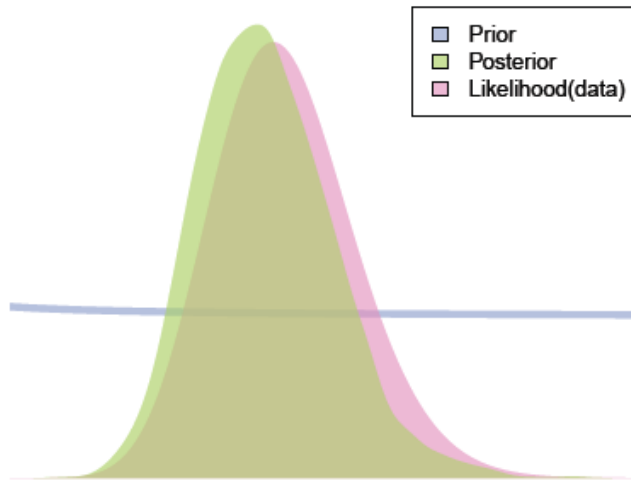
Results of different estimations





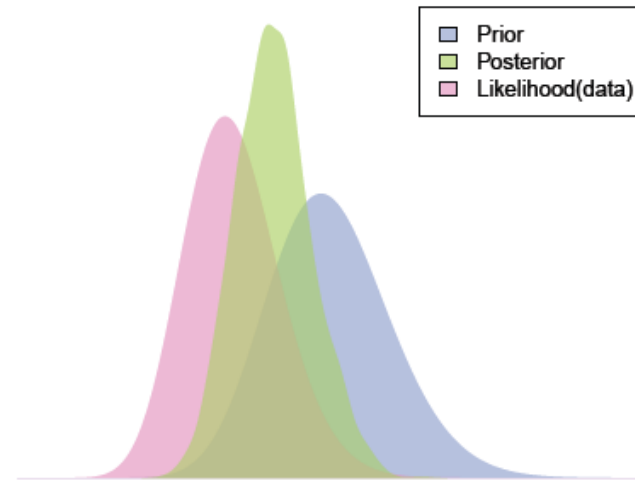
Visualization of posterior distribution

Distribution for Group 1:
Hierarchical structure, non-informative prior



Frequency parameter

Distribution for Group 1:
Informative prior with mean 0.04, small variance



Frequency parameter

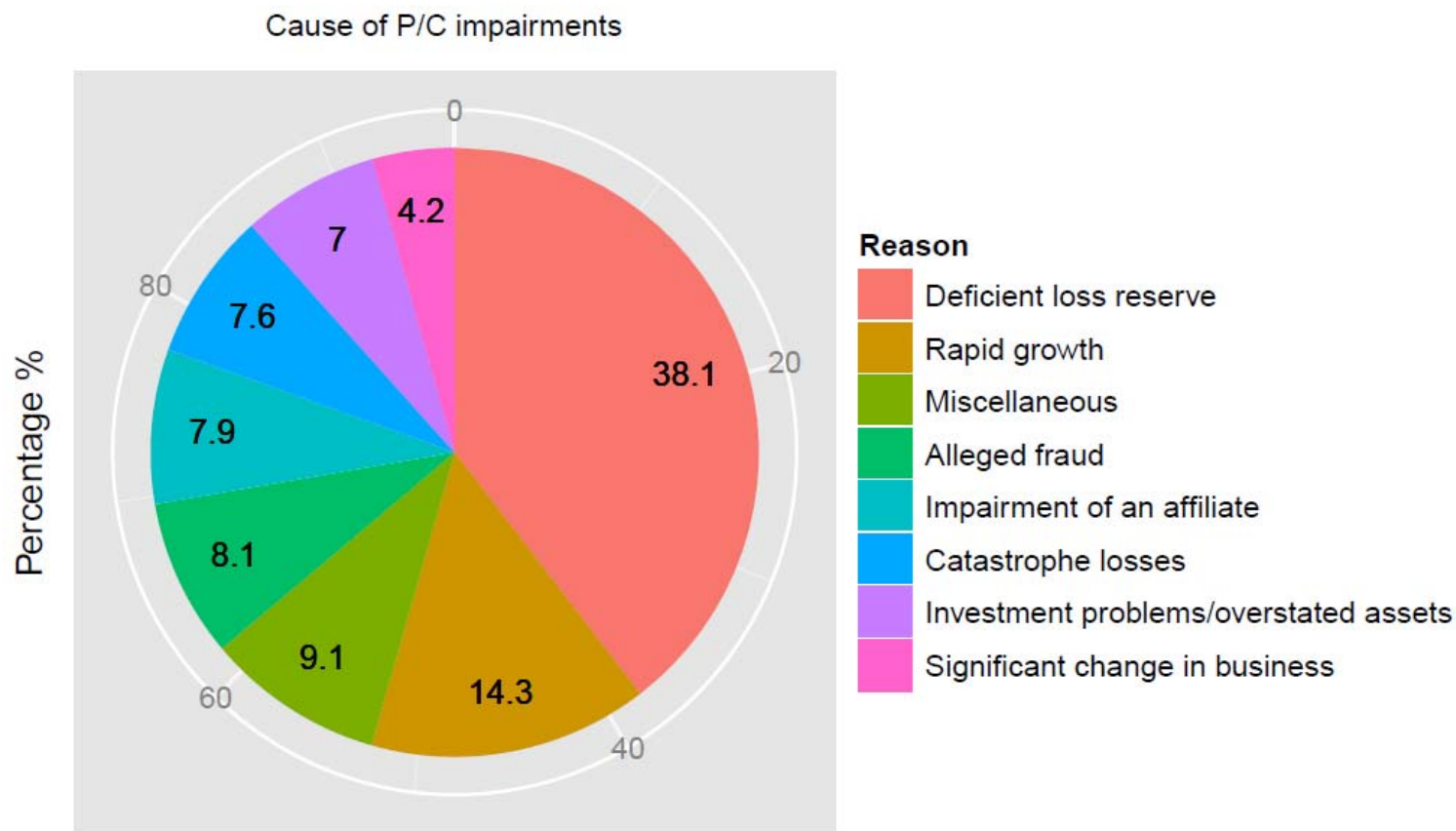


Part II

Case study in loss reserving



Importance of loss reserving





Challenges

- Challenges in current loss reserving practices:
 - Most stochastic models need to be supplemented by a tail factor, but the corresponding uncertainty is hard to be accounted for
 - Inference at an arbitrary point is hard to obtain, e.g., 3 months or 9 months
 - Too many parameters! Parsimony is the basic principle of statistics
 - Treat accident year, development lag, or both independently
 - Focus on one triangle, lack a method to blend industry data
 - Usually rely on post-model selection using judgment:
 - Input of point estimate is almost meaningless, but large leverage
 - Extra uncertainty is not accounted for



Benefits of the Bayesian model to be built

- Allows input of external information and expert opinion
- Blending of information across accident years and across companies
- Extrapolates development beyond the range of observed data
- Estimates at any time point can be made
- Uncertainty of extrapolation is directly included
- Full distribution is available, not just standard error
- Prediction of a new accident year can be achieved
- Minimizes the risk of underestimating the uncertainty in traditional models
- Estimation of company-level and accident-year-level variations



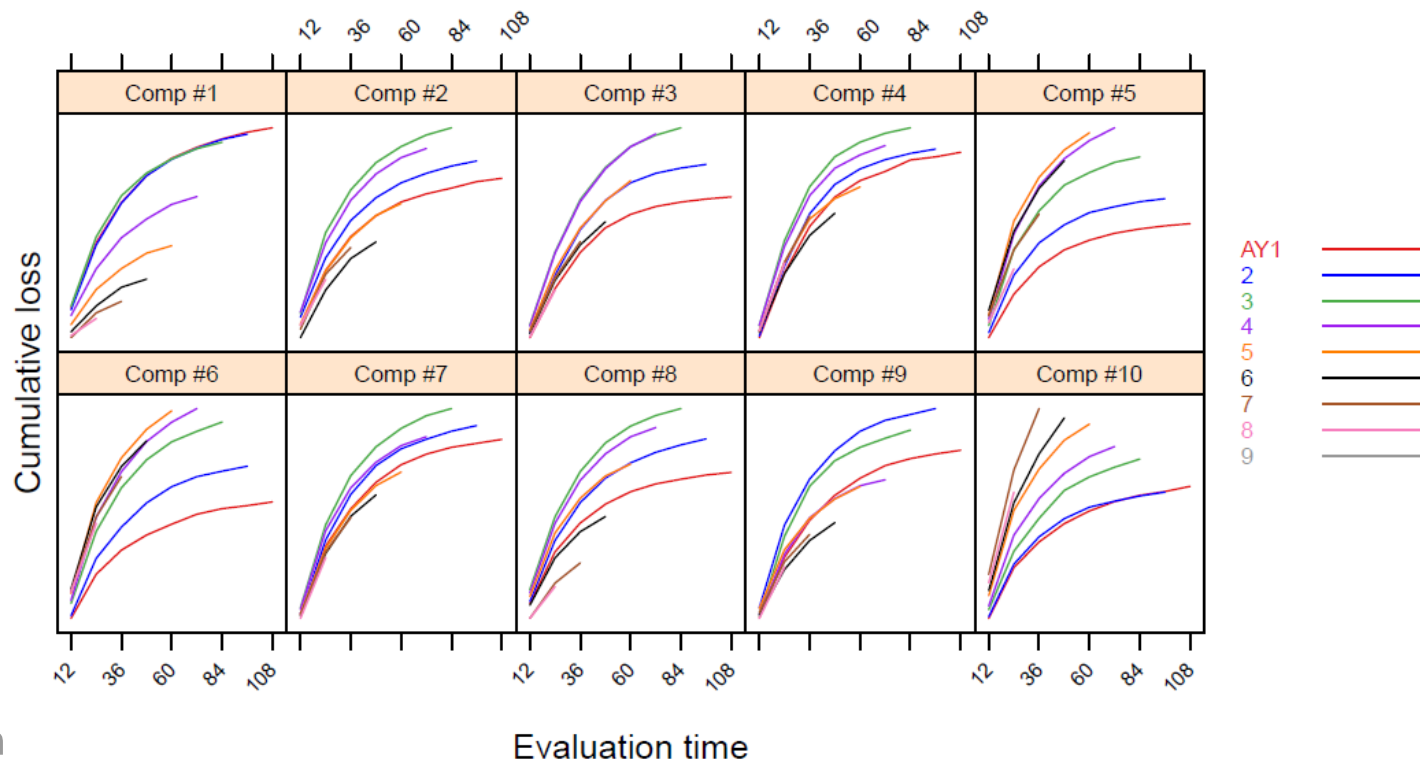
Steps in the Bayesian analysis

- Steps in a Bayesian analysis
 - Setting up the probability model
 - Specify the full distribution of data and the priors
 - Prior distribution could be either informative or non-informative, but need to result in a proper joint density
 - Computation and inference
 - Usually need to use sampling method to simulate values from the posterior distribution
 - Model checking
 - Residual plot
 - Out-of-Sample validation
 - Sensitivity analysis of prior distribution



Visualization of data

- Workers' Comp Schedule P data (1988-1997) from 10 large companies
- Use only 9 years' data, put the 10th year as hold-out validation set





Probability model

- We use the Log-Normal distribution to reflect the skewness and ensure that cumulative losses are positive
- We use a nonlinear mean structure with the log-logistic growth curve: $t^\omega / (t^\omega + \theta^\omega)$

Expected cumulative loss = premium * expected loss ratio * expected emergence
- We use an auto-correlated process along the development for forecasting
- We build a multi-level structure to allow the expected ultimate loss ratios to vary by accident year and company:
 - In one company, loss ratios from different years follow the same distribution with a mean of company-level loss ratio
 - Different company-level average loss ratios follow the same distribution with a mean of the industry-level loss ratio
- Growth curve is assumed to be the same within one company, but vary across companies, arising from the same industry average growth curve
- Assign non-informative priors to complete model specification



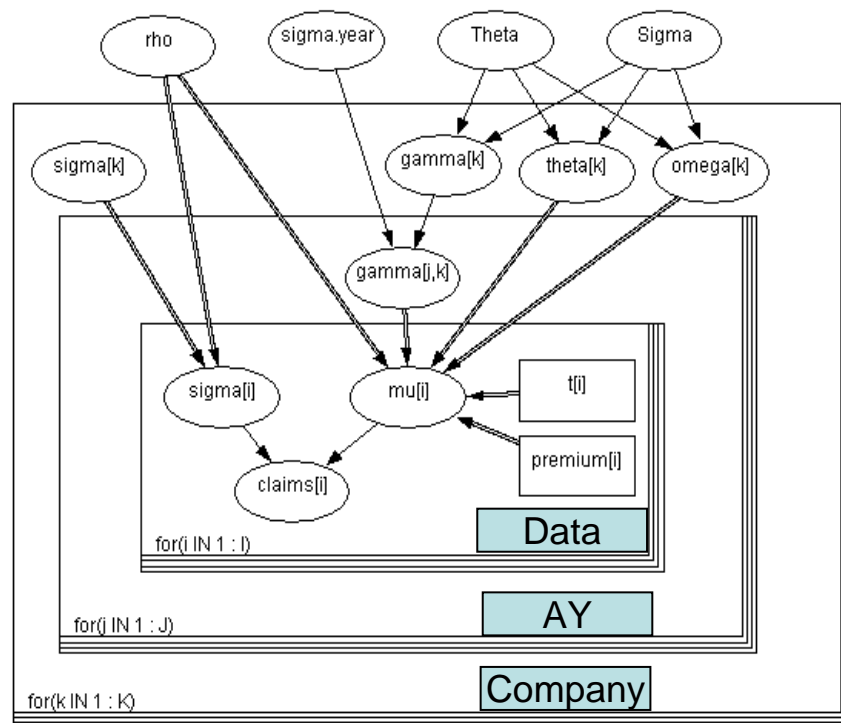
Visualization of the model

$$\text{Mean Structure} = \begin{cases} \log y_{ik}(t_j) & = \log \mu_{ik}(t_j) + \epsilon_{ik}(t_j), \\ \mu_{ik}(t_j) & = p_{ik} \cdot \gamma_{ik} \cdot G(t_j; \Theta_k), \\ G(t_j; \omega_k, \theta_k) & = \frac{t_j^{\omega_k}}{t_j^{\omega_k} + \theta_k^{\omega_k}}. \end{cases}$$

$$\text{Data distribution} = \begin{cases} \epsilon_{ik}(t_j) & = \rho \cdot \epsilon_{ik}(t_{j-1}) + \delta_{ik}(t_j), \\ \delta_{ik}(t_j) & \sim N[0, \sigma_k^2 \cdot (1 - \rho^2)], \\ \epsilon_{ik}(t_0) & \sim N(0, \sigma_k^2). \end{cases}$$

$$\text{Multi-level model} = \begin{cases} \log \gamma_{ik} & \sim N(\log \gamma_k, \sigma_{\gamma_{\text{year}}}^2) \quad \forall i, \\ \log(\gamma_k, \omega_k, \theta_k)' & \sim N[\log(\gamma, \omega, \theta)', \Sigma] \quad \forall k. \end{cases}$$

$$\text{Prior distribution} = \begin{cases} \log \gamma & \sim N(0, 100^2), \\ \log \omega & \sim N(0, 100^2), \\ \log \theta & \sim N(0, 100^2), \\ \sigma_\gamma & \sim U(0, 100), \\ \Sigma & \sim \text{Inv-Wishart}_3(I). \end{cases}$$



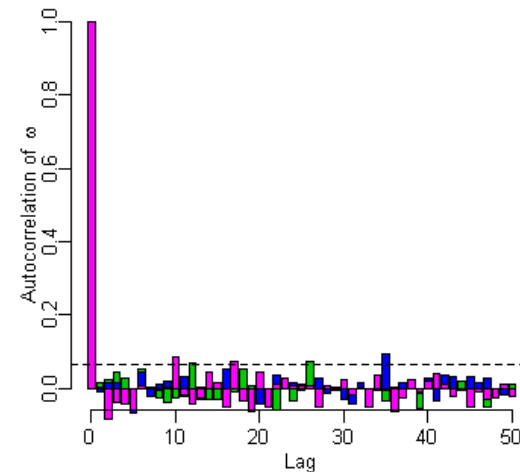
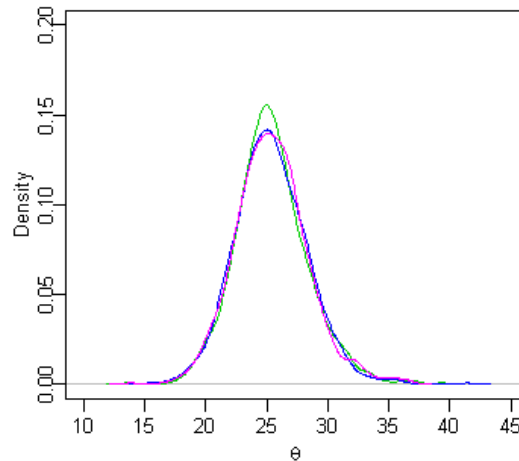
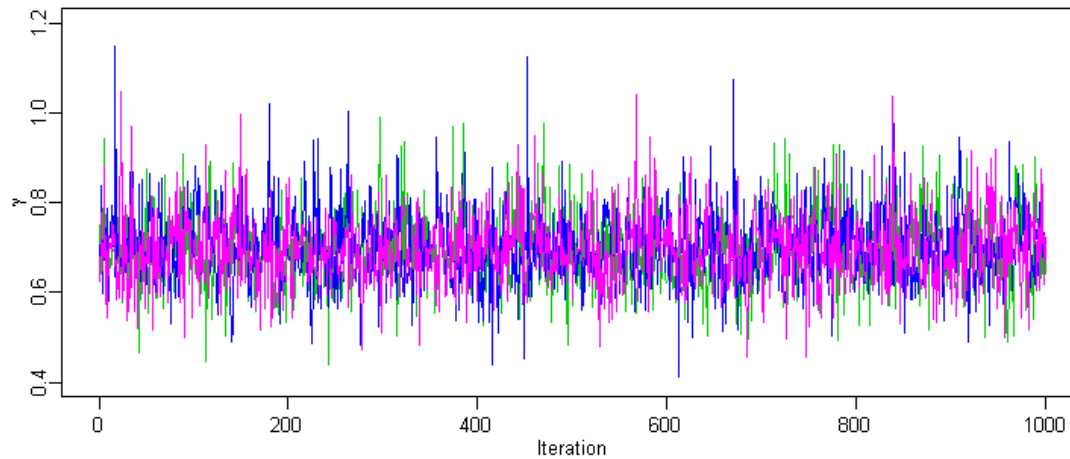


Computation and model estimation

- Such a specification does not result in a closed-form posterior distribution
- Must resort to sampling method to simulate the distribution
- We use Markov Chain Monte Carlo algorithms
 - Developed in the 50s, but became popular in early 90s
 - The software WinBUGS implements the MCMC method
 - Always need to check the convergence of the MCMC algorithm
 - Trajectory plot
 - Density plot
 - Autocorrelation plot

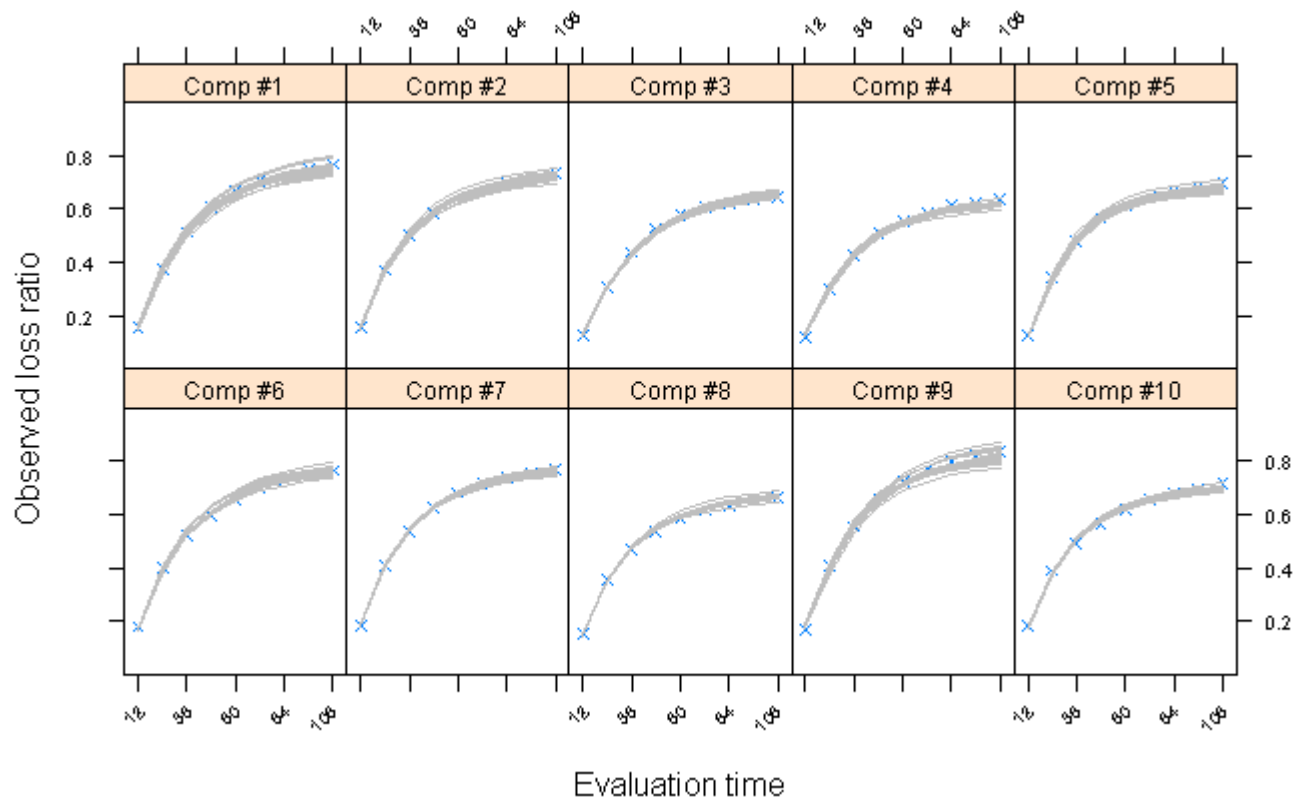


Checking convergence of the Markov chain





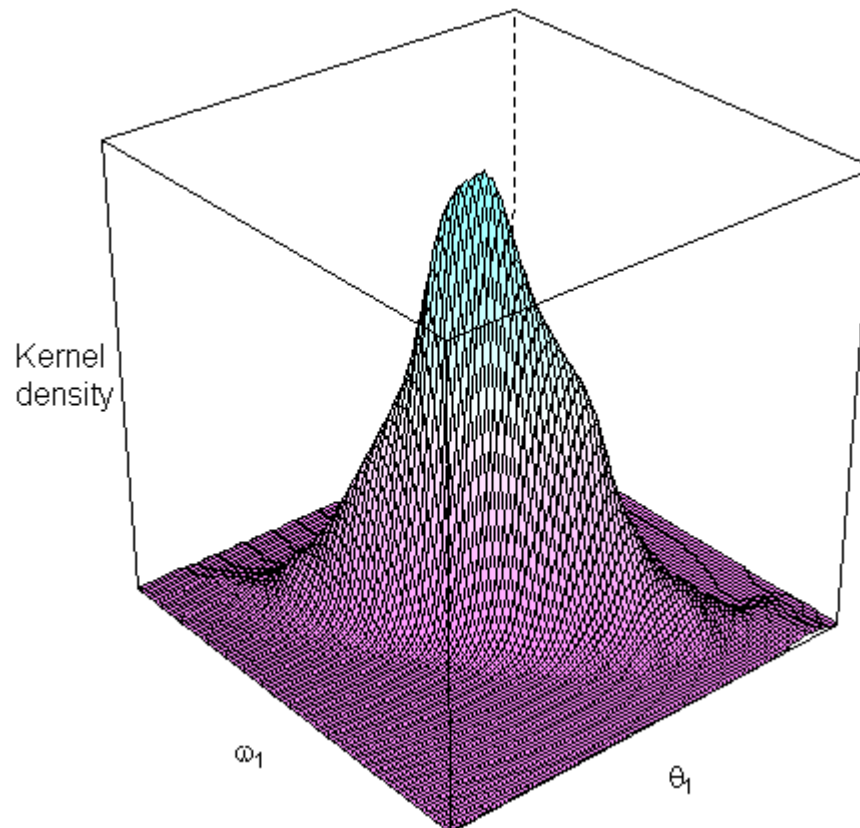
Fitted curves for the first accident year





Joint distribution of growth parameters

Density plot of the growth parameters
--company #1





Estimation of loss ratios

- Industry average loss ratio is 0.693 [0.644, 0.748]
- Variations across company is about twice as large as those across accident years



Comp #1



Comp #2



Comp #3



Comp #4



Comp #5



Comp #6



Comp #7



Comp #8



Comp #9



Comp #10





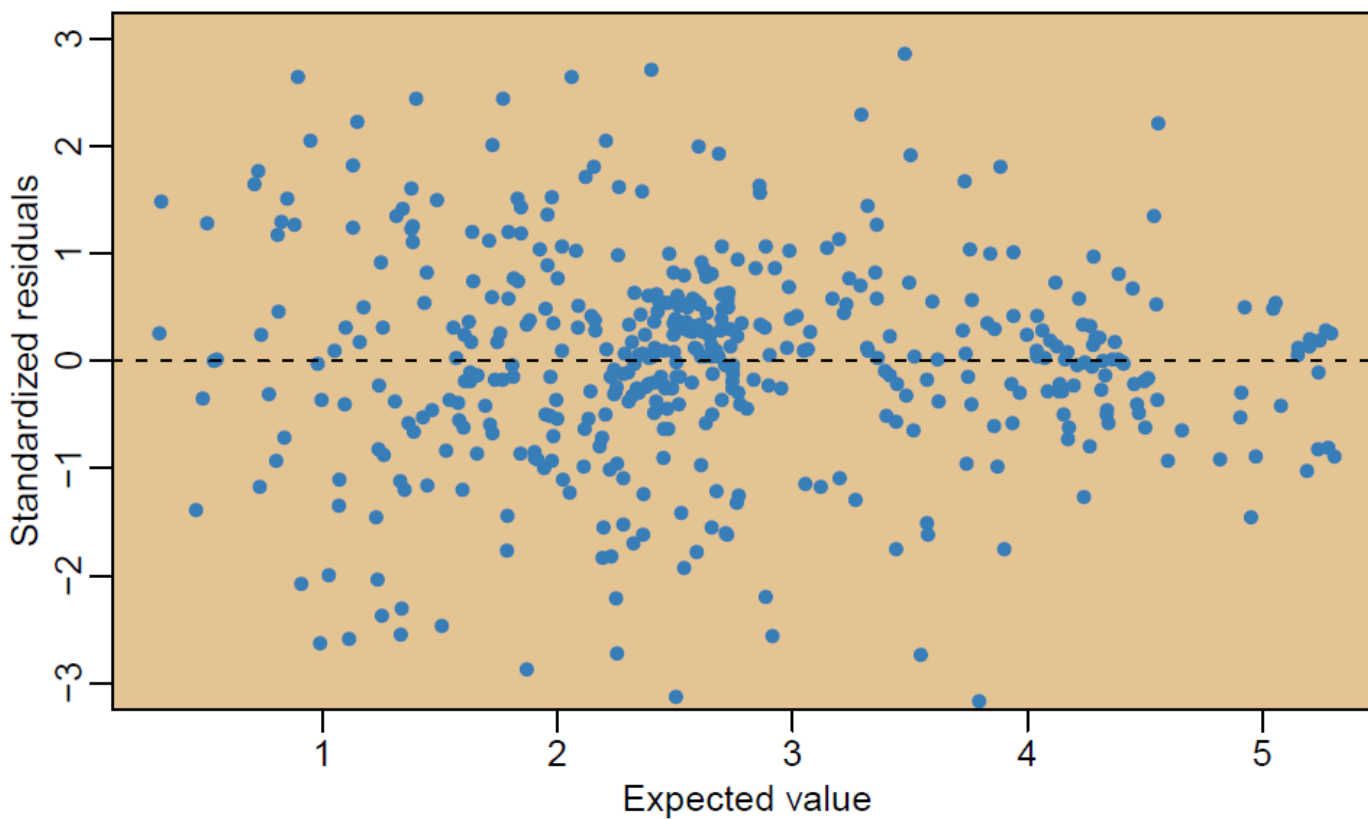
Loss reserve estimation results

- Autocorrelation is about 0.479 [0.445, 0.511]
- Industrial average emergence percentage at 108 months is about 93.5%
- Bayesian reserves projected to ultimate are greater than the GLM estimates projected to 108 months, by a factor about 1.4.

| Company | Estimate at ultimate | | | Estimate at the end of the 9th year | | | |
|---------|----------------------|----------|-----------------|-------------------------------------|----------|---------|----------|
| | Bayesian | | | Bayesian | | GLM-ODP | |
| | Reserve | Pred Err | 50% Interval | Reserve | Pred Err | Reserve | Pred Err |
| 1 | 260.98 | 46.84 | (230.80,292.54) | 170.33 | 25.98 | 155.99 | 10.90 |
| 2 | 173.13 | 22.00 | (159.37,188.60) | 136.20 | 15.13 | 139.63 | 7.11 |
| 3 | 216.19 | 13.95 | (206.70,224.83) | 151.82 | 9.01 | 130.71 | 4.53 |
| 4 | 81.95 | 7.39 | (77.17,87.14) | 63.28 | 4.80 | 54.69 | 3.46 |
| 5 | 44.60 | 6.69 | (40.33,49.21) | 37.95 | 5.14 | 33.56 | 2.12 |
| 6 | 48.86 | 5.27 | (45.48,52.41) | 38.31 | 3.97 | 37.00 | 2.05 |
| 7 | 34.45 | 2.19 | (33.03,35.90) | 26.21 | 1.49 | 25.11 | 0.91 |
| 8 | 22.91 | 2.06 | (21.62,24.32) | 16.46 | 1.37 | 16.83 | 0.72 |
| 9 | 30.66 | 5.62 | (27.11,34.42) | 22.58 | 3.22 | 18.39 | 1.52 |
| 10 | 19.88 | 1.35 | (18.94,20.80) | 15.47 | 0.91 | 17.71 | 0.68 |

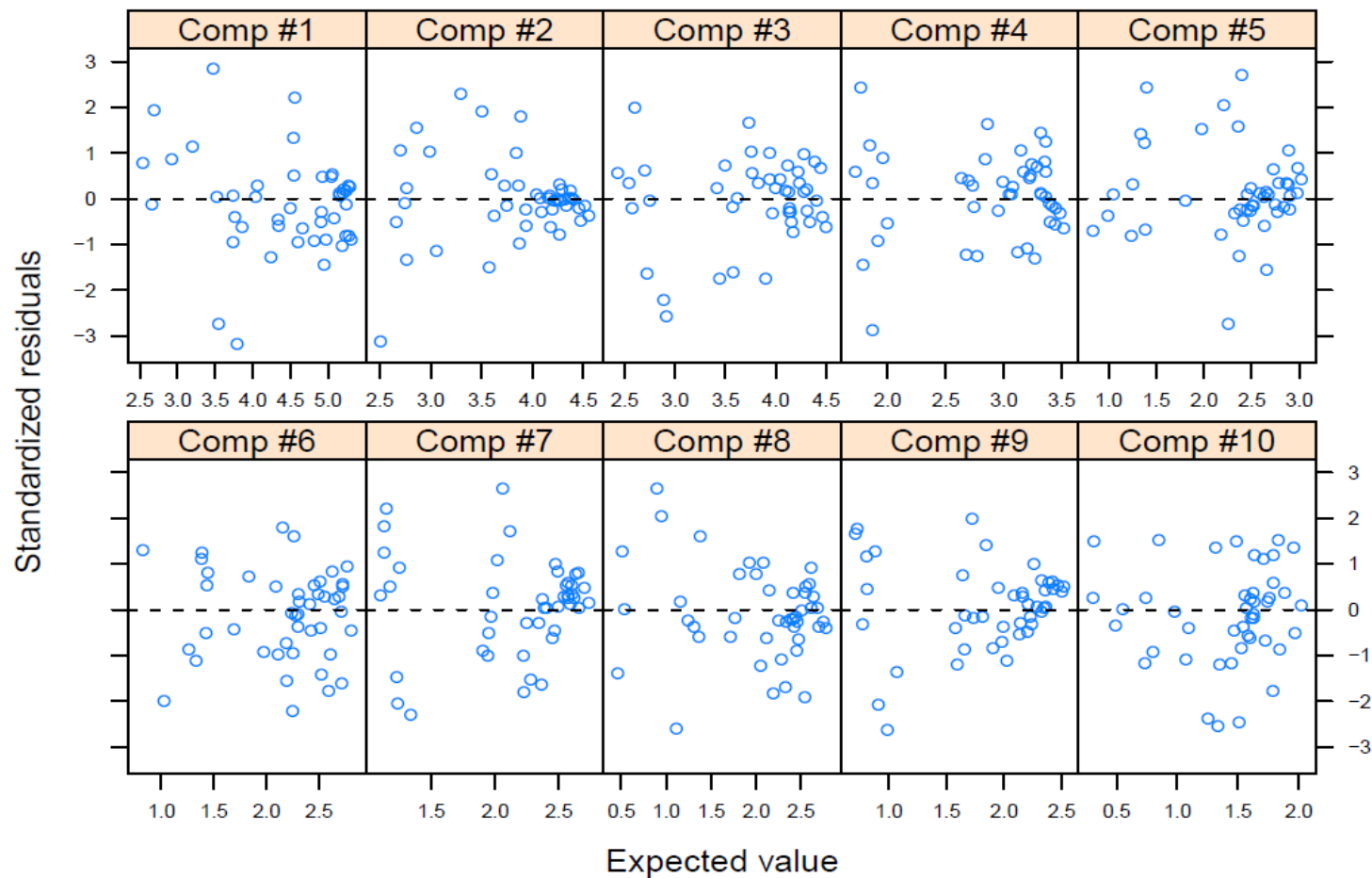


Residual Plot





Residual plot by company





Out-of-Sample test

- We use only 9 years of data to train the model, and validate on the 10th year
 - Note that this is the cash flow of the coming calendar year
 - Policies written in the past
 - Policies to be written in the coming year (need an estimated premium)
- For 4 companies, we also have observed data for the bottom right part
- The coverage rates of the 50% and 95% intervals in the two validation sets are

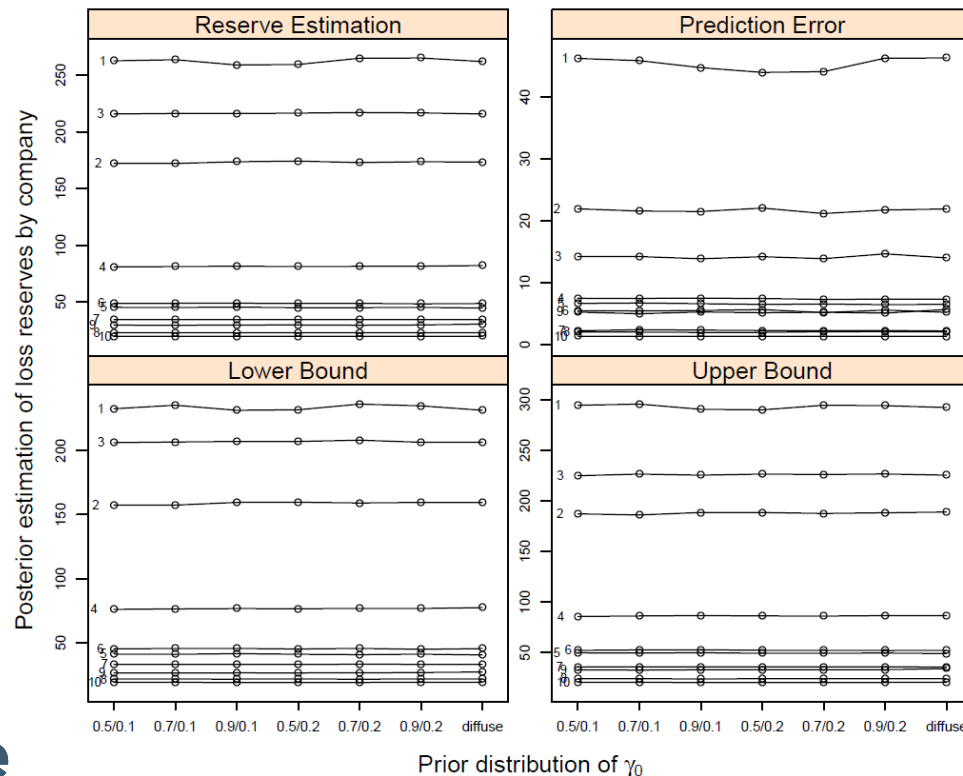
| | 50% Interval | 95% Interval |
|-------|--------------|--------------|
| Set 1 | 57% | 95% |
| Set 2 | 40% | 81% |

- The model performs fairly well overall, but long-term prediction is a little under expectation



Sensitivity analysis

- Change the prior distribution of the industry-level loss ratio to more realistic distributions
- 6 scenarios: Gamma distribution with mean 0.5, 0.7 and 0.9, variance 0.1 and 0.2, respectively





Discussion of the model

- The model used in this analysis provides solutions to many existing challenges
- The model can be further improved:
 - Inflation can be readily included with an appropriate model
 - Prior information can be incorporated on the accident-year or company level
 - Build in more hierarchies: states, lines of business, etc...
 - Include triangles that have more loss history to stabilize extrapolation
- For future research:
 - How to pick the form of the nonlinear pattern?
 - Include multiple lines of business with copula



Summary

- Introduced Bayesian hierarchical model as a full probability model that allows pooling of information and inputs of expert opinion
- Illustrated application of the Bayesian model in insurance with a case study of forecasting loss payments in loss reserving using data from multiple companies
- The application of Bayesian model in insurance is intuitive and promising. I hope more people will start exploiting it and applying it to their work.
- You may download this presentation, the paper and code from my website:

<http://www.actuaryzhang.com/publication/publication.html> ;

Or contact me at: Yanwei.Zhang@cna.com



Reference

- Bühlmann and Gisler (2005). A Course in Credibility Theory and its Applications.
- Clark D. R. (2003). LDF Curve-Fitting and Stochastic Reserving: A Maximum Likelihood Approach. Available at <http://www.casact.org/pubs/forum/03orum/03041.pdf>.
- Guszczka J. (2008). Hierarchical Growth Curve Models for Loss Reserving. Available at <http://www.casact.org/pubs/forum/08orum/7Guszczka.pdf>.
- Klugman S (1987). Credibility for Classification Ratemaking via The Hierarchical Normal Linear Model. Available at <http://www.casact.org/pubs/proceed/proceed87/87272.pdf>.
- Scollnik D. P. M. (2001). Actuarial Modeling with MCMC and BUGS, North American Actuarial Journal 5(2): 96-124.
- Zhang et al. (2010). A Bayesian Nonlinear Model for Forecasting Insurance Loss Payments. Available at <http://www.actuaryzhang.com/publication/bayesianNonlinear.pdf>.



Questions?



Appendix

WinBUGS in Excel



WinBUGS

- BUGS (**B**ayesian inference **U**sing **G**ibbs **S**ampling) was developed by the MRC Biostatistics Unit, and it has a number of versions. WinBUGS is one of them.
- We can work directly in WinBUGS, but better to submit batch run from other software
 - R: package R2WinBUGS
 - SAS: macro %WINBUGSIO
 - Excel: add-in BugsXLA
- R is most handy when working with WinBUGS, but we will focus on Excel here
- The excel add-in BugsXLA is developed by Phil Woodward, and provides a great user interface to work with WinBUGS
- It allows the specification of typical Bayesian hierarchical models, but enhancement is needed to fit more complicated and customized models
- I will illustrate this using the simple Workers' Comp Frequency model



BugsXLA

- Download and install WinBUGS at <http://www.mrc-bsu.cam.ac.uk/bugs/>
- Download and install the Excel add-in BugsXLA at <http://www.axrf86.dsl.pipex.com/>
- Put the data into long format

| Payroll | Claims | Group |
|---------|--------|-------|
| 280 | 9 | 1 |
| 320 | 7 | 1 |
| 265 | 6 | 1 |
| 340 | 13 | 1 |
| 260 | 6 | 2 |
| 275 | 4 | 2 |
| 240 | 2 | 2 |
| 265 | 8 | 2 |
| 145 | 8 | 3 |
| 120 | 3 | 3 |
| 105 | 4 | 3 |



BugsXLA

- Click the "Bayesian analysis" button
- Specify input data
- Specify categorical variables
- In the new window, move the variable "Group" to the "FACTORS" column
- Can specify the levels and the ordering with "Edit Factor Levels"

The screenshot shows the BugsXLA software interface. A data table is visible with columns C, D, and E. The data is as follows:

| Payroll | Claims | Group |
|---------|--------|-------|
| 280 | 9 | 1 |
| 320 | 7 | 1 |
| 265 | 6 | 1 |
| 340 | 13 | 1 |
| 260 | 6 | 2 |
| 275 | 4 | 2 |
| 240 | 2 | 2 |
| 265 | 8 | 2 |
| 145 | 8 | 3 |
| 120 | 3 | 3 |
| 105 | 4 | 3 |

The **Bayesian Model Specification** dialog box is open, showing the **Data Range** set to `Data!C3:E14`. The **Data is ...** section has **in columns** selected and **Names in first row** checked. A **Set Variable Types** button is highlighted. The **Model** section shows **Distribution** set to **Poisson** and **Link** set to **Log**. The **Response** is **Claims** and the **Offset** is **Payroll**.

The **Create Factors** dialog box is also open. The **FACTORS** list contains **Group**. The **VARIATES** list contains **Payroll** and **Claims**. The **CENSORS** list is empty. The **Edit Factor Levels** button is highlighted.



BugsXLA

- Specify Poisson Distribution for the response variable "Claims"
- Want to use identity link, but the only option is "log"
- But for this simple example, we can just re-parameterize the model
- Put "Payroll" as offset
- Put "Group" as random effect
- We are done specifying the model. Now, click "MCMC Options" to customize simulations

| Payroll | Claims | Group |
|---------|--------|-------|
| 280 | 9 | 1 |
| 320 | 7 | 1 |
| 265 | 6 | 1 |
| 340 | 13 | 1 |
| 260 | 6 | 2 |
| 275 | 4 | 2 |
| 240 | 2 | 2 |
| 265 | 8 | 2 |
| 145 | 8 | 3 |
| 120 | 3 | 3 |
| 105 | 4 | 3 |



BugsXLA

- Burn-in: number of simulations to discard from the beginning
- Samples: number of samples to draw
- Thin: sample every k^{th} simulations
- Chains: number of chains
- Import Stats: summary statistics for the parameters and simulations
- Import Sample: the simulated outcomes for each parameter

The image shows a screenshot of the 'WinBUGS MCMC & Output Options' dialog box. It is divided into several sections:

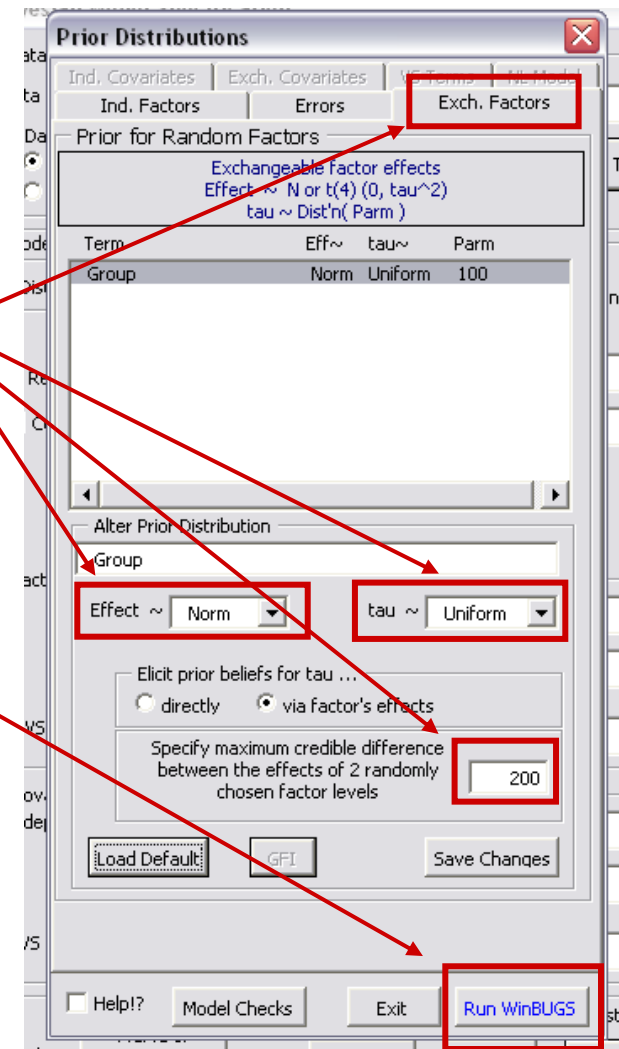
- MCMC Options:** Includes spinners for 'Burn-In' (10000), 'Samples' (5000), 'Thin' (5), and 'Chains' (3). There is a checkbox for 'Over Relax' and a spinner for 'Auto Quit' (No).
- Import Stats for ...:** A list of checkboxes for 'Fixed Terms', 'Covariate Coeffs', 'Intercept', 'Exch. Coeffs', 'Variance Compts', and 'Random Terms'. There are radio buttons for 'Var' and 'SD'.
- Import Samples for ...:** A similar list of checkboxes and radio buttons for 'Fixed Terms', 'Covariate Coeffs', 'Intercept', 'Exch. Coeffs', 'Variance Compts', and 'Random Terms'.
- Buttons:** 'Use Preset Values', 'Make R script', and 'Create BugFolio' are checkboxes. 'Restore Defaults', 'Exit (no save)', and 'Save Changes' are buttons.
- Footer:** Includes 'Help!?', 'MCMC & Output Options', 'Clear Form', 'Exit', 'OK', and 'Predictions or Contrasts'.



BugsXLA

- After clicking "OK" in the "Bayesian analysis" dialog, a "Prior Distribution" dialog pops up
- Change the distribution here so that the group effect is Normally distributed, with a large variance, say, the standard deviation is uniform on (0,100)
- Click "Run WinBUGS"
- Then,

WARNING: Wait until WinBUGS has quit before proceeding.





BugsXLA

- Simulation results are imported
 - Estimation summary
 - Model checks
 - Simulated outcomes
- Calculate the mean for each group
- Plot the result

H13 **=EXP(D5+\$D\$2)**

| | A | B | C | D | E | F | G | H | I | J | K |
|----|---------------|--------------------------------------|--------------------------------------|-------------|----------------|-------------|---------------|--------------|---|---|---------------------|
| 1 | | | Label | Mean | St.Dev. | 2.5% | Median | 97.5% | | | WinBUGS Name |
| 2 | | | CONSTANT | -3.8740 | 2.0170 | -11.0300 | -3.6130 | -0.7491 | | | Beta0 |
| 3 | | | SD(Group) | 2.0290 | 5.1820 | 0.0372 | 0.6123 | 14.7600 | | | sigma.Z[1] |
| 4 | | | Exchangeable (Random) Effects | | | | | | | | |
| 5 | | Group 1 | | 0.3091 | 2.0190 | -2.7600 | 0.0325 | 7.4360 | | | Z.Eff[1,1] |
| 6 | | Group 2 | | -0.0168 | 2.0130 | -3.1830 | -0.2057 | 7.0520 | | | Z.Eff[1,2] |
| 7 | | Group 3 | | 0.5348 | 2.0370 | -2.4490 | 0.1956 | 7.7620 | | | Z.Eff[1,3] |
| 8 | | | | | | | | | | | |
| 9 | Model | [Data!\$C\$3:\$E\$14] | | | | | | | | | |
| 10 | Distribution | Poisson | | | | | | | | | |
| 11 | Link | Log | | | | Separate | Average | Bayesian | | | |
| 12 | Response | Claims | | | 0.0000 | | 0.0268 | | | | |
| 13 | Offset | Payroll | | | 1.0000 | 0.0290 | 0.0268 | 0.0283 | | | |
| 14 | Random | Group | | | 2.0000 | 0.0192 | 0.0268 | 0.0204 | | | |
| 15 | | | | | 3.0000 | 0.0405 | 0.0268 | 0.0355 | | | |
| 16 | Priors | | | | 4.0000 | | 0.0268 | | | | |
| 17 | CONSTANT | N(mu=0, sigma=100) | | | | | | | | | |
| 18 | Group | Norm(0,tau^2); tau ~ Uniform(0, 100) | | | | | | | | | |

