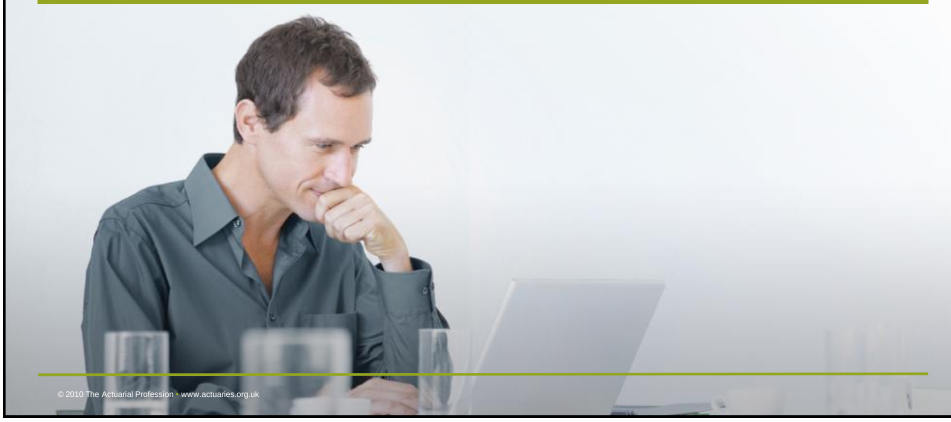
The Actuarial Profession
making financial sense of the future

A statistical basis for claims experience monitoring
Greg Taylor



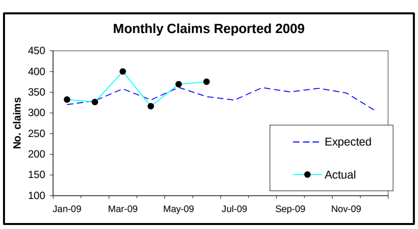
Situation summary

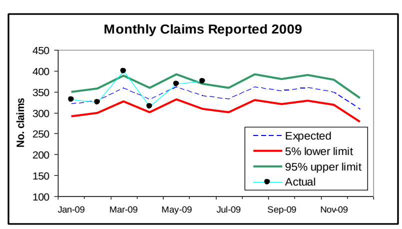
- Predictive model has been formulated and calibrated
  - Forecasts made
  - Subsequent data accumulated
- Data consistent with model or not?
  - Compare data $\boldsymbol{Y}$ with forecasts $\hat{\boldsymbol{Y}}$
- **BUT**
  - Which data?
  - Which forecasts?
  - What is the criterion for consistency?
- No body of theory

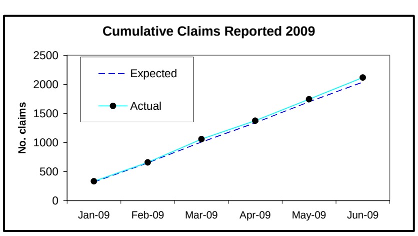
Stochastic and non-stochastic monitoring

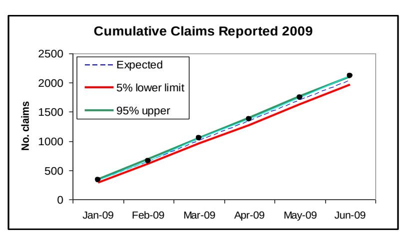
- Stochastic monitoring
  - Form some comparative function of Y and Ŷ
    - e.g. Y - Ŷ, Y / Ŷ, etc **("test statistic")**
  - Generates the probability distribution of the test statistic
  - Uses this distribution to test whether the difference between Y and Ŷ is statistically significant
- Non-stochastic monitoring
  - Same but without the distribution and significance test

Stochastic and non-stochastic monitoring - example









Definition of monitoring

- Non-stochastic monitoring provides no rigorous basis for decision-making
  - We do not concern ourselves with it further
- Formal definition of stochastic monitoring

Consider a stochastic claims model $\mathcal{M}$ of claims experience dependent on parameter vector $\theta = (\theta_1, \theta_2, \ldots, \theta_q)$. Suppose that $\mathcal{M}$ has been calibrated with an estimate $\hat{\theta}$ of θ. The calibrated model generates a forecast of the joint d.f. $G(y;\hat{\theta})$ of a random vector $Y = (Y_1, Y_2, \ldots, Y_r)$ with mean $\hat{Y}(\hat{\theta}) = (\hat{Y}_1(\hat{\theta}), \hat{Y}_2(\hat{\theta}), \ldots, \hat{Y}_r(\hat{\theta}))$. The d.f. G provides the capacity for testing the null hypothesis $\theta = \hat{\theta}$ on the basis of an observation on Y, and such a test will be called a **stochastic claims experience monitoring system**.

Two forms of test statistic

- Micro-testing
  - Tests fine detail of model
- Macro-testing
  - Fit-for-purpose testing
    - Tests model against its primary objectives

$\hat{\theta}$

Micro-testing

- Definition in terms of test of whether $\theta = \hat{\theta}$
  - Natural to formulate test in terms of hypothesis testing
- Let Y denote **new** data (excluding data from which model derived)
  - Assume $Y \sim G(y;\theta)$ [G a d.f.]
  - Null hypothesis H_0: $\theta = \hat{\theta}$
  - Alternative hypothesis H_1: $\theta \neq \hat{\theta}$
  - Formulate likelihood ratio test

Micro-testing a GLM

- Assume

  $$Y = h^{-1}(A\beta) + \varepsilon$$

  h = link function

  A = design matrix

  β = parameter vector (dimension q)

  ε = random error $E[\varepsilon] = 0$
- H_0: $\beta = \hat{\beta}$
- H_1: $\beta \neq \hat{\beta}$

Micro-testing a GLM (2)

- Define $\delta\beta = \beta - \hat{\beta}$ and re-define ($\hat{\beta}$ known)
  - $H_0: \delta\beta = 0$
  - $H_1: \delta\beta \neq 0$
  - $\delta\tilde{\beta}$ = the GLM estimate of $\delta\beta$ on the basis of Y
- The LR test statistic is

$$T = D(y;0) - D(y;\delta\tilde{\beta}) \sim \chi_q^2$$

where $D(y;\delta\tilde{\beta})$ is deviance of the model with estimate $\delta\tilde{\beta}$ of $\delta\beta$

Micro-testing a GLM (3)

- Testing using GLM software
  - Set offset of $\hat{\beta}$ for vector β and fit nothing
    - Deviance is D(y;0)
  - Re-fit all covariates
    - Estimated coefficients are the $\delta\tilde{\beta}$
    - Deviance is $D(y;\delta\tilde{\beta})$
  - LR test statistic of the model as a whole is $T = D(y;0) - D(y;\delta\tilde{\beta}) \sim \chi_q^2$

Micro-testing a GLM (4)

- Can easily adapt this to test a subset S of covariates at once (rather than testing all q covariates at once)
- S can be a singleton
  – Testing a single covariate
- Care needed in testing more than a singleton
  – e.g. the significance level of one or more failures in the test of all q parameters at the 100α% significance level at once is only $100[1-(1-\alpha)^q]\%$

Macro-control

- Need to test accuracy with which model is forecasting its primary target(s)
  – Otherwise the model may perform well in its fine structure but fail in its primary objective
- Identify primary target(s) of model, e.g.
  – **Valuation model**: primary target might be **quantum of liability**
  – **Pricing model**: primary target might be **forecast portfolio profit**

Macro-testing

- Essential idea
  - Calculate primary target
    - On the basis of the original model (before acquisition of data Y)
    - On the basis of the original model re-calibrated to take account of Y (as well as the original data)
      - Hindsight estimate
  - Calculate change between original and hindsight valuations
  - Test this difference for statistical significance

12

Macro-testing – more formally

- Notation
  - Denote the primary targets at time t by L_{it} (i=1,2,… denotes different targets)
  - Denote an estimate of it by

$$\hat{L}_{it|s}(\hat{\beta})$$

Hat denotes estimate

Value of target at time t

Estimated using data to time s

Target estimate based on parameter estimate

13

Macro-testing – more formally (2)

$\hat{L}_{it|t} \quad \hat{\beta}$

- Essential idea
  - Calculate primary target
    - On the basis of the original model (before acquisition of data Y)
    - On the basis of the original model re-calibrated to take account of Y (as well as the original data)
      - Hindsight estimate
  - Calculate change between original and hindsight valuations
  - Test this difference for statistical significance

14

Macro-testing – more formally (3)

$\hat{L}_{it|t} \quad \hat{\beta}$

- Essential idea
  - Calculate primary target
    - On the basis of the original model (before acquisition of data Y)
    - On the basis of the original model re-calibrated to take account of Y (as well as the original data)
      - Hindsight estimate

$\hat{L}_{it|t+1} \quad \hat{\beta} + \delta\tilde{\beta}$

  - Calculate change between original and hindsight valuations
  - Test this difference for statistical significance

15

Macro-testing – more formally (4)

- Essential idea
  - Calculate primary target
    - On the basis of the original model (before acquisition of data Y)
    - On the basis of the original model re-calibrated to take account of Y (as well as the original data)
      - Hindsight estimate

$$\hat{L}_{it|t} \quad \hat{\beta}$$

$$\hat{L}_{it|t+1} \quad \hat{\beta} + \delta\tilde{\beta}$$

  - Calculate change between original and hindsight valuations

$$\Delta_{it} = \hat{L}_{it|t+1} \quad \hat{\beta} + \delta\tilde{\beta} \quad - \hat{L}_{it|t} \quad \hat{\beta}$$

  - Test this difference for statistical significance

Significance testing for macro-control

- Test hypotheses
  - H_0: $E[\Delta_{it} | \mathcal{F}_t] = 0$
  - H_1: $E[\Delta_{it} | \mathcal{F}_t] \neq 0$

  where $\mathcal{F}_t$ denotes data up to time t

  The distribution of $\Delta_{it} | \mathcal{F}_t$ may be estimated from the model at time t
    - Denote d.f. by F
  - Then significance test based on

  Prob $[|\Delta_{it}| > \delta \,|\, \mathcal{F}_t] = F(-\delta) + [1 - F(\delta)]$

Categories of target (at time t)

- **Category I**
  - "Resolved" (estimates replaced by observations) by time t+1
    - **Example**: forecast payments during single post-valuation period
- **Category II**
  - "Unresolved" (still contain estimates) at time t+1
    - **Example**: valuation estimate of liabilities

Categories of target (at time t) (2)

$$\Delta_{it} = \hat{L}_{it|t+1} \quad \hat{\beta} + \delta\tilde{\beta} \quad - \hat{L}_{it|t} \quad \hat{\beta}$$

- For category I targets $\hat{L}_{it|t+1} \quad \hat{\beta} + \delta\tilde{\beta}$ is replaced by an observation (denote ℓ_{it})
- Then test statistic reduces to

$$\Delta_{it} = \ell_{it} - \hat{L}_{it|t} \quad \hat{\beta} \ .$$

  - The usual "actual less forecast"

Numerical examples

Examples - data

- Data from Mack (1993) Astin Bulletin
  - Mortgage guarantee business

Accident year	Cumulative claim payments to end of development year 1	2	3	4	5	6	7	8	9
1	58,046	127,970	476,599	1,027,692	1,360,489	1,647,310	1,819,179	1,906,852	1,950,105
2	24,492	141,767	984,288	2,142,656	2,961,978	3,683,940	4,048,898	4,115,760	
3	32,848	274,682	1,522,637	3,203,427	4,445,927	5,158,781	5,342,585		
4	21,439	529,828	2,900,301	4,999,019	6,460,112	6,853,904			
5	40,397	763,394	2,920,745	4,989,572	5,648,563				
6	90,748	951,994	4,210,640	5,866,482					
7	62,096	868,480	1,954,797						
8	24,983	284,441							
9	13,121								

Examples - data

- Valuation at end of payment year 8
  – Model 8x8 sub-triangle

Accident year	Cumulative claim payments to end of development year 1	2	3	4	5	6	7	8	9
1	58,046	127,970	476,599	1,027,692	1,360,489	1,647,310	1,819,179	1,906,852	1,950,105
2	24,492	141,767	984,288	2,142,656	2,961,978	3,683,940	4,048,898	4,115,760	
3	32,848	274,682	1,522,637	3,203,427	4,445,927	5,158,781	5,342,585		
4	21,439	529,828	2,900,301	4,999,019	6,460,112	6,853,904			
5	40,397	763,394	2,920,745	4,989,572	5,648,563				
6	90,748	951,994	4,210,640	5,866,482					
7	62,096	868,480	1,954,797						
8	24,983	284,441							
9	13,121								

22

Examples - data

- Then monitor 9th diagonal against that model
  – Bottom left cell plays no part

Accident year	Cumulative claim payments to end of development year 1	2	3	4	5	6	7	8	9
1	58,046	127,970	476,599	1,027,692	1,360,489	1,647,310	1,819,179	1,906,852	1,950,105
2	24,492	141,767	984,288	2,142,656	2,961,978	3,683,940	4,048,898	4,115,760	
3	32,848	274,682	1,522,637	3,203,427	4,445,927	5,158,781	5,342,585		
4	21,439	529,828	2,900,301	4,999,019	6,460,112	6,853,904			
5	40,397	763,394	2,920,745	4,989,572	5,648,563				
6	90,748	951,994	4,210,640	5,866,482					
7	62,096	868,480	1,954,797						
8	24,983	284,441							
9	13,121								

23

Valuation model

- Use chain ladder model (as did Mack)
  - Fit to 8x8 sub-triangle using GLM

Accident year	Cumulative claim payments to end of development year 1	2	3	4	5	6	7	8	9
1	58,046	127,970	476,599	1,027,692	1,360,489	1,647,310	1,819,179	1,906,852	1,950,105
2	24,492	141,767	984,288	2,142,656	2,961,978	3,683,940	4,048,898	4,115,760	
3	32,848	274,682	1,522,637	3,203,427	4,445,927	5,158,781	5,342,585		
4	21,439	529,828	2,900,301	4,999,019	6,460,112	6,853,904			
5	40,397	763,394	2,920,745	4,989,572	5,648,563				
6	90,748	951,994	4,210,640	5,866,482					
7	62,096	868,480	1,954,797						
8	24,983	284,441							
9	13,121								
$\hat{f}_j$		**11.1**	**4.09**	**1.71**	**1.28**	**1.14**	**1.069**	**1.026**	

$C_{ij} \sim \text{ODP}\,(C_{i,j-1} f_j\ ,\ \varphi_j)$

f_j age-to-age factors

Monitor valuation model

- Model just 9th diagonal
  - Model as adjustments to the valuation model

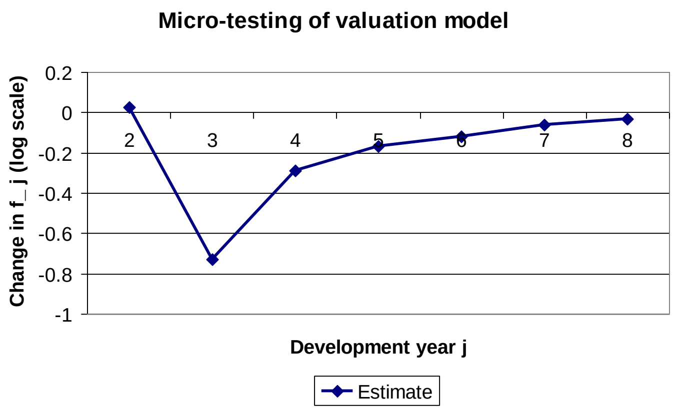
Accident year	Cumulative claim payments to end of development year 1	2	3	4	5	6	7	8	9
1	58,046	127,970	476,599	1,027,692	1,360,489	1,647,310	1,819,179	1,906,852	1,950,105
2	24,492	141,767	984,288	2,142,656	2,961,978	3,683,940	4,048,898	4,115,760	
3	32,848	274,682	1,522,637	3,203,427	4,445,927	5,158,781	5,342,585		
4	21,439	529,828	2,900,301	4,999,019	6,460,112	6,853,904			
5	40,397	763,394	2,920,745	4,989,572	5,648,563				
6	90,748	951,994	4,210,640	5,866,482					
7	62,096	868,480	1,954,797						
8	24,983	284,441							
9	13,121								

$C_{ij} \sim \text{ODP}\,(C_{i,j-1}\hat{f}_j\,(\delta f_j),\ \hat{\varphi}_j)$

δf_j now to be estimated

Should be insignificantly different from 1

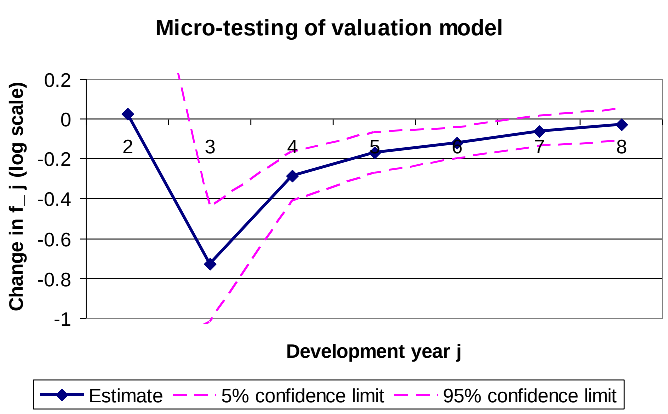
Monitoring – micro-testing



Note immediately that 7 out of 8 increments are negative

Probability = 3%

Monitoring – micro-testing (2)



In 4 out of 8 cases zero lies outside 90% confidence interval for increment in ln f_j

i.e. lies outside 90% confidence interval for δf_j

Macro-testing

Accident year	Loss reserve at end of payment year 8			
	Estimated end year 8	Hindsight estimate	Change	Significance
2	195,131			
3	793,116			
4	2,456,607			
5	4,232,286			
6	10,251,788			
7	13,048,875			
8	4,412,105			
Total	35,389,909			

28

Macro-testing

Accident year	Loss reserve at end of payment year 8			
	Estimated end year 8	Hindsight estimate	Change	Significance
2	195,131	66,862	-66%	
3	793,116	324,500	-59%	
4	2,456,607	1,057,529	-57%	
5	4,232,286	2,066,644	-51%	
6	10,251,788	5,139,865	-50%	
7	13,048,875	4,452,906	-66%	
8	4,412,105	3,143,711	-29%	
Total	35,389,909	16,252,018	-54%	

29

Macro-testing

Accident year	Loss reserve at end of payment year 8			
	Estimated end year 8	Hindsight estimate	Change	Significance
2	195,131	66,862	-66%	27%
3	793,116	324,500	-59%	10%
4	2,456,607	1,057,529	-57%	**1%**
5	4,232,286	2,066,644	-51%	**0%**
6	10,251,788	5,139,865	-50%	**0.0%**
7	13,048,875	4,452,906	-66%	**0.00%**
8	4,412,105	3,143,711	-29%	43%
Total	35,389,909	16,252,018	-54%	**0.00%**

Model fails spectacularly on addition of 9th diagonal

Conclusion

- First recall that all testing in this paper requires a stochastic model
- Monitoring of new claims experience against a pre-existing model has been formulated in rigorous statistical terms
  - A statistical hypothesis that can be tested for significance
- This has enabled both:
  - Micro-testing: the testing of the fine detail of the model
  - Macro-testing: the overall stability of its predictive performance
- A numerical example has illustrated model failure in a convincing manner

Questions or comments?

Expressions of individual views by members of The Actuarial Profession and its staff are encouraged.

The views expressed in this presentation are those of the presenter.