One year uncertainty in claims reserving

Arthur Charpentier

Université Rennes 1 (& Université de Montréal)

arthur.charpentier@univ-rennes1.fr

http ://freakonometrics.blog.free.fr/

Workshop on Insurance Mathematics, Montreal, January 2011.

Agenda of the talk

- **Formalizing claims reserving problem**
- **From Mack to Merz & Wüthrich**
- **From Mack (1993) to Merz & Wüthrich (2009)**
- **Updating Poisson-ODP bootstrap technique**

	one year	ultimate
China ladder	Merz & Wüthrich (2008)	Mack (1993)
GLM+boostrap	×	Hacheleister & Stanard (1975)
		England & Verrall (1999)

'one year horizon for the reserve risk'

AISAM-ACME study on non-life long tail liabilities

Reserve risk and risk margin assessment under Solvency II

17 October 2007

‘one year horizon for the reserve risk’

4 The concept of the one year horizon for the reserve risk

The uncertainty measurement of reserves in the balance sheet (called risk margin in the Solvency II framework) and the reserve risk do not have the same time horizon. It seems important to underline this point because it may be a source of confusion when the calibration is discussed.

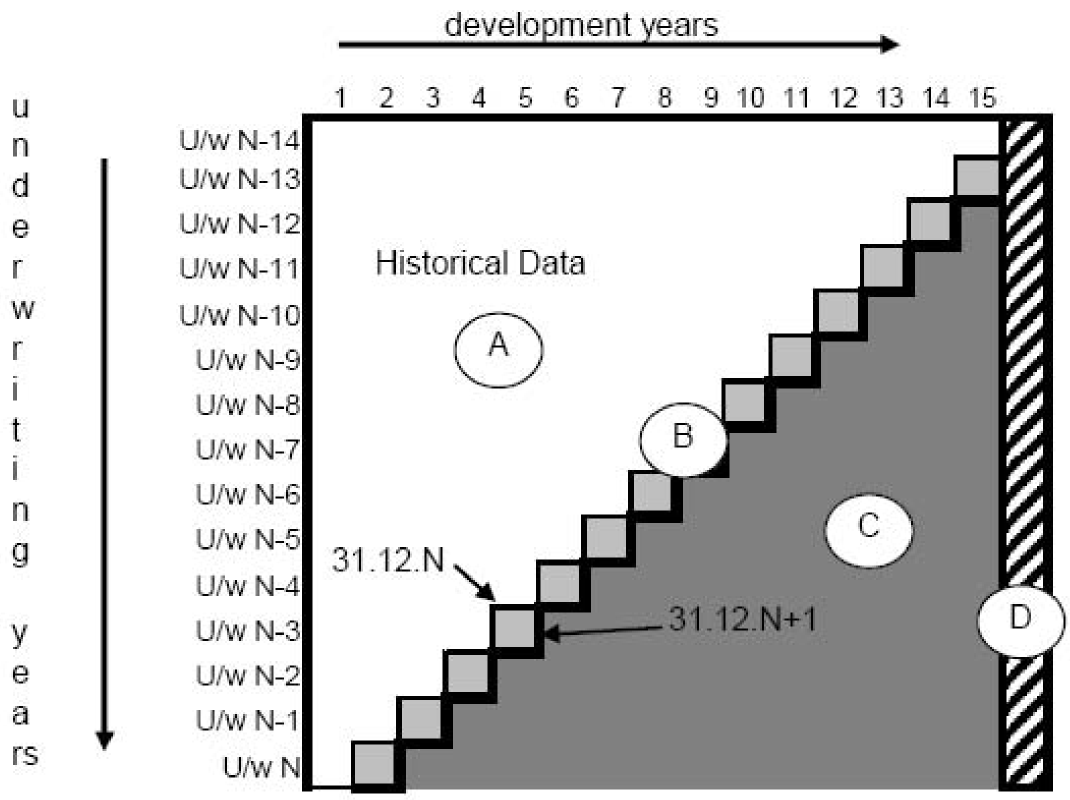
4.1.2 The reserve risk captures uncertainty over a one year period

4.1.2.1 The Solvency II draft Directive framework

The SCR has the following definition[3]:

> *“The SCR corresponds to the economic capital a (re)insurance undertaking needs to hold in order to limit the probability of ruin to 0.5%, i.e. ruin would occur once every 200 years (see Article 100). The SCR is calculated using Value-at-Risk techniques, either in accordance with the standard formula, or using an internal model: all potential losses, including adverse revaluation of assets and liabilities over the next 12 months are to be assessed. The SCR reflects the true risk profile of the undertaking, taking account of all quantifiable risks, as well as the net impact of risk mitigation techniques.”*

'one year horizon for the reserve risk'



'one year horizon for the reserve risk'

CEIOPS
Committee of European Insurance and Occupational Pensions Supervisors

Consultation Paper No. 71

CEIOPS-CP-71-09
2 November 2009

Draft CEIOPS' Advice for Level 2 Implementing Measures on Solvency II:

SCR Standard Formula

Calibration of non-life underwriting risk

'one year horizon for the reserve risk'

Method 4

3.242 This approach is consistent with the undertaking specific estimate assumptions from the Technical Specifications for QIS4 for premium risk.

3.243 This method involves a three stage process:

a. **Involves by undertaking calculating the mean squared error of prediction of the claims development result over the one year.**

- The mean squared errors are calculated using the approach detailed in "Modelling The Claims Development Result For Solvency Purposes" by Michael Merz and Mario V Wuthrich, Casualty Actuarial Society E-Forum, Fall 2008.
- Furthermore, in the claims triangles:
- cumulative payments $C_{i,j}$ in different accident years i are independent
- for each accident year, the cumulative payments $(C_{i,j})_j$ are a Markov process and there are constants f_j and s_j such that $E(C_{i,j}|C_{i,j-1})=f_jC_{i,j-1}$ and $Var(C_{i,j}|C_{i,j-1})=s_j^2C_{i,j-1}$.

1 Formalizing the claims reserving problem

- $X_{i,j}$ denotes incremental payments, with delay j, for claims occurred year i,
- $C_{i,j}$ denotes cumulated payments, with delay j, for claims occurred year i, $C_{i,j} = X_{i,0} + X_{i,1} + \cdots + X_{i,j}$,

	0	1	2	3	4	5
0	3209	1163	39	17	7	21
1	3367	1292	37	24	10	
2	3871	1474	53	22		
3	4239	1678	103			
4	4929	1865				
5	5217					

	0	1	2	3	4	5
0	3209	4372	4411	4428	4435	4456
1	3367	4659	4696	4720	4730	
2	3871	5345	5398	5420		
3	4239	5917	6020			
4	4929	6794				
5	5217					

- $\mathcal{H}_n$ denotes information available at time n,

$$\mathcal{H}_n = \{(C_{i,j}), 0 \leq i+j \leq n\} = \{(X_{i,j}), 0 \leq i+j \leq n\}$$

Actuaries have to predict the total amount of payments for accident year i, i.e.

$$\widehat{C}_{i,n}^{(n-i)} = \mathbb{E}[C_{i,\infty}|\mathcal{H}_n] = \mathbb{E}[C_{i,n}|\mathcal{H}_n]$$

The difference betwwen what should be paid, and what has been paid will be the claims reserve, $\widehat{R}_i = \widehat{C}_{i,n}^{(n-i)} - C_{i,n-i}$.

A classical measure of uncertainty in claims reserving is $\text{mse}[C_{i,n}|\mathcal{F}_{i,n-i}]$ called ultimate uncertainty.

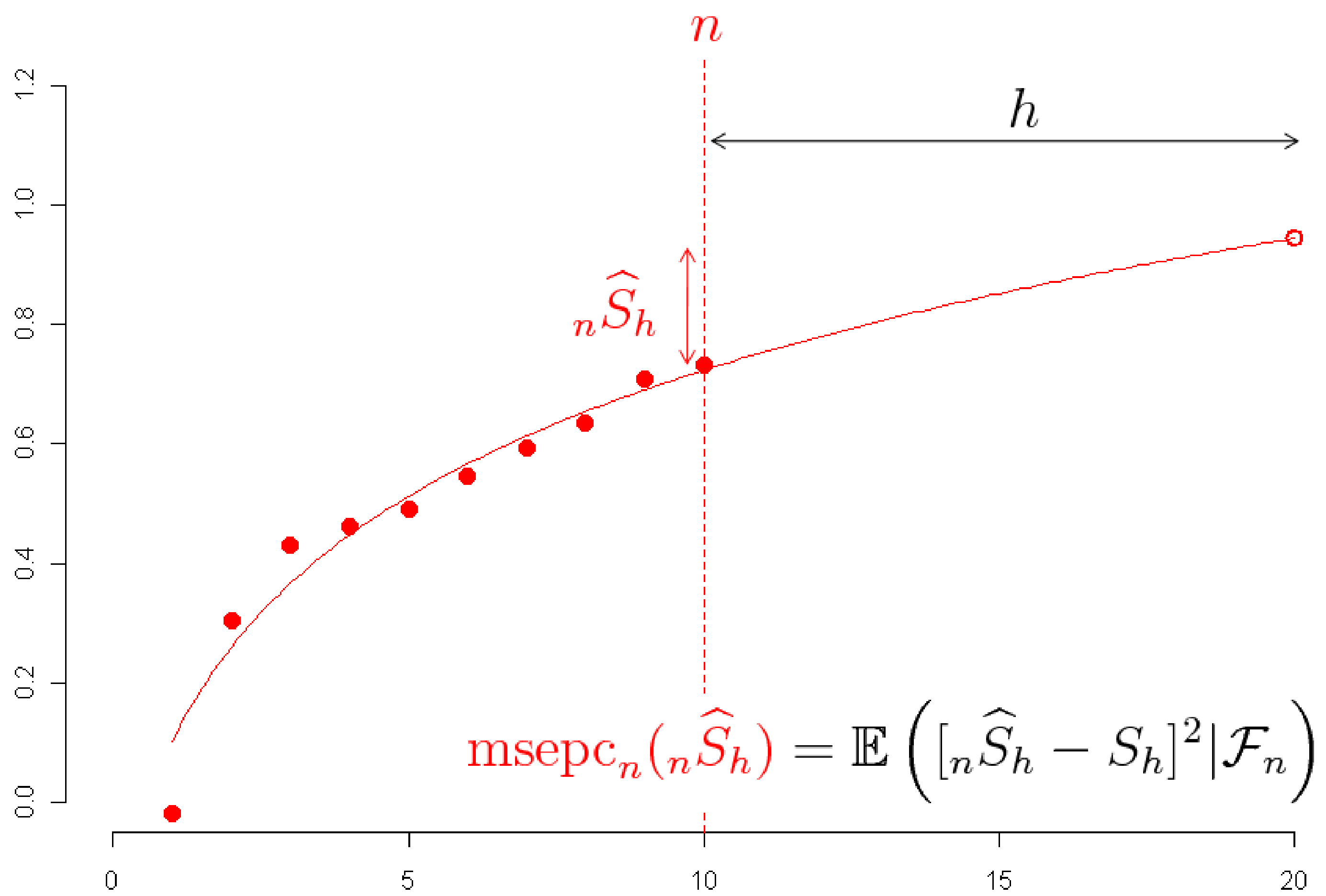
In Solvency II, it is now requiered to quantify one year uncertainy. The starting point is that in one year we will update our prediction

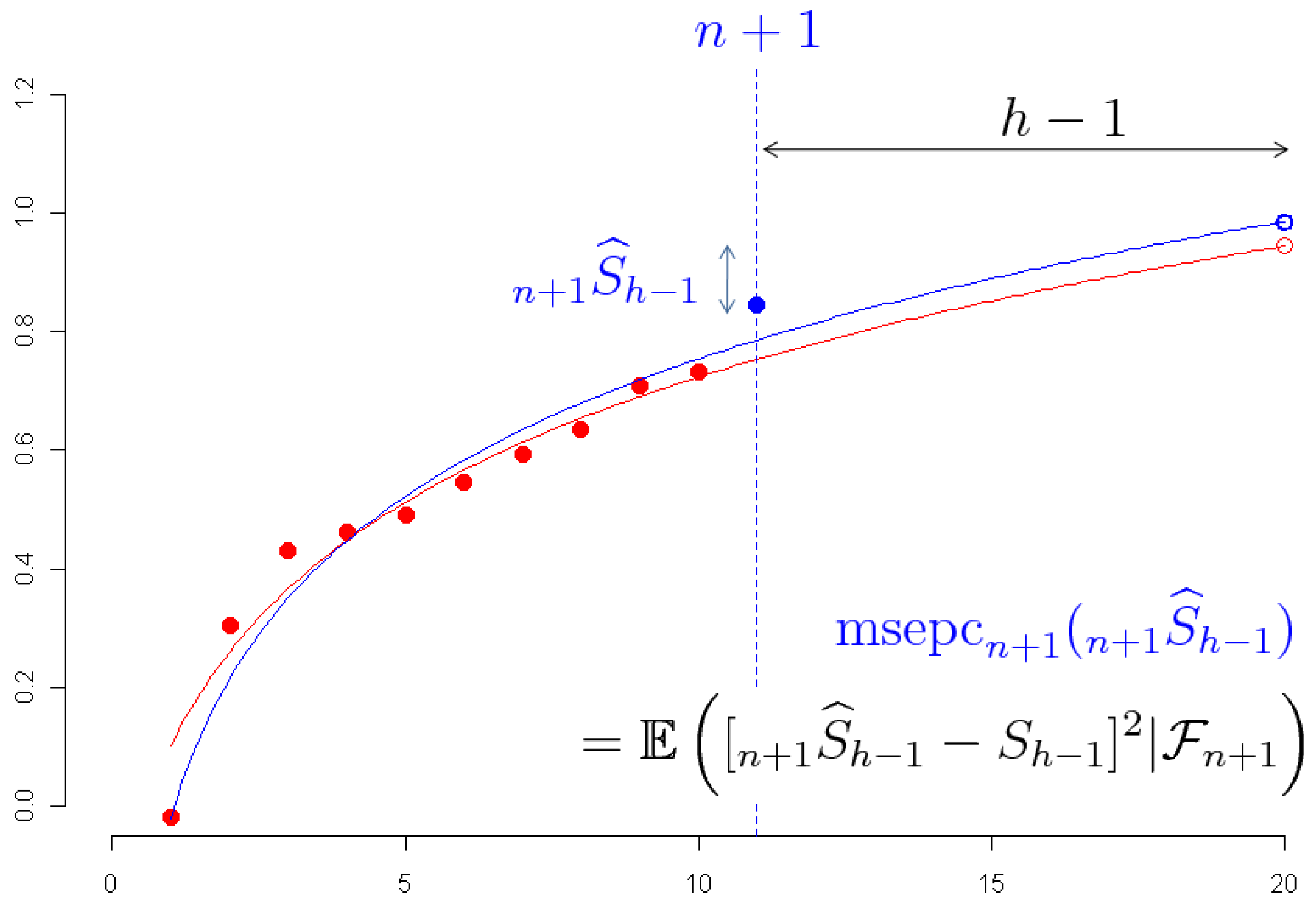
$$\widehat{C}_{i,n}^{(n-i+1)} = \mathbb{E}[C_{i,n}|\mathcal{H}_{n+1}]$$

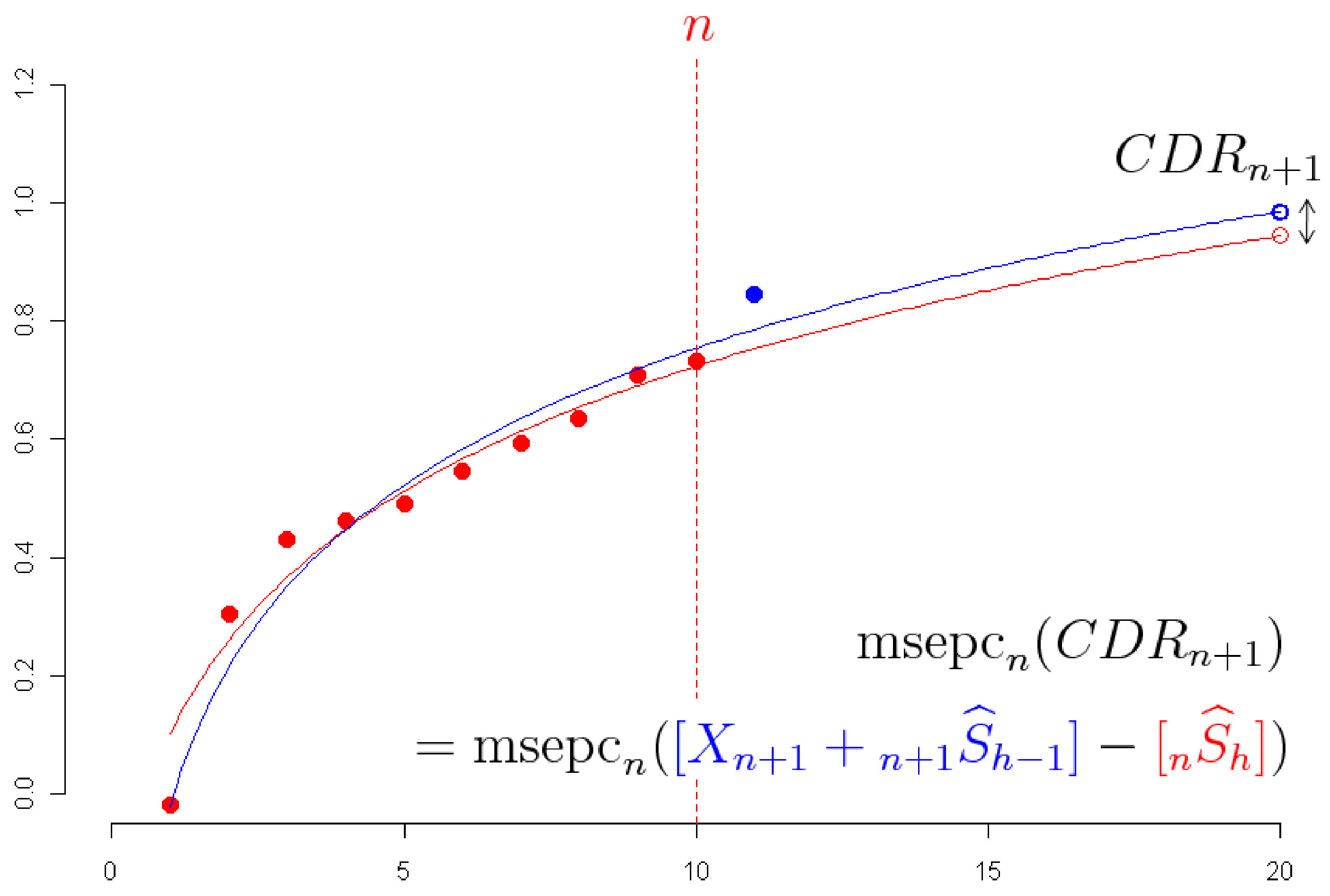
so that there is a change (compared with now)

$$\Delta_i^n = CDR_{i,n} = \widehat{C}_{i,n}^{(n-i+1)} - \widehat{C}_{i,n}^{(n-i)}.$$

If that difference is positive, the insurance will experience a mali (the insurer will pay more than what was expected), and a boni if the difference is negative. Note that $\mathbb{E}[\Delta_i^n|\mathcal{H}_n] = 0$. In Solvency II, actuaries have to calculate uncertainty associated to Δ_i^n, i.e. $\text{mse}[\Delta_i^n|\mathcal{H}_n]$.


n
h
${}_n\widehat{S}_h$
$\text{msepc}_n({}_n\widehat{S}_h) = \mathbb{E}\left([{}_n\widehat{S}_h - S_h]^2|\mathcal{F}_n\right)$
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
5
10
15
20


$n+1$
$h-1$
${}_{n+1}\widehat{S}_{h-1}$
$\text{msepc}_{n+1}({}_{n+1}\widehat{S}_{h-1})$
$= \mathbb{E}\left(\left[{}_{n+1}\widehat{S}_{h-1} - S_{h-1}\right]^2 \middle| \mathcal{F}_{n+1}\right)$
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
5
10
15
20


n
1.2
1.0
0.8
0.6
0.4
0.2
0.0
CDR_{n+1}
$\text{msepc}_n(CDR_{n+1})$
$= \text{msepc}_n([X_{n+1} + {}_{n+1}\widehat{S}_{h-1}] - [{}_n\widehat{S}_h])$
0
5
10
15
20

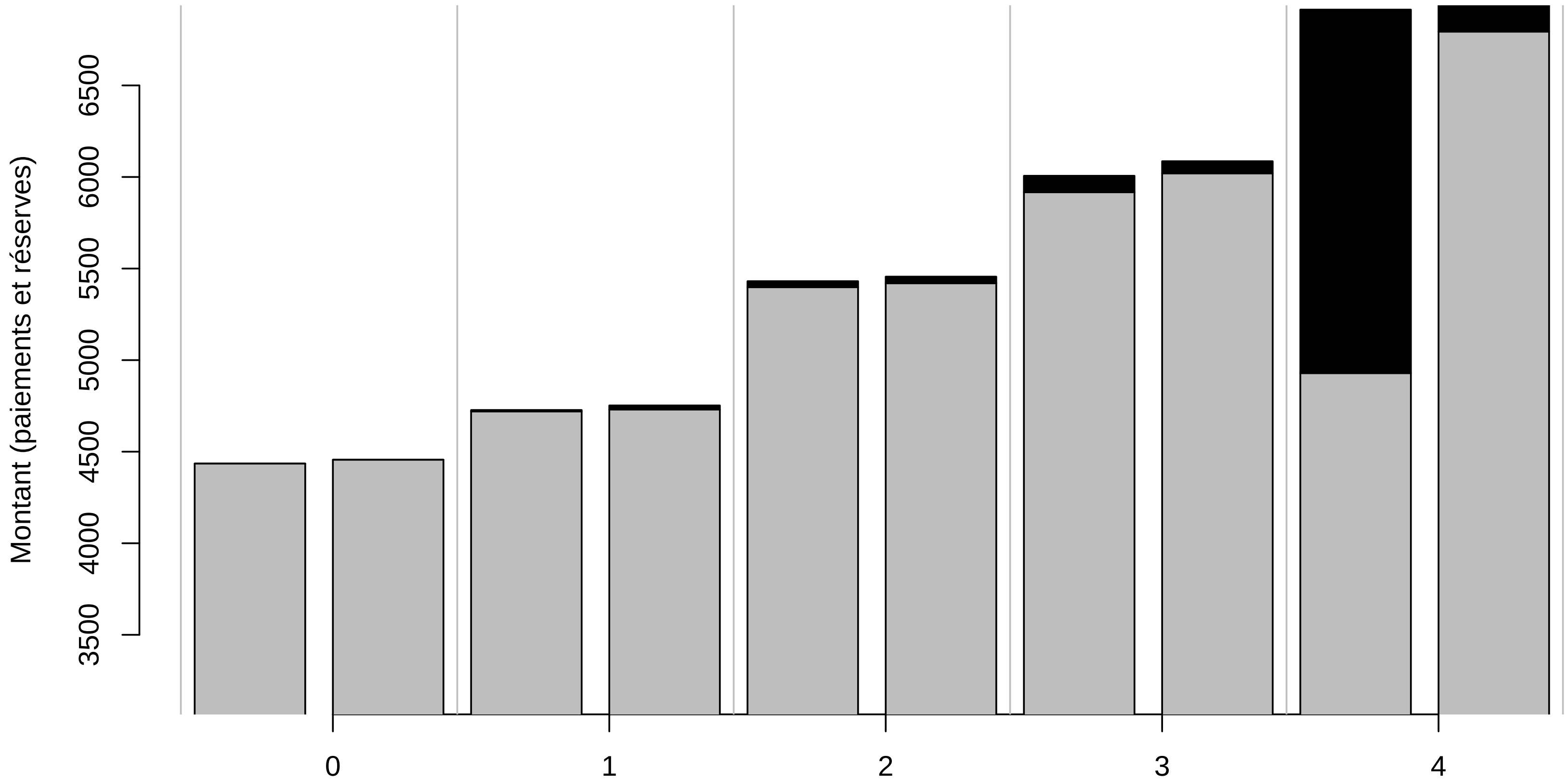


Figure 1 – Estimation of total payments $\widehat{C}_{i,n}$ two consecutive years.

2 Chain Ladder and claims reserving

A standard approach is to assume that

$$C_{i,j+1} = \lambda_j \cdot C_{i,j} \text{ for all } i, j = 1, \cdots, n.$$

A natural estimator for λ_j, based on past experience is

$$\widehat{\lambda}_j = \frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}} \text{ for all } j = 1, \cdots, n-1.$$

Future payments can be predicted as follows,

$$\widehat{C}_{i,j} = \left[\widehat{\lambda}_{n-i} \ldots \widehat{\lambda}_{j-1}\right] C_{i,n-i}.$$

	0	1	2	3	4	5
λ_j	1,38093	1,01143	1,00434	1,00186	1,00474	1,0000

TABLE 1 – Development factors, $\widehat{\boldsymbol{\lambda}} = (\widehat{\lambda}_i)$.

Observe that

$$\widehat{\lambda}_j = \sum_{i=1}^{n-j} \omega_{i,j} \lambda_{i,j} \text{ où } \omega_{i,j} = \frac{C_{i,j}}{\sum_{i=1}^{n-j} C_{i,j}} \text{ et } \lambda_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}.$$

i.e. it is a weighted least squares estimator

$$\widehat{\lambda}_j = \underset{\lambda \in \mathbb{R}}{\text{argmin}} \left\{ \sum_{i=1}^{n-j} C_{i,j} \left[\lambda - \frac{C_{i,j+1}}{C_{i,j}} \right]^2 \right\},$$

or

$$\widehat{\lambda}_j = \underset{\lambda \in \mathbb{R}}{\text{argmin}} \left\{ \sum_{i=1}^{n-j} \frac{1}{C_{i,j}} \left[\lambda C_{i,j} - C_{i,j+1} \right]^2 \right\}.$$

	0	1	2	3	4	5
1	3209	4372	4411	4428	4435	4456
2	3367	4659	4696	4720	4730	4752.4
3	3871	5345	5398	5420	5430.1	5455.8
4	4239	5917	6020	6046.15	6057.4	6086.1
5	4929	6794	6871.7	6901.5	6914.3	6947.1
6	5217	7204.3	7286.7	7318.3	7331.9	7366.7

Table 2 – Cumulated payments $\boldsymbol{C} = (C_{i,j})_{i+j\leq n}$ and future projected payments $\widehat{\boldsymbol{C}} = (\widehat{C}_{i,j})_{i+j>n}$, $\widehat{C}_{i,j} = \widehat{\lambda}_{j-1}\cdots\widehat{\lambda}_{n-i}C_{i,n-i}$.

Proposition1

If there are $\boldsymbol{A} = (A_0, \cdots, A_n)$ and $\boldsymbol{B} = (B_0, \cdots, B_n)$, with $B_0 + \cdots + B_n = 1$, such that

$$\sum_{i=1}^{n-j} A_i B_j = \sum_{i=1}^{n-j} Y_{i,j} \text{ for all } j, \text{ and } \sum_{j=0}^{n-i} A_i B_j = \sum_{j=0}^{n-i} Y_{i,j} \text{ for all } i,$$

then

$$\widehat{C}_{i,n} = A_i = C_{i,n-i} \cdot \prod_{k=n-i}^{n-1} \lambda_k$$

where

$$B_k = \prod_{j=k}^{n-1} \frac{1}{\lambda_j} - \prod_{j=k-1}^{n-1} \frac{1}{\lambda_j}, \text{ avec } B_0 = \prod_{j=k}^{n-1} \frac{1}{\lambda_j}.$$

cf. margin method (Bailey (1963)), and Poisson regression.

3 From Mack to Merz & Wüthrich

3.1 Quantifying uncertainty

We wish to compare $\widehat{R}$ and R (where R is random - and unknown). The mse of prediction can be written

$$\mathbb{E}([\widehat{R}-R]^2) \approx \underbrace{\mathbb{E}([\widehat{R}-\mathbb{E}(R)]^2)}_{\text{mse}(\widehat{R})} + \underbrace{\mathbb{E}([R-\mathbb{E}(R)]^2)}_{\text{Var}(R)}$$

with a first order approximation.

Actually, what we are looking for is the msep conditional to the information we have

$$\text{msep}_n(\widehat{R}) = \mathbb{E}([\widehat{R}-R]^2|\mathcal{H}_n).$$

3.2 Mack's model

Mack (1993) proposed a probabilistic model to justify Chain-Ladder. Assume that $(C_{i,j})_{j\geq 0}$ is a Markovian process, and there are $\boldsymbol{\lambda}=(\lambda_j)$ and $\boldsymbol{\sigma}=(\sigma_j^2)$ such that

$$\begin{cases} \mathbb{E}(C_{i,j+1}|\mathcal{H}_{i+j})=\mathbb{E}(C_{i,j+1}|C_{i,j})=\lambda_j\cdot C_{i,j} \\ \text{Var}(C_{i,j+1}|\mathcal{H}_{i+j})=\text{Var}(C_{i,j+1}|C_{i,j})=\sigma_j^2\cdot C_{i,j} \end{cases}$$

Under those assumptions

$$\mathbb{E}(C_{i,j+k}|\mathcal{H}_{i+j})=\mathbb{E}(C_{i,j+k}|C_{i,j})=\lambda_j\cdot\lambda_{j+1}\cdots\lambda_{j+k-1}C_{i,j}$$

Mack (1993) assumed further that accident year are independent : $(C_{i,j})_{j=1,\ldots,n}$ and $(C_{i',j})_{j=1,\ldots,n}$ are independent for all $i\neq i'$.

It is possible to write $C_{i,j+1}=\lambda_j C_{i,j}+\sigma_j\sqrt{C_{i,j}}\varepsilon_{i,j}$ where residuals $(\varepsilon_{i,j})$ are i.i.d., centred, with unit variance.

$\Longrightarrow$ it is possible to used weigthed least squares,

$$\min\left\{\sum_{i=1}^{n-j}\frac{1}{C_{i,j}}\left(C_{i,j+1}-\lambda_j C_{i,j}\right)^2\right\}$$

$$\widehat{\lambda}_j=\frac{\sum_{i=1}^{n-j}C_{i,j+1}}{\sum_{i=1}^{n-j}C_{i,j}},\forall j=1,\cdots,n-1.$$

is an unbiased estimator of λ_j.

Further, a natural estimator for the variance parameter is

$$\widehat{\sigma}_j^2=\frac{1}{n-j-1}\sum_{i=1}^{n-j-1}\left(\frac{C_{i,j+1}-\widehat{\lambda}_j C_{i,j}}{\sqrt{C_{i,j}}}\right)^2$$

which can be written

$$\widehat{\sigma}_j^2=\frac{1}{n-j-1}\sum_{i=1}^{n-j-1}\left(\frac{C_{i,j+1}}{C_{i,j}}-\widehat{\lambda}_j\right)^2\cdot C_{i,j}$$

3.3 Uncertainty on $\widehat{R}_i$ and $\widehat{R}$

Recall that

$$\text{mse}(\widehat{R}_i) = \text{mse}(\widehat{C}_{i,n} - C_{i,n-i}) = \text{mse}(\widehat{C}_{i,n}) = \mathbb{E}\left(\left[\widehat{C}_{i,n} - C_{i,n}\right]^2 |\mathcal{H}_n\right)$$

Proposition2

The mean squared error of reserve estimate $\widehat{R}_i$ $\text{mse}(\widehat{R}_i)$, for accident year i can be estimated by

$$\widehat{\text{mse}}(\widehat{R}_i) = \widehat{C}_{i,n}^2 \sum_{k=n-i}^{n-1} \frac{\widehat{\sigma}_k^2}{\widehat{\lambda}_k^2}\left(\frac{1}{\widehat{C}_{i,k}} + \frac{1}{\sum_{j=1}^{n-k} C_{j,k}}\right).$$

And the reserve *all years* $\widehat{R} = \widehat{R}_1 + \cdots + \widehat{R}_n$ satisfies

$$\text{mse}(\widehat{R}) = \mathbb{E}\left(\left[\sum_{i=2}^{n} \widehat{R}_i - \sum_{i=2}^{n} R_i\right]^2 |\mathcal{H}_n\right)$$

Proposition3

The mean squared error of all year reserves $\text{mse}(\widehat{R})$, is given by

$$\widehat{\text{mse}}(\widehat{R}) = \sum_{i=2}^{n} \widehat{\text{mse}}(\widehat{R}_i) + 2 \sum_{2\leq i<j\leq n} \widehat{C}_{i,n}\widehat{C}_{j,n} \sum_{k=n-i}^{n-1} \frac{\widehat{\sigma}_k^2/\widehat{\lambda}_k^2}{\sum_{l=1}^{n-k} C_{l,k}}.$$

Example1

On our triangle $\widehat{\text{mse}}(\widehat{R}) = 79.30$, while $\widehat{\text{mse}}(\widehat{R}_n) = 68.45$, $\widehat{\text{mse}}(\widehat{R}_{n-1}) = 31.3$ or $\widehat{\text{mse}}(\widehat{R}_{n-2}) = 5.05$.

3.4 One year uncertainty, as in Merz & Wüthrich (2007)

	0	1	2	3	4
1	3209	4372	4411	4428	4435
2	3367	4659	4696	4720	4727.4
3	3871	5345	5398	5422.3	5430.9
4	4239	5917	5970.0	5996.9	6006.4
5	4929	6810.8	6871.9	6902.9	6939.0

Table 3 – Triangle of cumulated payments, seen as at year $n = 5$, $\boldsymbol{C} = (C_{i,j})_{i+j\leq n-1, i\leq n-1}$ avec les projection future $\widehat{\boldsymbol{C}} = (\widehat{C}_{i,j})_{i+j>n-1}$.

	0	1	2	3	4	5
1	3209	4372	4411	4428	4435	4456
2	3367	4659	4696	4720	4730	4752.4
3	3871	5345	5398	5420	5430.1	5455.8
4	4239	5917	6020	6046.15	6057.4	6086.1
5	4929	6794	6871.7	6901.5	6914.3	6947.1

Table 4 – Triangle of cumulated payments, seen as at year $n = 6$, on prior years, $\boldsymbol{C} = (C_{i,j})_{i+j\leq n, i\leq n-1}$ avec les projection future $\widehat{\boldsymbol{C}} = (\widehat{C}_{i,j})_{i+j>n}$.

We have to consider different transition factors

$$\widehat{\lambda}_j^n = \frac{\sum_{i=1}^{n-i-1} C_{i,j+1}}{\sum_{i=1}^{n-i-1} C_{i,j}} \text{ and } \widehat{\lambda}_j^{n+1} = \frac{\sum_{i=1}^{n-i} C_{i,j+1}}{\sum_{i=1}^{n-i} C_{i,j}}$$

We know, e.g. that $\mathbb{E}(\widehat{\lambda}_j^n|\mathcal{H}_n) = \lambda_j$ and $\mathbb{E}(\widehat{\lambda}_j^{n+1}|\mathcal{H}_{n+1}) = \lambda_j$.. But here, $n+1$ is the future, so we have to quantify $\mathbb{E}(\widehat{\lambda}_j^{n+1}|\mathcal{H}_{n})$.

Set $S_j^n = C_{1,j} + C_{2,j} + \cdots, C_{j,n-j}$, so that

$$\widehat{\lambda}_j^{n+1} = \frac{\sum_{i=1}^{n-i} C_{i,j+1}}{\sum_{i=1}^{n-i} C_{i,j}} = \frac{\sum_{i=1}^{n-i} C_{i,j+1}}{S_j^{n+1}} = \frac{\sum_{i=1}^{n-1-i} C_{i,j+1}}{S_j^{n+1}} + \frac{C_{n-j,j+1}}{S_j^{n+1}}$$

i.e.

$$\widehat{\lambda}_j^{n+1} = \frac{S_j^n \cdot \widehat{\lambda}_j^n}{S_j^{n+1}} + \frac{C_{n-j,j+1}}{S_j^{n+1}}.$$

Lemma1

Under Mack's assumptions

$$\mathbb{E}(\widehat{\lambda}_j^{n+1}|\mathcal{H}_n) = \frac{S_j^n}{S_j^{n+1}} \cdot \widehat{\lambda}_j^n + \lambda_j \cdot \frac{C_{n-j,n}}{S_j^{n+1}}.$$

Now let us defined the CDR,

Definition1

The *claims development result* $CDR_i(n+1)$,for accident year i, between dates n and $n+1$, is

$$CDR_i(n+1) = \mathbb{E}(R_i^n|\mathcal{H}_n) - \left[Y_{i,n-i+1} + \mathbb{E}(R_i^{n+1}|\mathcal{H}_{n+1})\right],$$

where $Y_{i,n-i+1}$ is the future incremental payment $Y_{i,n-i+1} = C_{i,n-i+1} - C_{i,n-i}$.

Note that $CDR_i(n+1)$ is a $\mathcal{H}_{n+1}$-mesurable martingale, and

$$CDR_i(n+1) = \mathbb{E}(C_{i,n}|\mathcal{H}_n) - \mathbb{E}(C_{i,n}|\mathcal{H}_{n+1}).$$

Lemma2

Under Mack's assumption, an estimator for $\mathbb{E}(\widehat{CDR}_i(n+1)^2|\mathcal{H}_n)$ is

$$\widehat{\text{mse}}(\widehat{CDR}_i(n+1)|\mathcal{H}_n) = \widehat{C}_{i,n}^2 \left(\widehat{\Gamma}_{i,n} + \widehat{\Delta}_{i,n}\right)$$

where

$$\widehat{\Delta}_{i,n} = \frac{\widehat{\sigma}_{n-i+1}^2}{\widehat{\lambda}_{n-i+1}^2 S_{n-i+1}^{n+1}} + \sum_{j=n-i+2}^{n-1} \left(\frac{C_{n-j+1,j}}{S_j^{n+1}}\right)^2 \frac{\widehat{\sigma}_j^2}{\widehat{\lambda}_j^2 S_j^n}$$

and

$$\widehat{\Gamma}_{i,n} = \left(1 + \frac{\widehat{\sigma}_{n-i+1}^2}{\widehat{\lambda}_{n-i+1}^2 C_{i,n-i+1}}\right) \prod_{j=n-i+2}^{n-1} \left(1 + \frac{\widehat{\sigma}_j^2}{\widehat{\lambda}_j^2 [S_j^{n+1}]^2} C_{n-j+1,j}\right) - 1$$

Merz & Wüthrich (2008) observed that

$$\widehat{\Gamma}_{i,n} \approx \frac{\widehat{\sigma}_{n-i+1}^2}{\widehat{\lambda}_{n-i+1}^2 C_{i,n-i+1}} + \sum_{j=n-i+2}^{n-1} \left(\frac{C_{n-j+1,j}}{S_j^{n+1}}\right)^2 \frac{\widehat{\sigma}_j^2}{\widehat{\lambda}_j^2 C_{n-j+1,j}}$$

since $\prod(1+u_i) \approx 1+\sum u_i$, if u_i is small, i.e.

$$\frac{\widehat{\sigma}_j^2}{\widehat{\lambda}_j^2} << C_{n-j+1,j}.$$

We have finally

Proposition4

Under Mack's assumptions

$$\begin{aligned}\widehat{\text{mse}}_n(\widehat{CDR}_i(n+1)) \quad &\approx \quad [\widehat{C}_{i,n}^n]^2\Bigg[\frac{[\widehat{\sigma}_{n-i+1}^n]^2}{[\widehat{\lambda}_{n-i+1}^n]^2}\left(\frac{1}{\widehat{C}_{i,n-i+1}}+\frac{1}{\widehat{S}_{n-i+1}^n}\right)\\ &+\sum_{j=n-i+2}^{n-1}\frac{[\widehat{\sigma}_j^n]^2}{[\widehat{\lambda}_j^n]^2}\left(\frac{1}{\widehat{S}_j^n}\left(\frac{\widehat{C}_{n-j+1,j}}{\widehat{S}_j^{n+1}}\right)^2\right)\Bigg].\end{aligned}$$

while Mack proposed

$$\widehat{\text{mse}}_n(\widehat{R}_i) = [\widehat{C}_{i,n}^n]^2\Bigg[\frac{[\widehat{\sigma}_{n-i+1}^n]^2}{[\widehat{\lambda}_{n-i+1}^n]^2}\left(\frac{1}{\widehat{C}_{i,n-i+1}}+\frac{1}{\widehat{S}_{n-i+1}^n}\right)+\sum_{j=n-i+2}^{n-1}\frac{[\widehat{\sigma}_j^n]^2}{[\widehat{\lambda}_j^n]^2}\left(\frac{1}{\widehat{C}_{i,j}}+\frac{1}{\widehat{S}_j^n}\right)\Bigg].$$

i.e. only the first term of the model error is considered here, and only the first diagonal is considered for the process error ($i+j=n+1$) (the other terms are neglected compared with $\widehat{C}_{n-j+1,j}/\widehat{S}_j^{n+1}$).

Finally, all year, we have

$$\widehat{\text{mse}}_n(CDR(n+1)) \approx \sum_{i=1}^{n} \widehat{\text{mse}}_n(CDR_i(n+1))$$

$$+2\sum_{i<l} \widehat{C}_{i,n}^{n}\widehat{C}_{l,n}^{n} \left(\frac{[\widehat{\sigma}_{n-i}^{n}]^2/[\widehat{\lambda}_{n-i}^{n}]^2}{\sum_{k=0}^{i-1} C_{k,n-i}} + \sum_{j=n-i+1}^{n-1} \frac{C_{n-j,j}}{\sum_{k=0}^{n-j} C_{k,j}} \frac{[\widehat{\sigma}_{j}^{n}]^2/[\widehat{\lambda}_{j}^{n}]^2}{\sum_{k=0}^{n-j-1} C_{k,j}} \right).$$

again if $C_{n-j+1,j} \leq S_j^{n+1}$.

Example2

On our triangle $\widehat{\text{mse}}_n(\text{CDR}(n+1)) = 72.57$, while $\widehat{\text{mse}}_n(\text{CDR}_n(n+1)) = 60.83$, $\widehat{\text{mse}}_n(\text{CDR}_{n-1}(n+1)) = 30.92$ or $\widehat{\text{mse}}_n(\text{CDR}_{n-2}(n+1)) = 4.48$. La formule approchée donne des résultats semblables.

	Process error (intrisic volatilitity)			Estimation error (model error)			Prediction error (total)		
	Whole run-off	One year horizon	Variation (%)	Whole run-off	One year horizon	Variation (%)	Whole run-off	One year horizon	Variation (%)
participant n°1 (WCp1)	4.60%	4.34%	*-6%*	2.10%	1.81%	*-14%*	5.10%	4.70%	*-8%*
participant n°1 (WCp2)	1.48%	1.23%	*-17%*	1.45%	1.30%	*-10%*	2.07%	1.79%	*-14%*
participant n°2 (GL1)	4.40%	1.90%	*-57%*	6.60%	3.00%	*-55%*	7.90%	3.60%	*-54%*
participant n°2 (GL2)	4.80%	2.50%	*-48%*	6.80%	3.20%	*-53%*	8.30%	4.10%	*-51%*
participant n°3 (GL)	4.65%	2.54%	*-45%*	6.15%	2.80%	*-54%*	7.70%	3.78%	*-51%*
participant n°5 (GL)	5.23%	2.03%	*-61%*	9.19%	4.96%	*-46%*	10.58%	5.36%	*-49%*
participant n°5 (WCp)	6.91%	5.56%	*-20%*	5.51%	3.42%	*-38%*	8.84%	6.53%	*-26%*
participant n°9 (GL)	6.80%	4.80%	*-29%*	11.60%	6.60%	*-43%*	13.50%	8.20%	*-39%*
participant n°10 (GL)	5.05%	3.77%	*-25%*	3.62%	3.17%	*-12%*	6.21%	4.93%	*-21%*

4 Poisson regression

We have seen that a factor model $Y_{i,j} = a_i \times b_j$ should be interesting (and can be related to Chain Ladder)

4.1 Hachemeister & Stanard

Hachemeister & Stanard (1975), Kremer (1985) and Mack (1991) considered a log-Poisson regression on incremental payments

$$\mathbb{E}(Y_{i,j}) = \mu_{i,j} = \exp[r_i + c_j] = a_i \cdot b_j$$

Then our *best estimate* is

$$\widehat{Y}_{i,j} = \widehat{\mu}_{i,j} = \exp[\widehat{r}_i + \widehat{c}_j] = \widehat{a}_i \cdot \widehat{b}_j.$$

Example3

On our triangle, we have

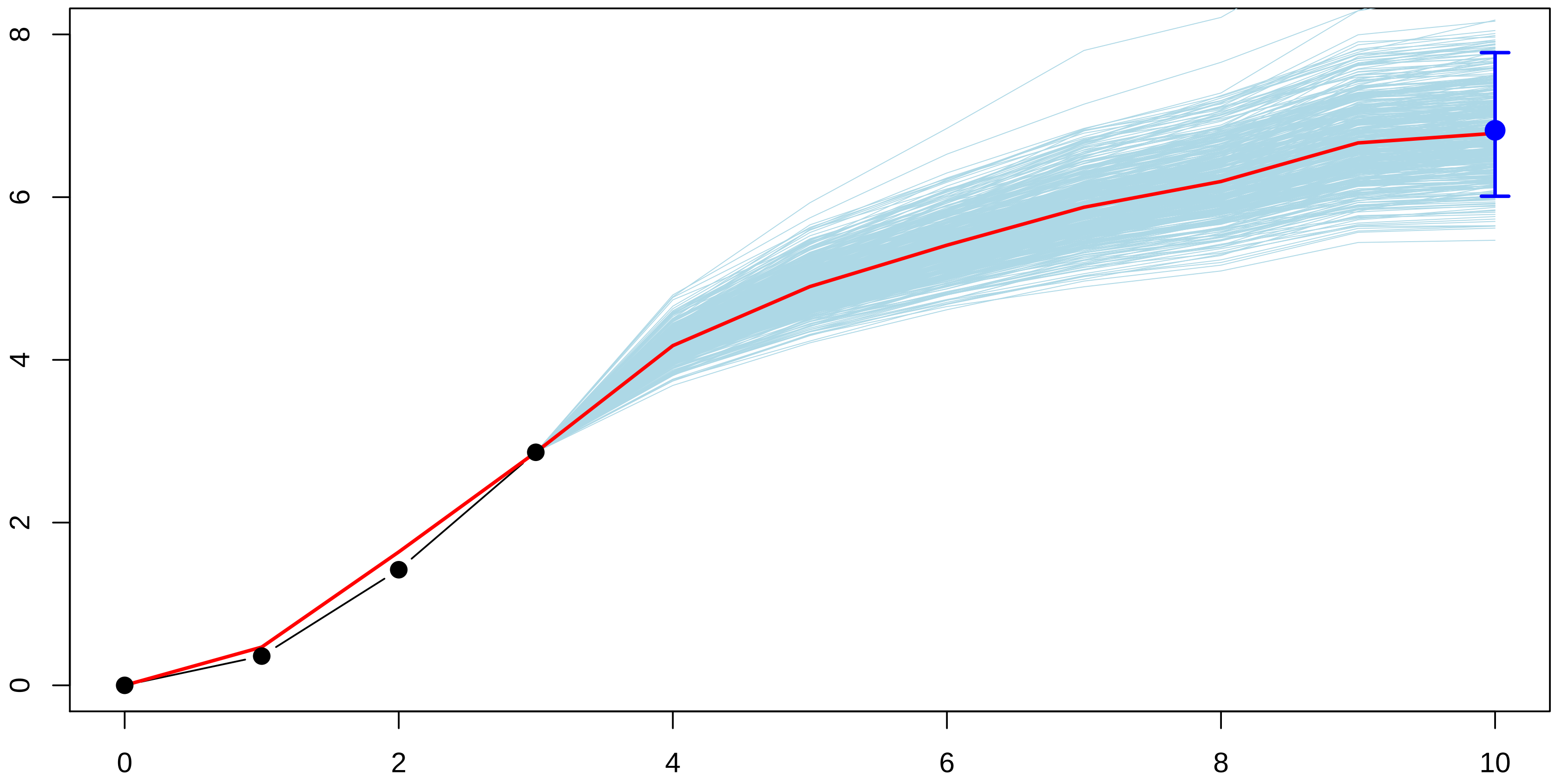
```
Call:
glm(formula = Y ~ lig + col, family = poisson("log"), data = base)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  8.05697    0.01551 519.426  < 2e-16 ***
lig2         0.06440    0.02090   3.081  0.00206 **
lig3         0.20242    0.02025   9.995  < 2e-16 ***
lig4         0.31175    0.01980  15.744  < 2e-16 ***
lig5         0.44407    0.01933  22.971  < 2e-16 ***
lig6         0.50271    0.02079  24.179  < 2e-16 ***
col2        -0.96513    0.01359 -70.994  < 2e-16 ***
col3        -4.14853    0.06613 -62.729  < 2e-16 ***
col4        -5.10499    0.12632 -40.413  < 2e-16 ***
col5        -5.94962    0.24279 -24.505  < 2e-16 ***
col6        -5.01244    0.21877 -22.912  < 2e-16 ***
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 46695.269  on 20  degrees of freedom
Residual deviance:    30.214  on 10  degrees of freedom
  (15 observations deleted due to missingness)
AIC: 209.52

Number of Fisher Scoring iterations: 4
```

	0	1	2	3	4	5
1	3209	4372	4411	4428	4435	4456
2	3367	4659	4696	4720	4730	4752.4
3	3871	5345	5398	5420	5430.1	5455.8
4	4239	5917	6020	6046.15	6057.4	6086.1
5	4929	6794	6871.7	6901.5	6914.3	6947.1
6	5217	7204.3	7286.7	7318.3	7331.9	7366.7

Table 5 – Triangle of cumulated payments, based on sums of $\widehat{\boldsymbol{Y}} = (\widehat{Y}_{i,j})_{0\leq i,j\leq n}$'s obtained from the Poisson regression.

4.2 Uncertainty in our regression model



4.2.1 Les formules économétriques fermées

Using standard GLM notions,

$$\widehat{Y}_{i,j} = \widehat{\mu}_{i,j} = \exp[\widehat{\eta}_{i,j}].$$

Using delta method we can write

$$\text{Var}(\widehat{Y}_{i,j}) \approx \left|\frac{\partial \mu_{i,j}}{\partial \eta_{i,j}}\right|^2 \cdot \text{Var}(\widehat{\eta}_{i,j}),$$

i.e. (with a log link)

$$\frac{\partial \mu_{i,j}}{\partial \eta_{i,j}} = \mu_{i,j}$$

In an overdispersed Poisson regression (as in Renshaw (1998)),

$$\mathbb{E}\left([Y_{i,j} - \widehat{Y}_{i,j}]^2\right) \approx \widehat{\phi} \cdot \widehat{\mu}_{i,j} + \widehat{\mu}_{i,j}^2 \cdot \widehat{\text{Var}}(\widehat{\eta}_{i,j})$$

for the lower part of the triangle. Further, since

$$\text{Cov}(\widehat{Y}_{i,j}, \widehat{Y}_{k,l}) \approx \widehat{\mu}_{i,j} \cdot \widehat{\mu}_{k,l} \cdot \widehat{\text{Cov}}(\widehat{\eta}_{i,j}, \widehat{\eta}_{k,l}).$$

we get

$$\mathbb{E}\left([R-\widehat{R}]^2\right) \approx \left(\sum_{i+j>n} \widehat{\phi}\cdot\widehat{\mu}_{i,j}\right) + \widehat{\boldsymbol{\mu}}' \cdot \widehat{\text{Var}}(\widehat{\boldsymbol{\eta}}) \cdot \widehat{\boldsymbol{\mu}}$$

Example4

On our triangle, the mean square error is 131.77 (to be compared with 79.30).

4.2.2 Using bootstrap techniques

From triangle of incremental payments, $(Y_{i,j})$ assume that

$$Y_{i,j} \sim \mathcal{P}(\widehat{Y}_{i,j}) \text{ where } \widehat{Y}_{i,j} = \exp(\widehat{L}_i + \widehat{C}_j)$$

1. Estimate parameters $\widehat{L}_i$ and $\widehat{C}_j$, define Pearson's (pseudo) residuals

$$\widehat{\varepsilon}_{i,j} = \frac{Y_{i,j} - \widehat{Y}_{i,j}}{\sqrt{\widehat{Y}_{i,j}}}$$

2. Generate pseudo triangles on the past, $\{i + j \leq t\}$

$$Y^\star_{i,j} = \widehat{Y}_{i,j} + \widehat{\varepsilon}^\star_{i,j}\sqrt{\widehat{Y}_{i,j}}$$

3. (re)Estimate parameters $\widehat{L}^\star_i$ and $\widehat{C}^\star_j$, and derive expected payments for the future, $\widehat{Y}^\star_{i,j}$.

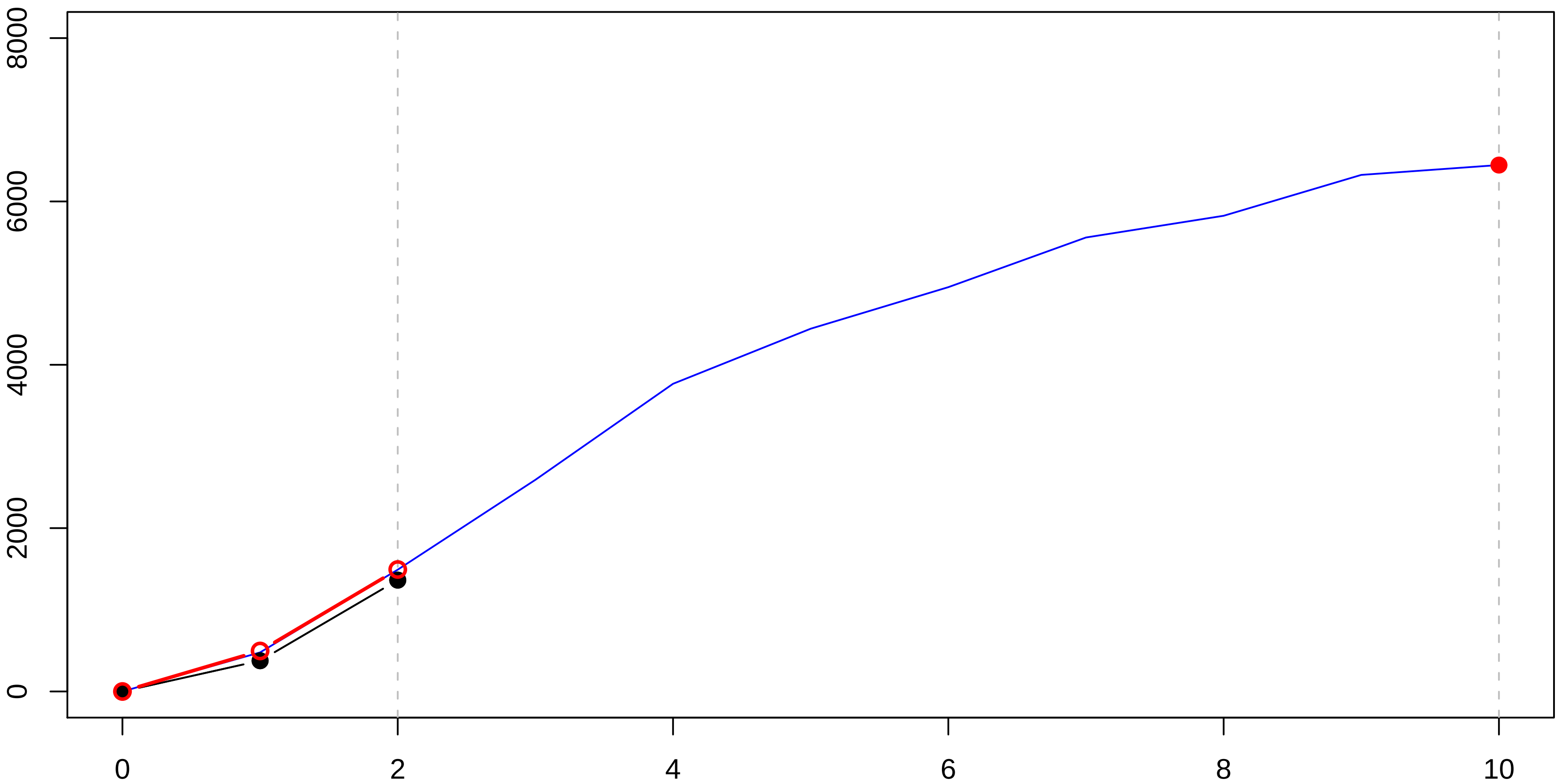
$$\widehat{R} = \sum_{i+j>t} \widehat{Y}^\star_{i,j}$$

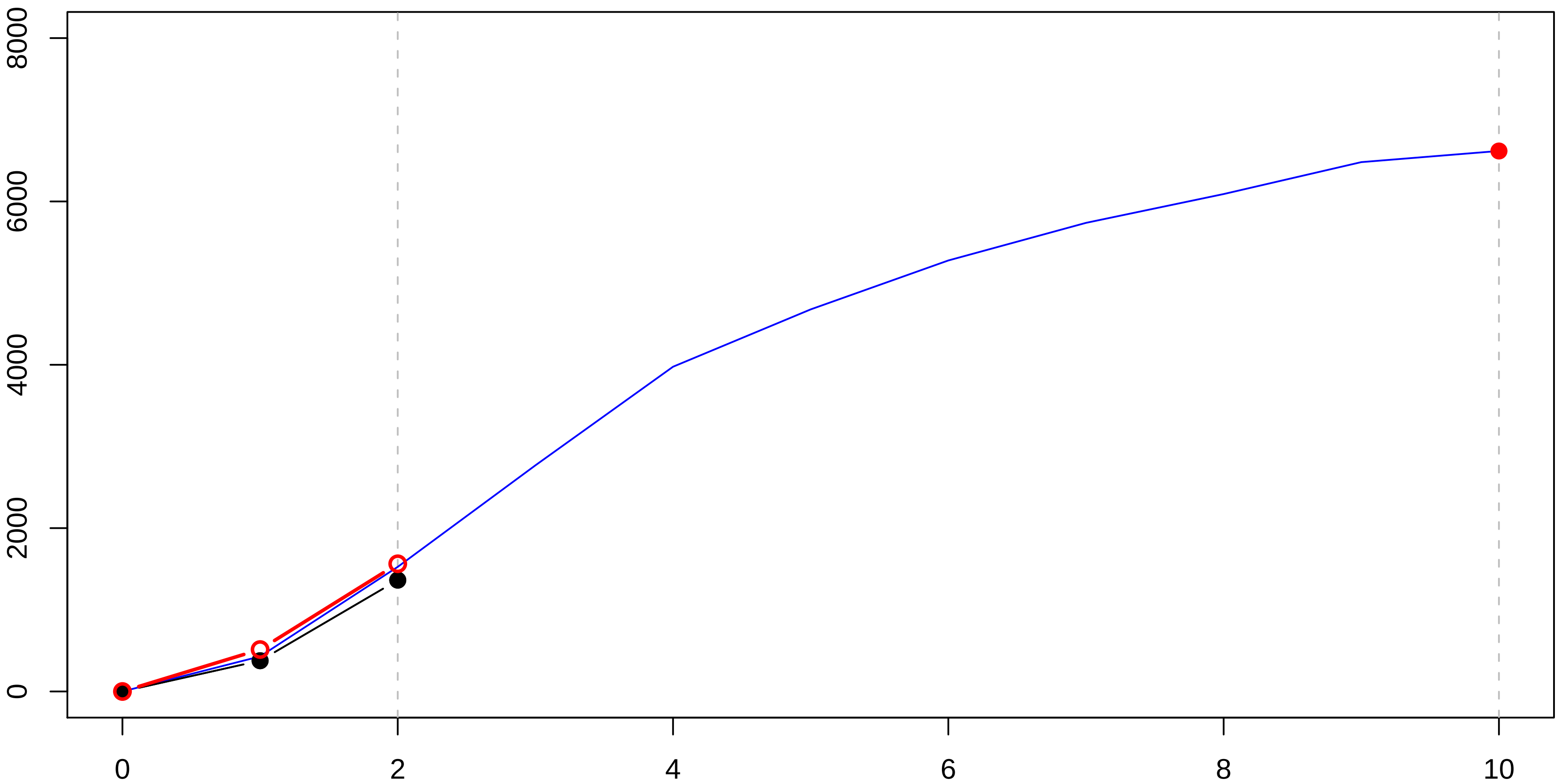
is the best estimate.

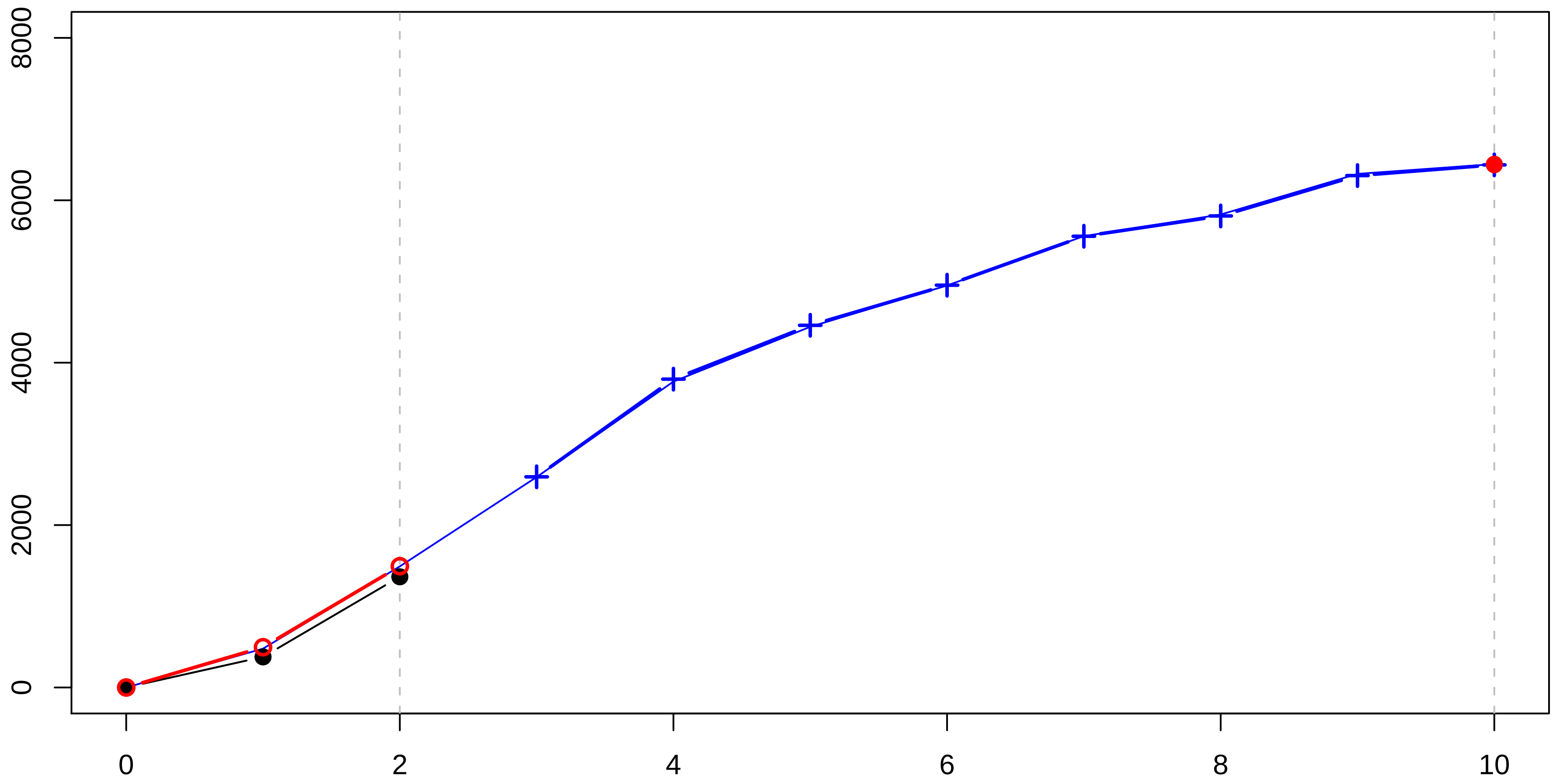
4. Generate a scenario for future payments, $Y^\star_{i,j}$ e.g. from a Poisson distribution $\mathcal{P}(\widehat{Y}^\star_{i,j})$

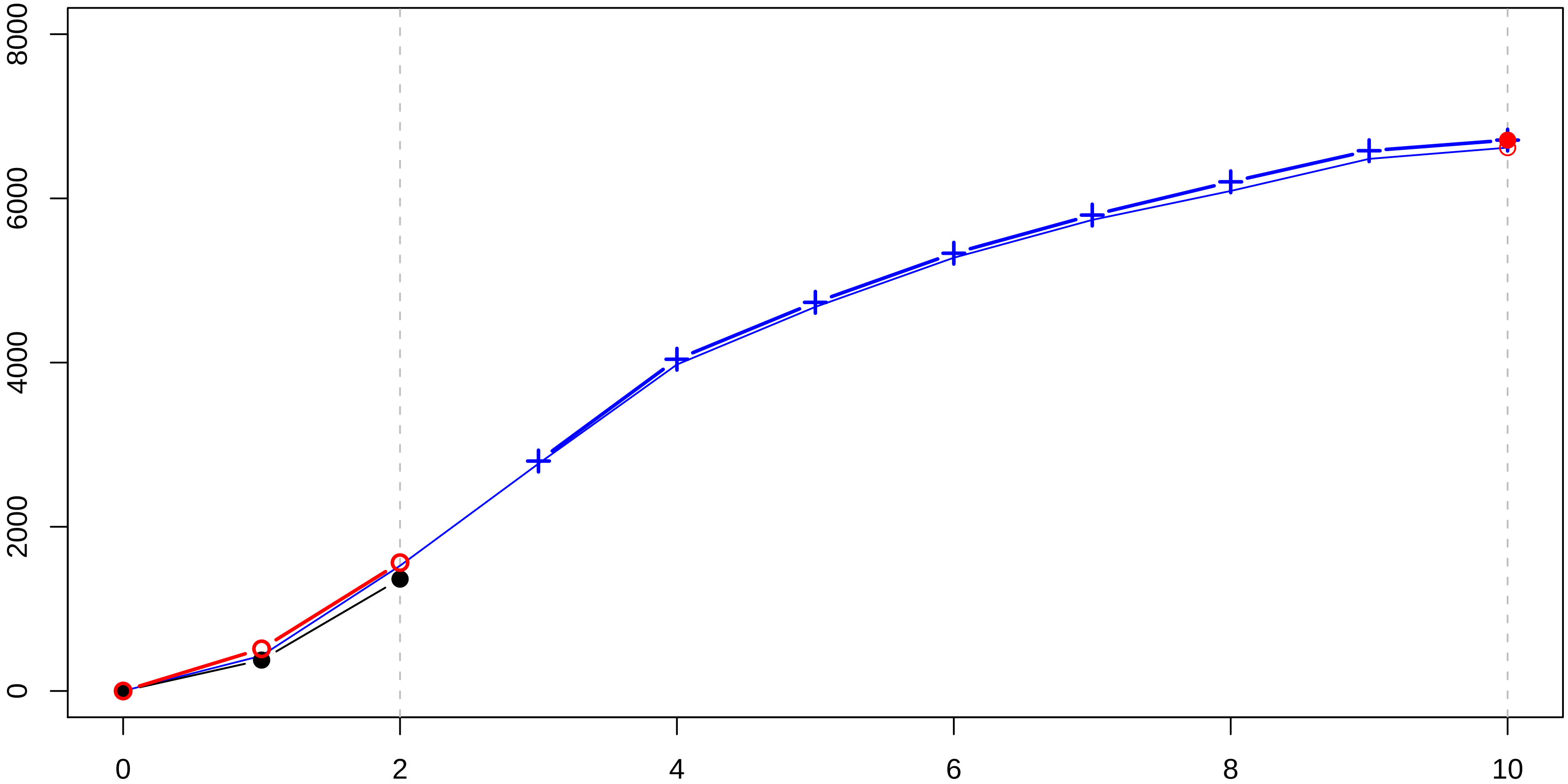
$$R = \sum_{i+j>t} Y^\star_{i,j}$$

One needs to repeat steps 2-4 several times to derive a distribution for R.

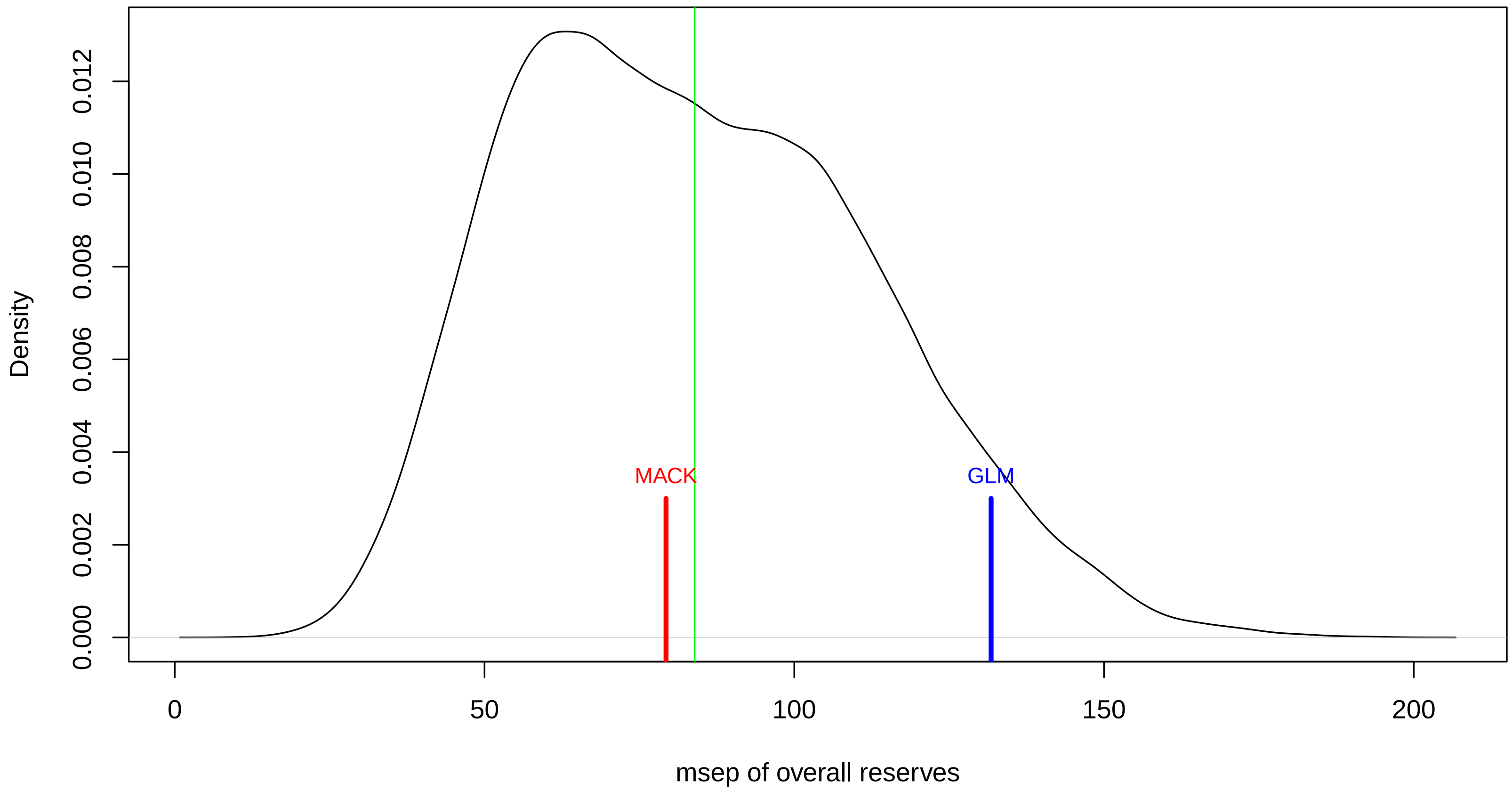

0
2000
4000
6000
8000
0
2
4
6
8
10


0
2000
4000
6000
8000
0
2
4
6
8
10


0
2000
4000
6000
8000
0
2
4
6
8
10


0
2000
4000
6000
8000
0
2
4
6
8
10

If we repeat it 50,000 times, we obtain the following distribution for the mse.



4.2.3 Bootstrap and one year uncertainty

2. Generate pseudo triangles on the past *and next year* $\{i+j \leq t+1\}$

$$Y^{\star}_{i,j} = \widehat{Y}_{i,j} + \widehat{\varepsilon}^{\star}_{i,j}\sqrt{\widehat{Y}_{i,j}}$$

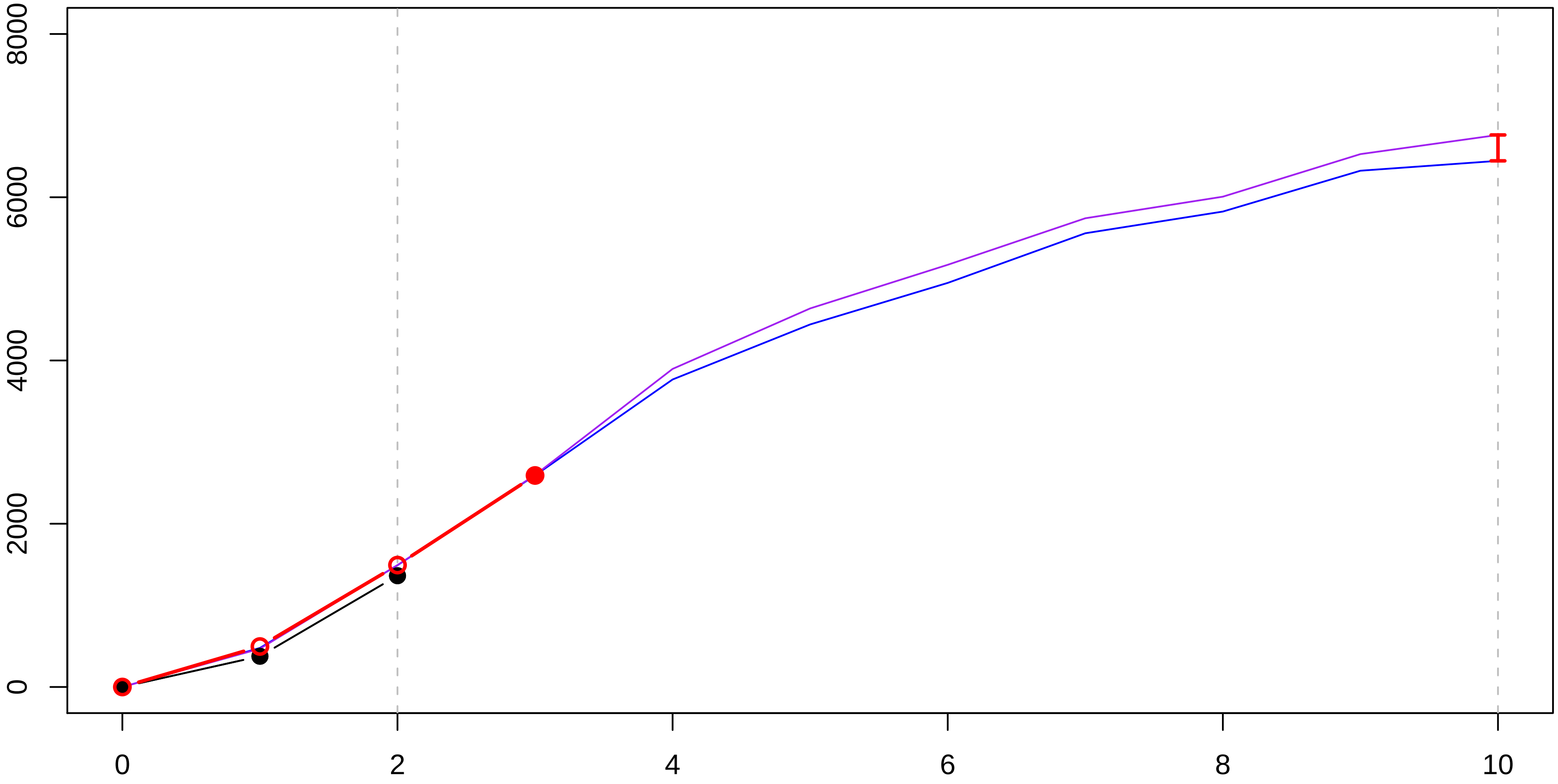
3. Estimate parameters $\widehat{L}^{\star}_i$ and $\widehat{C}^{\star}_j$, on the past, $\{i+j \leq t\}$, and derive expected payments for the future, $\widehat{Y}^{\star}_{i,j}$.

$$\widehat{R}_t = \sum_{i+j>t} \widehat{Y}_{i,j}$$

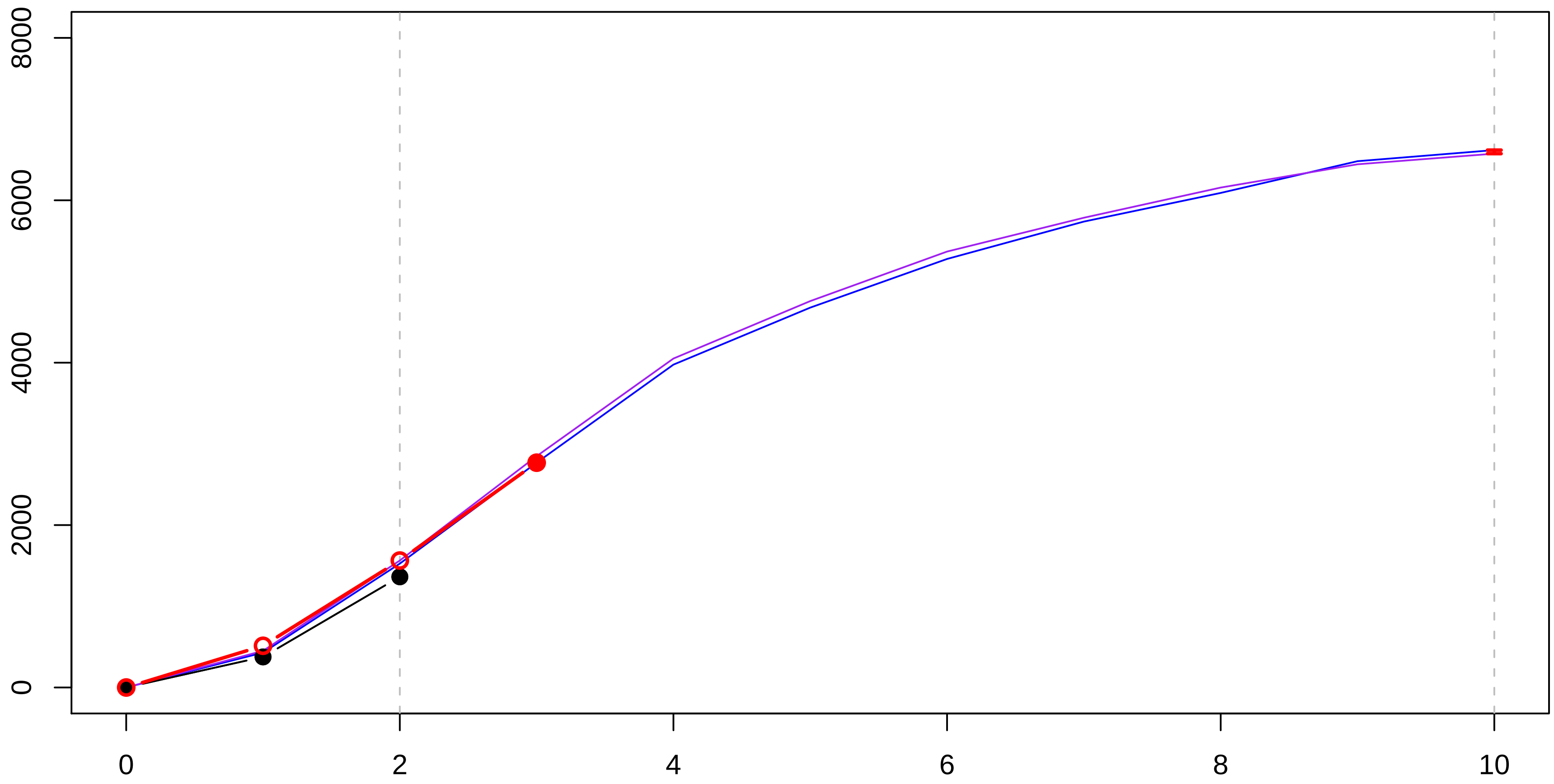
4. Estimate parameters $\widehat{L}^{\star}_i$ and $\widehat{C}^{\star}_j$, on the past and next year, $\{i+j \leq t+1\}$, and derive expected payments for the future, $\widehat{Y}^{\star}_{i,j}$.

$$\widehat{R}_{t+1} = \sum_{i+j>t} \widehat{Y}_{i,j}$$

5. Calculate CDR as CDR$=\widehat{R}_{t+1} - \widehat{R}_t$.



0
2000
4000
6000
8000
0
2
4
6
8
10



0
2000
4000
6000
8000
0
2
4
6
8
10

ultimate $(R - \mathbb{E}(R))$ versus one year uncertainty,