Solvency II' *newspeak*
'one year uncertainty for IBNR'

Arthur Charpentier[1] & Laurent Devineau[2]

[1] Université Rennes 1 - CREM & École Polytechnique, [2] Milliman

arthur.charpentier@univ-rennes1.fr

http ://blogperso.univ-rennes1.fr/arthur.charpentier/index.php/

AXA GRM, March 2010

Agenda of the talk

- **Solvency II : CP 71 and the *one year horizon***
- **The Chain-Ladder estimator**
- **Understanding the actuarial *wewspeak* in Solvency II**
  - From MSE to MSEP (MSE of prediction)
  - From MSEP to MSEPC (conditional MSEP)
  - CDR, claims development result
- **From Mack (1993) to Merz & Wüthrich (2009)**
- **Updating Poisson-ODP bootstrap technique**

	one year	ultimate
mse model	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.

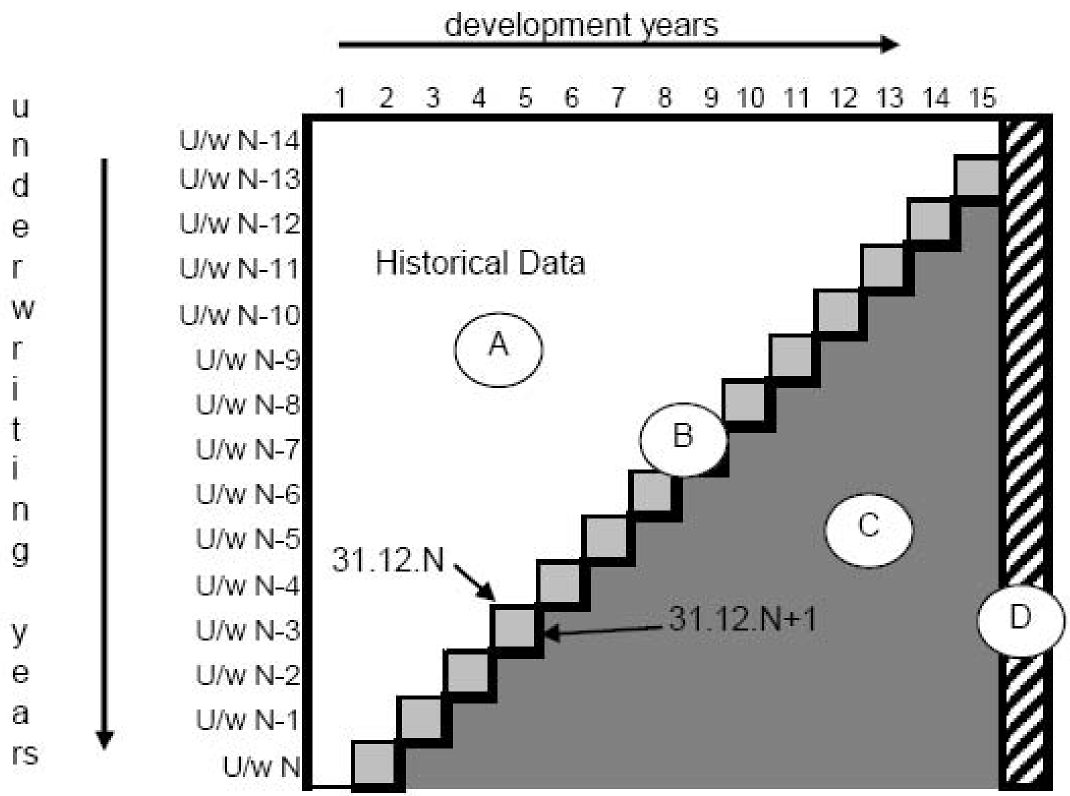
4.1.2 The reserve risk captures uncertainty over a one year period

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'

	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%*

'one year horizon for the reserve risk'

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 $Var(C_{i,j}|C_{i,j-1})=s_j^2C_{i,j-1}$.

Notations for triangle type data

- $X_{i,j}$ 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					

- $\mathcal{F}_t$ denotes information available at time t,

$$\mathcal{F}_t = \{(C_{i,j}), 0 \leq i+j \leq t\} = \{(X_{i,j}), 0 \leq i+j \leq t\}$$

Notations for triangle type data

- $X_{i,j}$ 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, $C_{i,j} = X_{i,0} + X_{i,1} + \cdots + X_{i,j}$,

$$C_{i,j} = X_{i,0} + X_{i,1} + \cdots + X_{i,j}$$

	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					

- ${\mathcal{F}_t}^k$ denotes partial information available at time t, based on the first k years, only

$${\mathcal{F}_t}^k = \{(C_{i,j}), 0 \leq i+j \leq t, i \leq k\} = \{(X_{i,j}), 0 \leq i+j \leq t, i \leq k\}$$

Chain Ladder estimation

	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					

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

with the following link ratios

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

One the triangle has been completed, we obtain the amount of reserves, with respectively 22, 36, 66, 153 and 2150 per accident year, i.e. the total is 2427.

Mack's stochastic model

Mack (1993) proposed the following stochastic model for claims reserving.

Three assumptions were made on incremental payments

$$\mathbb{E}(C_{i,j+1}|\mathcal{F}_{i+j}) = \mathbb{E}(C_{i,j+1}|C_{i,j}) = \lambda_j \cdot C_{i,j}$$

...

$$\text{Var}(C_{i,j+1}|\mathcal{F}_{i+j}) = \text{Var}(C_{i,j+1}|C_{i,j}) = \sigma_j^2 \cdot C_{i,j}$$

and independence between occurrence years (i.e. rows in the triangle).

Under those assumptions

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

Chain Ladder's standard estimator was

$$\widehat{\lambda}_j = \frac{\sum_{i=0}^{n-j-1} C_{i,j+1}}{\sum_{i=0}^{n-j-1} C_{i,j}}$$

which is an unbiased estimator of λ_j, given $\mathcal{F}_n$.

$$\mathbb{E}(\widehat{\lambda}_j|\mathcal{F}_j) = \lambda_j$$

Further $\widehat{\lambda}_j$ and $\widehat{\lambda}_{j+h}$ non-correlated, given $\mathcal{F}_j$.

Thus, an unbiased estimator for $\mathbb{E}(C_{i,j+k}|\mathcal{F}_{i+j})$ is

$$\widehat{C}_{i,j+k} = \widehat{\lambda}_j \cdot \widehat{\lambda}_{j+1} \cdots \widehat{\lambda}_{j+k-1} C_{i,j}$$

$$\mathbb{E}(\widehat{C}_{i,j+k}|\mathcal{F}_j) = C_{j+k}$$

Moreover, $\widehat{\lambda}_j$ is the estimator with minimum variance in the class of linear combination of link ratios $\lambda_{i,j} = C_{i,j+1}/C_{i,j}$.

Finally,

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

is an unbiased estimator of σ_j^2, given $\mathcal{F}_j$.

$$\mathbb{E}(\widehat{\sigma}_j^2|\mathcal{F}_j) = \sigma_j^2$$

In practice, on border (i.e. in $n-1$), extrapolation satisfies

$$\frac{\widehat{\sigma}_{n-3}^2}{\widehat{\sigma}_{n-2}^2} = \frac{\widehat{\sigma}_{n-2}^2}{\widehat{\sigma}_{n-1}^2} \text{ , i.e. } \widehat{\sigma}_{n-1}^2 = \min\left\{\frac{\sigma_{n-2}^4}{\sigma_{n-3}^2}, \min\left\{\sigma_{n-3}^2, \sigma_{n-2}^2\right\}\right\}$$

How to quantify uncertainty in triangles

In statistics, the mean squared error is a standard measure to quantify the uncertainty of an estimator, i.e.

$$\text{mse}(\widehat{\theta}) = \mathbb{E}\left(\left[\widehat{\theta} - \theta\right]^2\right)$$

$$\theta$$

In order to formalize the prediction process in claims reserving consider the following simpler case.

Let $\{x_1, \cdots, x_n\}$ denote an i.i.d. $\mathcal{B}(p)$ sample. We want to predict $S_h = X_{n+1} + \cdots + X_{n+h}$.

Let ${}_n\widehat{S}_h = \psi(X_{n+1}, \cdots, X_{n+h}) = h \cdot \widehat{p}_n$ denote the *natural* predictor for S_h, at time n.

Since S_h is a random variable (θ was a constant) define

$$\text{mse}({}_n\widehat{S}_h) = \mathbb{E}\left(\left[{}_n\widehat{S}_h - \mathbb{E}(S_h)\right]^2\right)$$

and

$$\text{msep}({}_n\widehat{S}_h) = \mathbb{E}\left(\left[{}_n\widehat{S}_h - S_h\right]^2\right)$$

Note that

$$\text{msep}({}_n\widehat{S}_h) = \mathbb{E}\left(\left[{}_n\widehat{S}_h - \mathbb{E}(S_h)\right]^2\right) + \mathbb{E}\left([\mathbb{E}(S_h) - S_h]^2\right)$$

$$\text{msep}({}_n\widehat{S}_h) = \text{mse}({}_n\widehat{S}_h) + \text{Var}(S_h)$$

where the first term is a process error and the second term a estimation error.

In Solvency II requirements,

$$CDR_{n+1} = [{}_n\widehat{S}_h] - [x_{n+1} + {}_{n+1}\widehat{S}_{h-1}]$$

This defines a martingale since

$$\mathbb{E}(CDR_{n+1}|\mathcal{F}_n) = 0$$

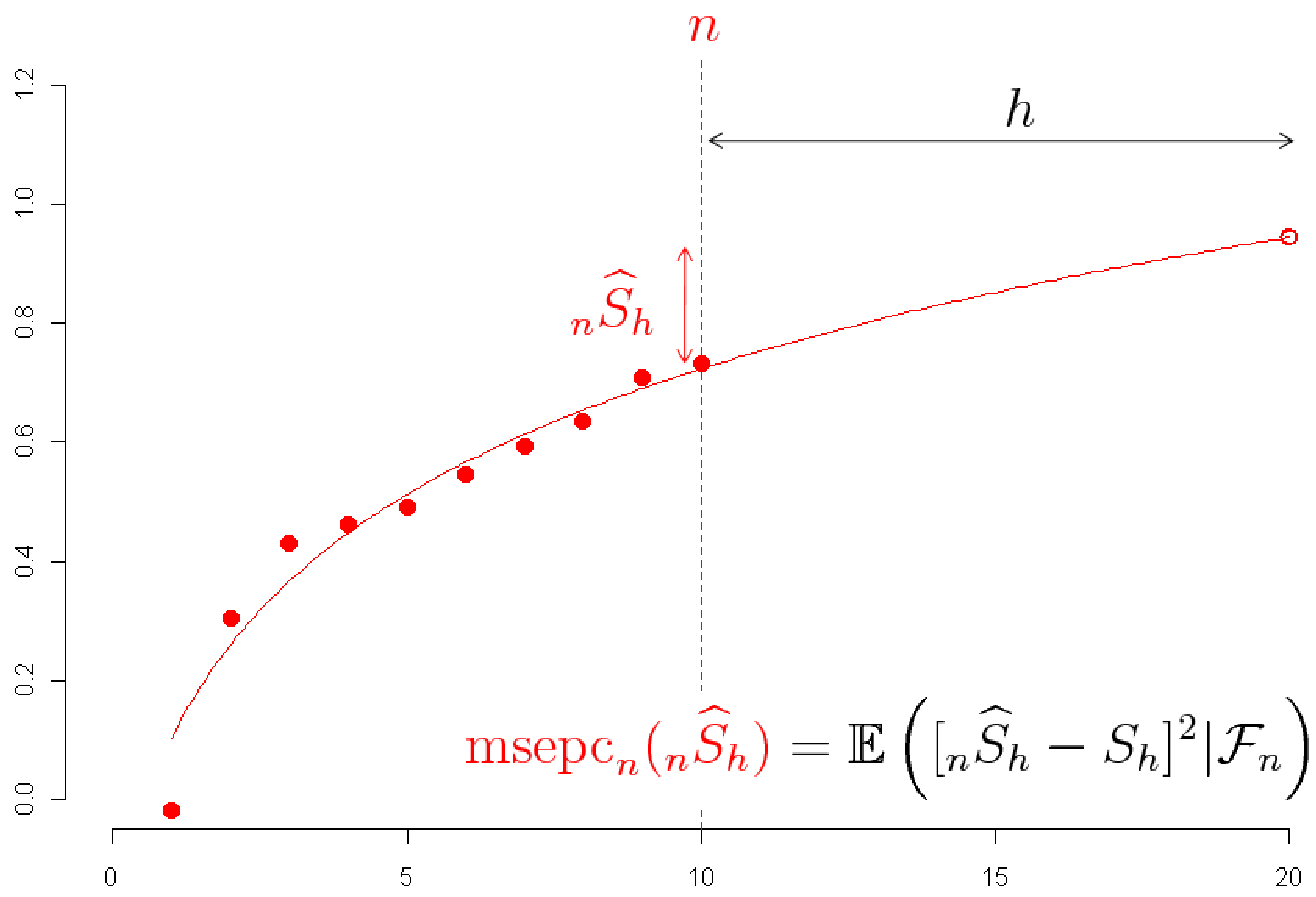
and what is required is to estimate

$$\text{msepc}_n(CDR_{n+1})$$

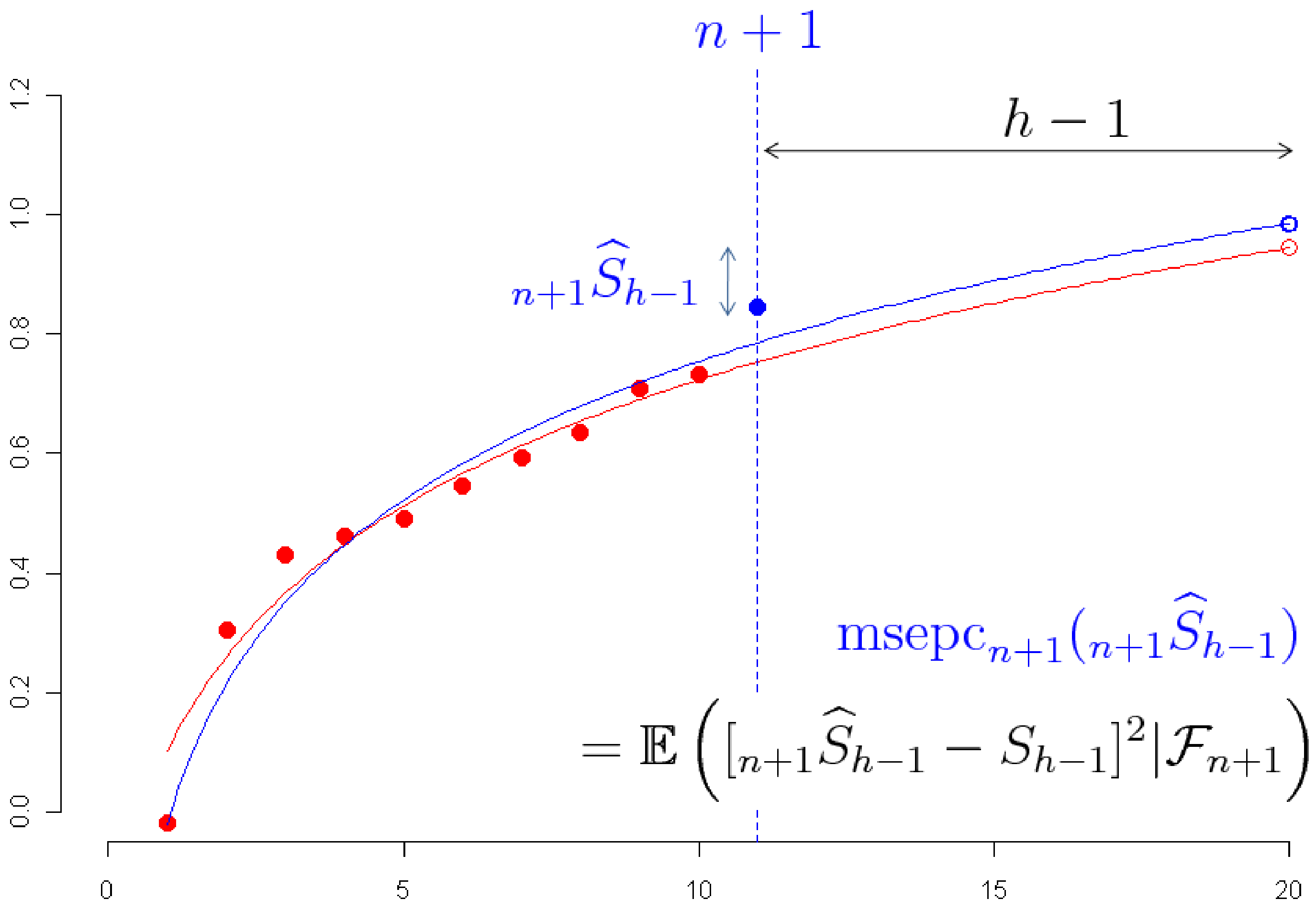
i.e. find $\widehat{\text{msepc}}_n(CDR_{n+1})$.

$$\text{msepc}_n({}_n\widehat{S}_h) = \mathbb{E}\left(\left[{}_n\widehat{S}_h - S_h\right]^2 |\mathcal{F}_n\right)$$

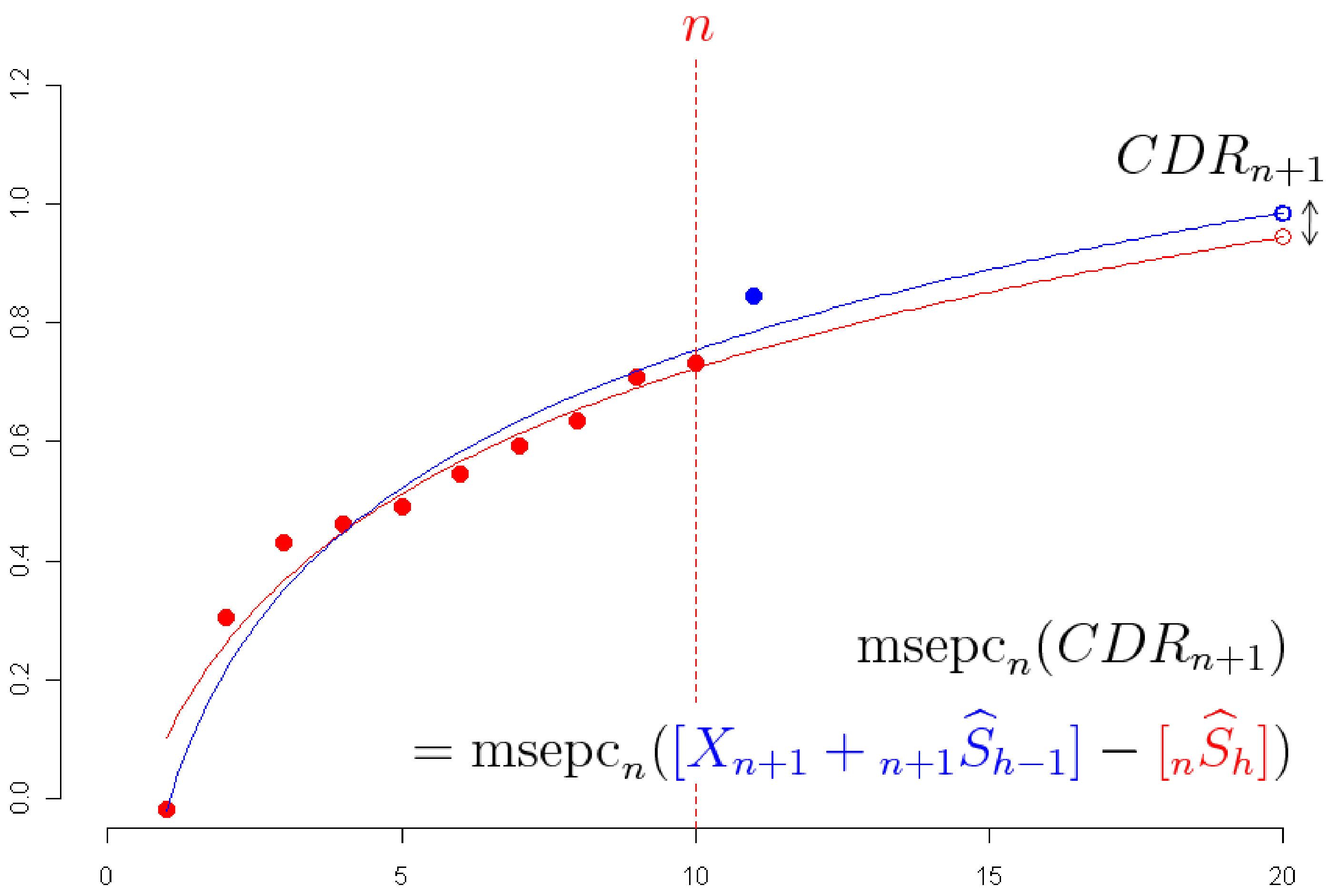
What are we looking for ?



What are we looking for ?



What are we looking for ?



$$\text{msepc}_n(CDR_{n+1})$$

$$= \text{msepc}_n([X_{n+1} + {}_{n+1}\widehat{S}_{h-1}] - [{}_n\widehat{S}_h])$$

A simplified model for claims reserving

Let us continue with our repeated tails/heads game. Let $\widehat{p}_n = [x_1 + \cdots + x_n]/n$, so that

$$\text{Var}(\widehat{p}_n) = \frac{p(1-p)}{n}$$

thus

$$\text{mse}({}_n\widehat{S}_h) = \text{mse}(h \cdot \widehat{p}_n) = h^2 \cdot \text{mse}(\widehat{p}_n) = \frac{h^2}{n} p(1-p)$$

or

$$\text{msep}({}_n\widehat{S}_h) = nhp(1-p) + \frac{h^2}{n} p(1-p) = \frac{nh + h^2}{n} p(1-p)$$

i.e.

$$\text{msep}({}_n\widehat{S}_h) = \frac{h(n+h)}{n} p(1-p)$$

Thus, this quantity can be estimated as

$$\widehat{\text{msep}}({}_n\widehat{S}_h) = \frac{h(n+h)}{n} \widehat{p}_n(1-\widehat{p}_n)$$

while the mse estimator was

$$\widehat{\text{mse}}({}_n\widehat{S}_h) = \frac{h^2}{n}\widehat{p}_n(1-\widehat{p}_n)$$

Looking that the msepc at time n, we have

$$\text{msepc}_n({}_n\widehat{S}_h) = \text{Var}(S|\mathcal{F}_n) + \text{mse}({}_n\widehat{S}_h|\mathcal{F}_n)$$

where

$$\begin{aligned}\text{Var}(S|\mathcal{F}_n) &= \text{Var}(X_{n+1} + \cdots + X_{n+h}|x_1, \cdots, x_n) \\ &= \text{Var}(X_{n+1} + \cdots + X_{n+h}) = hp(1-p)\end{aligned} \tag{2}$$

and

$$\text{mse}({}_n\widehat{S}_h|\mathcal{F}_n) = \left(\mathbb{E}(S_h|\mathcal{F}_n) - {}_n\widehat{S}_h\right)^2$$

which can be written

$$\text{msepc}_n({}_n\widehat{S}_h) = hp(1-p) + h^2\left(p - \widehat{p}_n\right)^2$$

This quantity can be estimated as

$$\widehat{\text{msepc}}_n({}_n\widehat{S}_h) = h\widehat{p}_n(1-\widehat{p}_n) + 0$$

i.e. we keep only the *variance process* term.

Mack (1993) suggested to use partial information to estimate the second term. Define $D = \{X_i, i \leq n\}$ and $B_k = \{X_i, i \leq n, i \leq k \leq n\}$ with $k \leq n$. Define

$$\widehat{\text{msepc}}_n^k({}_n\widehat{S}_h) = h\widehat{p}_n(1-\widehat{p}_n) + h^2\left(\widehat{p}_n - \widehat{p}_k\right)^2$$

In the following, we considered $k = n-1$.

Boostrap estimation of those quantities

The *problem* with mse's estimators is that if $\widehat{\theta}$ is an unbiased estimator of θ, then $\widehat{\text{mse}}(\widehat{\theta})$ is usually a biased estimator of

$$\text{mse}(\widehat{\theta}) = \mathbb{E}\left(\left[\widehat{\theta} - \theta\right]^2\right)$$

(Jensen's inequality). For instance,

$$\mathbb{E}\left(\widehat{\text{msep}}({}_n\widehat{S}_h)\right) = \mathbb{E}\left(\frac{h(n+h)}{n}\widehat{p}_n(1-\widehat{p}_n)\right)$$

i.e.

$$\mathbb{E}\left(\widehat{\text{msep}}({}_n\widehat{S}_h)\right) = \frac{h(n+h)}{n}\left(\mathbb{E}(\widehat{p}_n) - \mathbb{E}(\widehat{p}_n^2)\right)$$

with $\mathbb{E}(\widehat{p}_n) = p$ and

$$\mathbb{E}(\widehat{p}_n^2) = \mathbb{E}\left(\frac{1}{n^2}\sum X_i \sum X_j\right) = \frac{1}{n^2}\mathbb{E}\left(\sum X_i X_j\right)$$

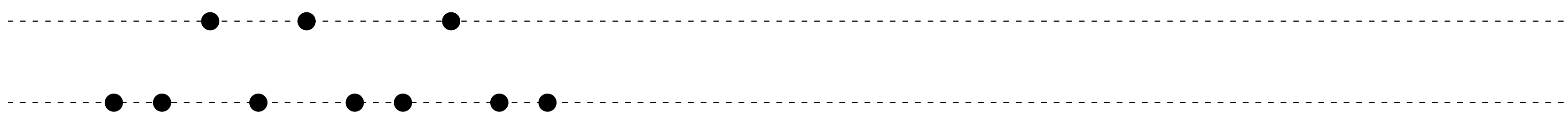
i.e.

$$\mathbb{E}(\widehat{p}_n^2) = \frac{1}{n^2}\left(np + n(n-1)p^2\right) = p^2 + \frac{p(1-p)}{n}$$

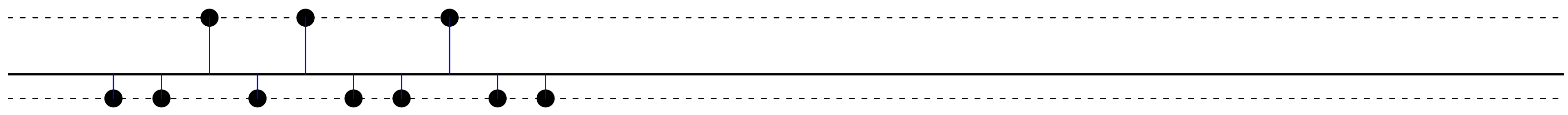
Thus

$$\mathbb{E}\left(\widehat{\text{msep}}({}_n\widehat{S}_h)\right) = \text{msep}({}_n\widehat{S}_h) + \frac{h(n+h)p(1-p)}{n^2}$$

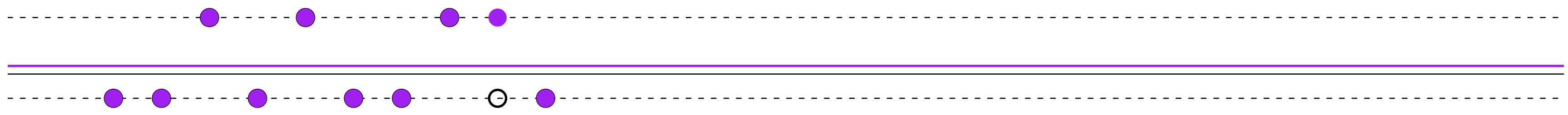
The estimated mse has a biais.



Let $(\widehat{\varepsilon}_1, \cdots, \widehat{\varepsilon}_n) = (x_1 - \widehat{p}_n, \cdots, x_n - \widehat{p}_n)$, i.e.



A standard technique is to boostrap the error term (instead of the sample data)



Here bootstrap techniques can be used to remove the biais of the mse estimator.

The one year horizon uncertainty

In Solvency II, insurance companies are requiered to estimate the msepc, at time n, of the difference between $X_{n+1} + {}_{n+1}\widehat{S}_{(h-1)}$ and ${}_{n}\widehat{S}_{(h)}$, i.e.

$$[X_{n+1} + {}_{n+1}\widehat{S}_{(h-1)}] - {}_{n}\widehat{S}_{(h)}$$

while before the interest was to estimate $\text{msepc}_n({}_{n}\widehat{S}_{(h)})$

Those two quantities estimate the same things, at different dates,

- ${}_{n}\widehat{S}_{(h)}$ is a predictor for S_h at time n
- $X_{n+1} + {}_{n+1}\widehat{S}_{(h-1)}$ is a predictor for S_h at time $n+1$,

Recall that

$$\widehat{p}_{n+1} = \frac{1}{n+1} X_{n+1} + \frac{n}{n+1} \widehat{p}_n$$

so that

$$X_{n+1} + (h-1) \cdot \widehat{p}_{n+1} - h \cdot \widehat{p}_n = \frac{n+1+(h-1)}{n+1} X_{n+1} + \frac{n-h}{n+1} \widehat{p}_n$$

If we admit that we are looking for the following quantity (as in Merz & Wüthrich (2008))

$$\text{msepc}_n(\widehat{CDR}_{n+1}) = \mathbb{E}\left([X_{n+1} + (h-1)\cdot \widehat{p}_{n+1} - h \cdot \widehat{p}_n]^2 \,|\mathcal{F}_n\right)$$

then

$$\text{msepc}_n(\widehat{CDR}_{n+1}) = \mathbb{E}\left(\left[\frac{n+h}{n+1}X_{n+1} + \frac{n-h}{n+1}\widehat{p}_n\right]^2 |\mathcal{F}_n\right)$$

Assuming that $\{X_1, \cdots, X_n\}$ and X_{n+1} are independent, then

$$\begin{aligned}\text{msepc}_n &= \mathbb{E}\left(\left[\frac{n+h}{n+1}X_{n+1}\right]^2 |\mathcal{F}_n\right) + \mathbb{E}\left(\left[\frac{n-h}{n+1}\widehat{p}_n\right]^2 |\mathcal{F}_n\right)\\ &+ \mathbb{E}\left(\frac{n+h}{n+1}X_{n+1}|\mathcal{F}_n\right)\cdot \mathbb{E}\left(\frac{n-h}{n+1}\widehat{p}_n|\mathcal{F}_n\right)\end{aligned}$$

i.e.

$$\begin{aligned}\text{msepc}_n &= \frac{(n+h)^2}{(n+1)^2}\mathbb{E}\left(X_{n+1}^2|\mathcal{F}_n\right)+\frac{(n-h)^2}{(n+1)^2}\mathbb{E}\left(\widehat{p}_n^2|\mathcal{F}_n\right)\\ &+ \frac{(n+h)(n-h)}{(n+1)^2}\mathbb{E}\left(X_{n+1}|\mathcal{F}_n\right)\cdot\mathbb{E}\left(\widehat{p}_n|\mathcal{F}_n\right)+\end{aligned}$$

Since $\mathbb{E}\left(\widehat{p}_n|\mathcal{F}_n\right)=\widehat{p}_n$, we can write

$$\text{msepc}_n = \frac{n+h)^2}{(n+1)^2}p+\frac{(n+h)(n-h)}{(n+1)^2}p\cdot\widehat{p}_n+\frac{(n-h)^2}{(n+1)^2}\widehat{p}_n^2$$

A *natural* estimator of that quantity is obtain as

$$\widehat{\text{msepc}}_n = \frac{(n+h)^2}{(n+1)^2}\widehat{p}_n+\frac{(n+h)(n-h)}{(n+1)^2}\widehat{p}_n^2+\frac{(n-h)^2}{(n+1)^2}\widehat{p}_n^2$$

i.e.

$${}_n\widehat{S}_{(h)}\widehat{\text{msepc}}_n = \frac{n^2-2nh}{(n+1)^2}\widehat{p}_n^2+\frac{(n+h)^2}{(n+1)^2}\widehat{p}_n$$

It is *usually* compared with the quantity that was calculated before, i.e.

$$\widehat{\text{msepc}}_n^{\color{red}k}({}_n\widehat{S}_h) = h\widehat{p}_n(1-\widehat{p}_n) + h^2\left(\widehat{p}_n - \widehat{p}_{\color{red}k}\right)^2$$

Mack's ultimate uncertainty

As shown in Mack (1993),

$$\widehat{\text{msep}}(\widehat{R}_i) = \widehat{C}_{i,\infty}^2 \sum_{j=n-i+1}^{n-1} \frac{\widehat{\sigma}_j^2}{\widehat{\lambda}_j^2} \left(\frac{1}{\widehat{C}_{i,j}} + \frac{1}{\widehat{S}_j} \right)$$

where S_j is the sum of cumulated payments on accident years before year $n-j$,

$$S_j = \sum_{i=1}^{n-j} C_{i,j}$$

Finally, it is possible also to derive an estimator for the aggregate msep (all accident years)

$$\widehat{\text{msep}}(\widehat{R}) = \sum \widehat{\text{msep}}(\widehat{R}_i) + 2\widehat{C}_{i,\infty}^2 \sum_{k=i+1}^{n} \widehat{C}_{k,n} \sum_{j=n-i+1}^{n-1} \frac{\widehat{\sigma}_j^2}{\widehat{\lambda}_j^2 S_j}$$

$$n \ x_1 = 1 \ x_2 = 0 \ x_n = 0 \ x_{n+1} = 0$$

X_{n+1} X_{n+2} X_{n+h} $n+1$

Mack's ultimate uncertainty

```
> library(ChainLadder)
>  source("http://perso.univ-rennes1.fr/arthur.charpentier/bases.R")
>  MackChainLadder(PAID)
MackChainLadder(Triangle = PAID)
  Latest Dev.To.Date Ultimate    IBNR Mack.S.E CV(IBNR)
1  4,456       1.000    4,456     0.0    0.000      NaN
2  4,730       0.995    4,752    22.4    0.639   0.0285
3  5,420       0.993    5,456    35.8    2.503   0.0699
4  6,020       0.989    6,086    66.1    5.046   0.0764
5  6,794       0.978    6,947   153.1   31.332   0.2047
6  5,217       0.708    7,367 2,149.7   68.449   0.0318
                  Totals
Latest:     32,637.00
Ultimate:   35,063.99
IBNR:        2,426.99
Mack S.E.:      79.30
CV(IBNR):        0.03
```

i.e. $\text{msepc}_6(\widehat{R}) = 79.30$.

GLM log-Poisson in triangles

Recall that we while to estimate

$$\mathbb{E}([R-\widehat{R}]^2) = \left[\mathbb{E}(R)-\mathbb{E}(\widehat{R})\right]^2 + \text{Var}(R-\widehat{R}) \approx \text{Var}(R)+\text{Var}(\widehat{R})$$

Classically, consider a log-Poisson model, were incremental payments satisfy

$$X_{i,j} \sim \mathcal{P}(\mu_{i,j}) \text{ where } \mu_{i,j} = \exp[\eta_{i,j}] = \exp[\gamma+\alpha_i+\beta_j]$$

Using the delta method, we get that *asymptotically*

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

where, since we consider a log link,

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

i.e., with an ODP distribution (i.e. $\text{Var}(X_{i,j}) = \varphi\mathbb{E}(X_{i,j})$), Thus, since the overall amount of reserves satisfies

$$\mathbb{E}\left([R-\widehat{R}]^2\right) \approx \sum_{i+j-1>n} \widehat{\varphi}\cdot\widehat{\mu}_{i,j} + \widehat{\boldsymbol{\mu}}'\widehat{\text{Var}}(\widehat{\boldsymbol{\eta}})\widehat{\boldsymbol{\mu}}$$

```
> an <- 6; ligne = rep(1:an, each=an); colonne = rep(1:an, an)
> passe = (ligne + colonne - 1)<=an; np = sum(passe)
> futur = (ligne + colonne - 1)> an; nf = sum(passe)
> INC=PAID
> INC[,2:6]=PAID[,2:6]-PAID[,1:5]
> Y = as.vector(INC)
> lig = as.factor(ligne); col = as.factor(colonne)
>
> CL <- glm(Y~lig+col, family=quasipoisson)
> Y2=Y; Y2[is.na(Y)]=.001
> CL2 <- glm(Y2~lig+col, family=quasipoisson)
> YP = predict(CL)
> p = 2*6-1;
> phi.P = sum(residuals(CL,"pearson")^2)/(np-p)
```

```
>  Sig = vcov(CL)
>  X = model.matrix(CL2)
>  Cov.eta = X%*%Sig%*%t(X)
>  mu.hat = exp(predict(CL,newdata=data.frame(lig,col)))*futur
>  pe2 = phi.P * sum(mu.hat) + t(mu.hat) %*% Cov.eta %*% mu.hat
>  cat("Total reserve =", sum(mu.hat), "prediction error =", sqrt(pe2),"\n")
Total reserve = 2426.985 prediction error = 131.7726
```

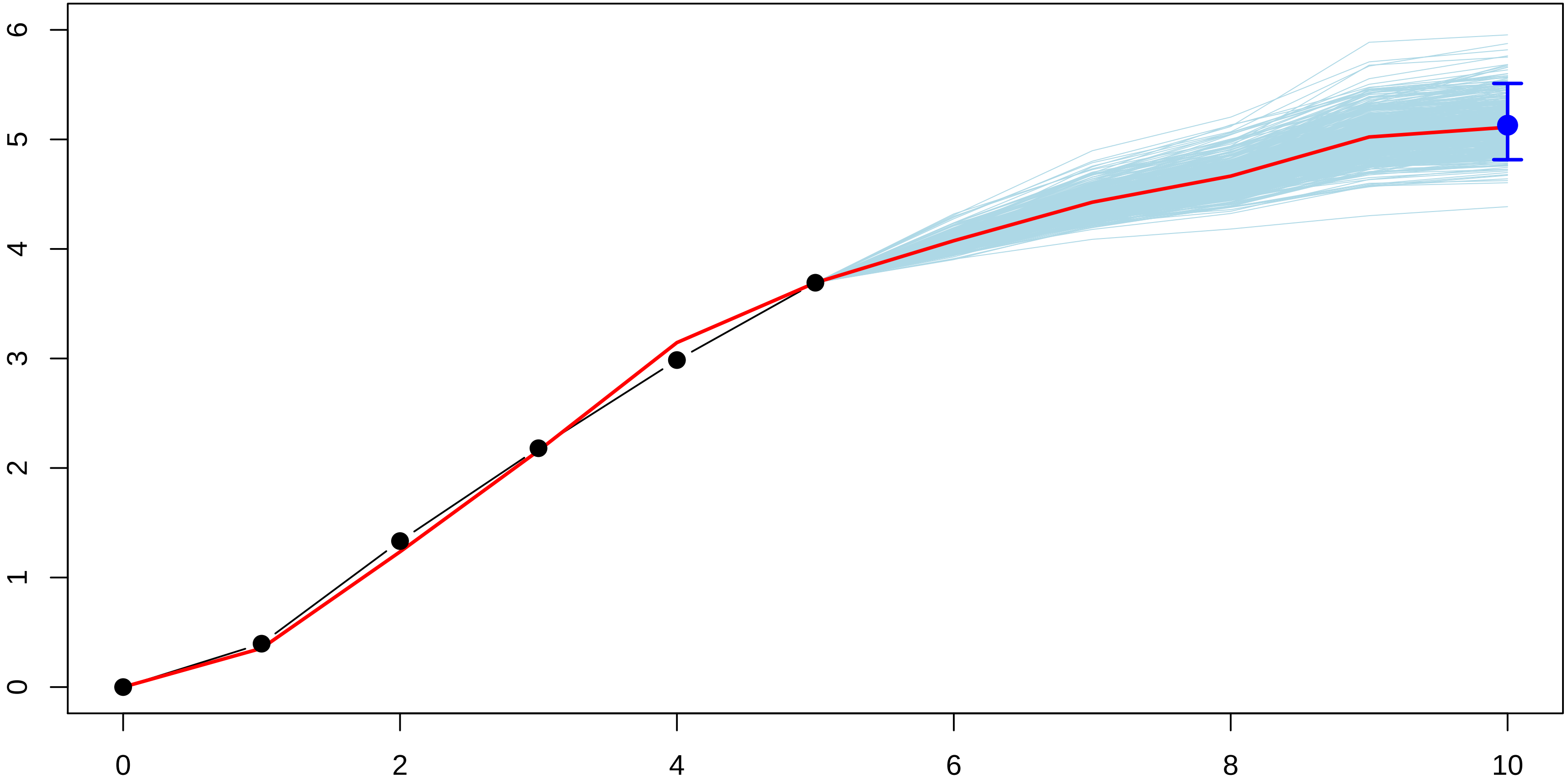
i.e. $\widehat{\mathbb{E}}(\widehat{R}-R)=131.77$.

GLM log-Poisson in triangles

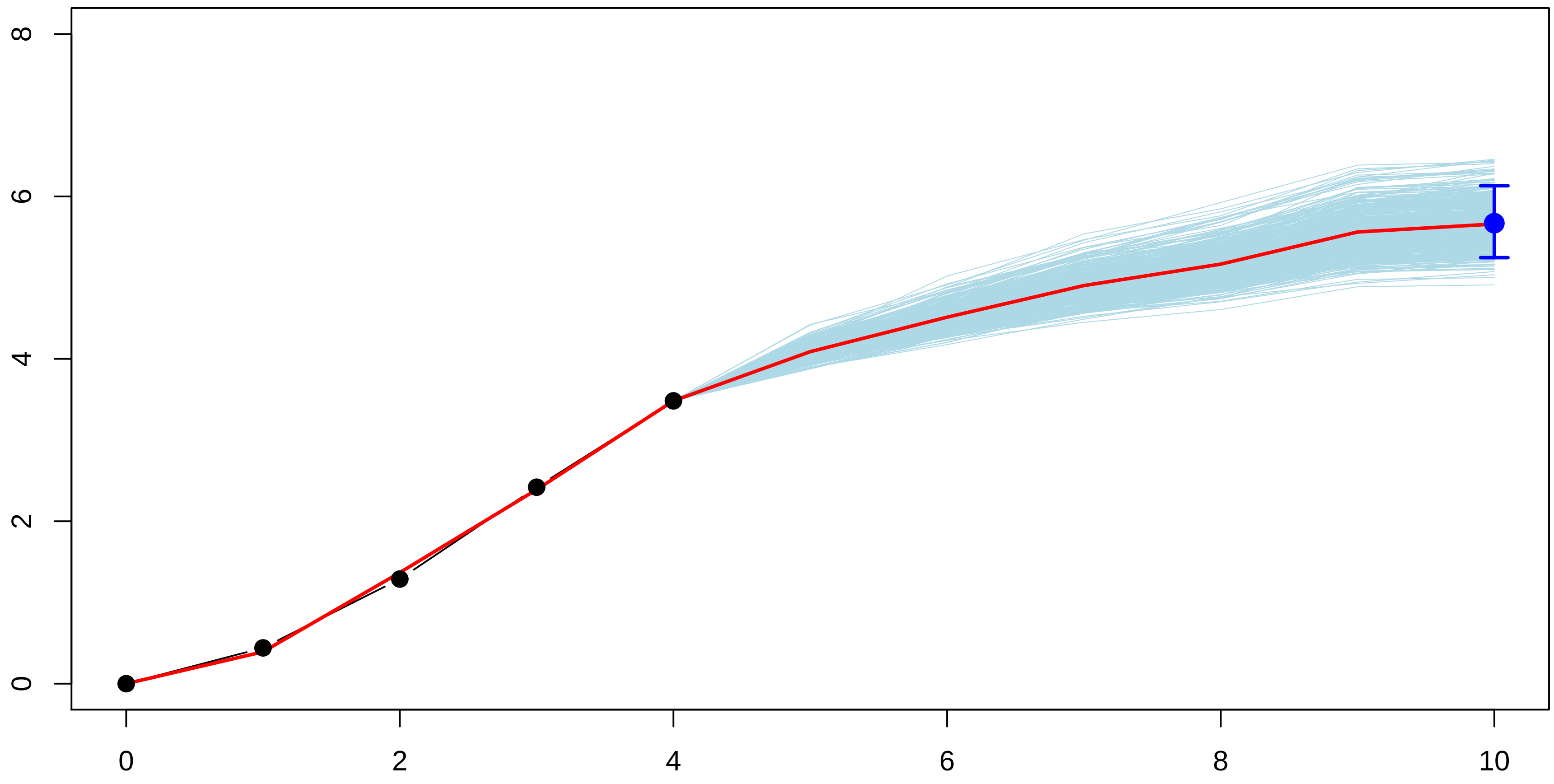
It is also possible to boostrap residuals to obtain *pseudo* triangles,

```
> CL <- glm(Y~lig+col, family=quasipoisson)
> E=residuals(CL,"pearson")
> Y0=predict(CL,newdata=data.frame(lig,col),type="response")
> Eb=sample(E,size=length(Y),replace=TRUE)
> Yb=Y0+Eb*sqrt(Y0)
> Yb[is.na(Y)]=NA
```

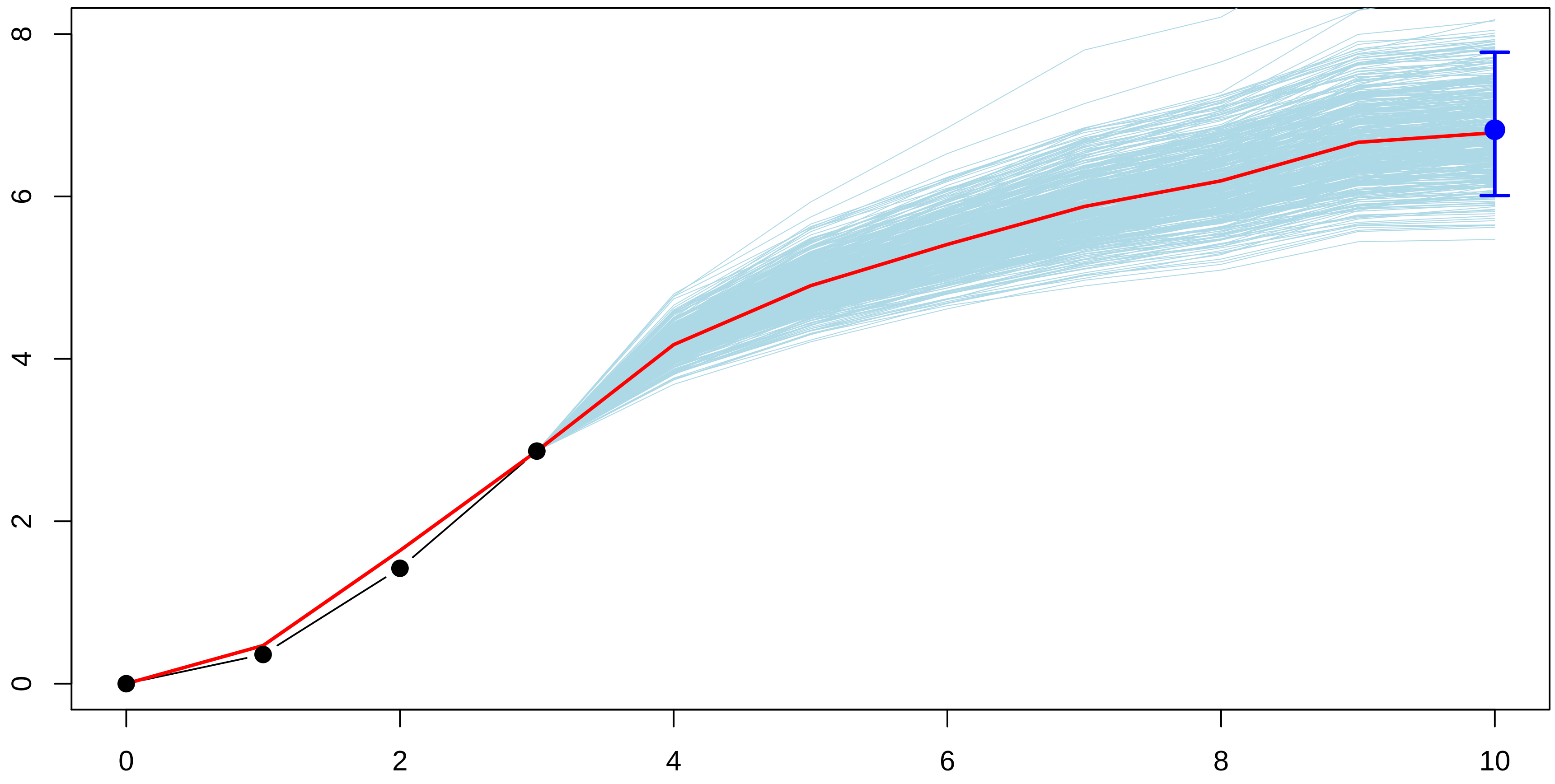
Bootstrap and GLM log-Poisson in triangles



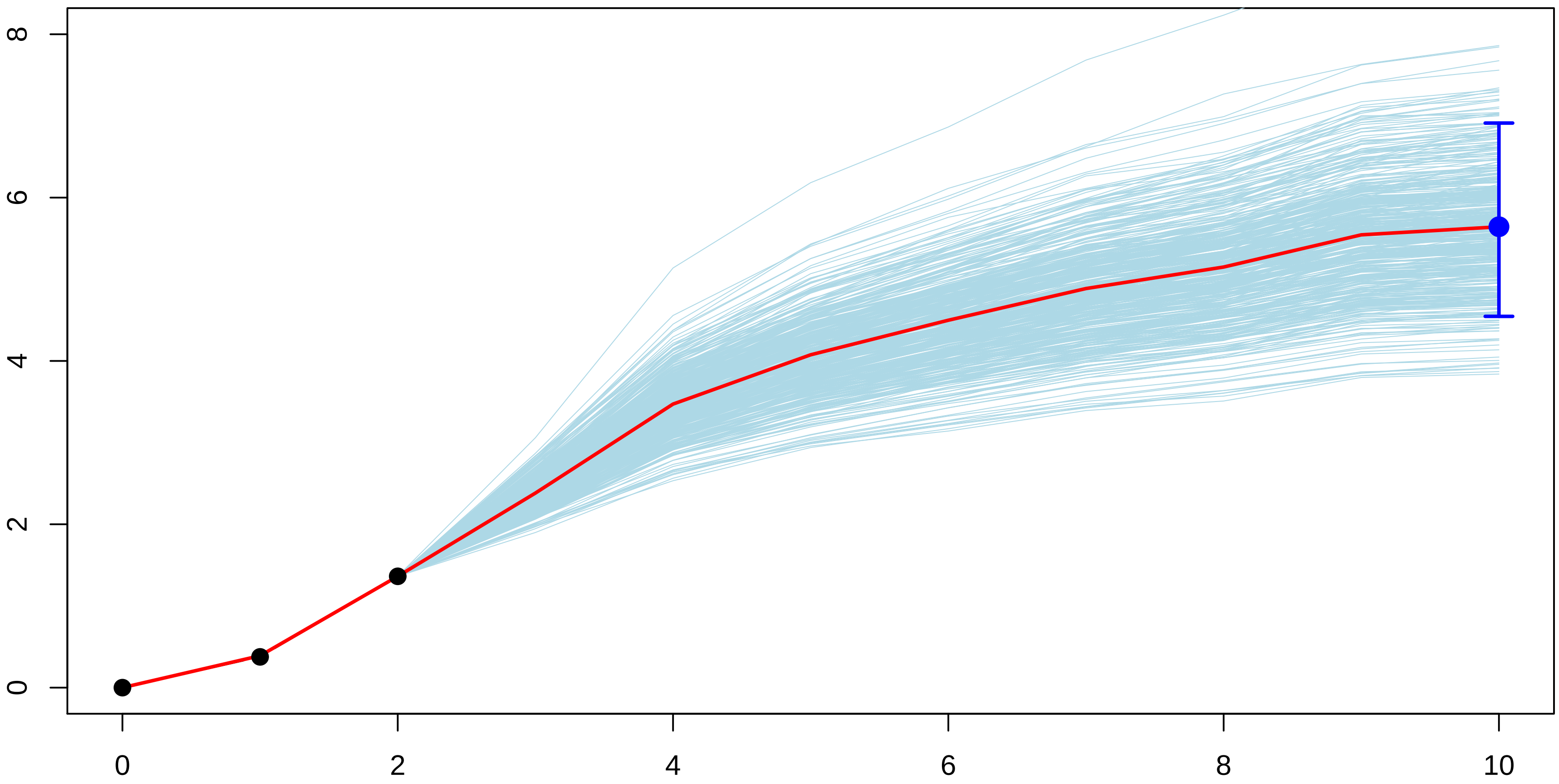
Bootstrap and GLM log-Poisson in triangles



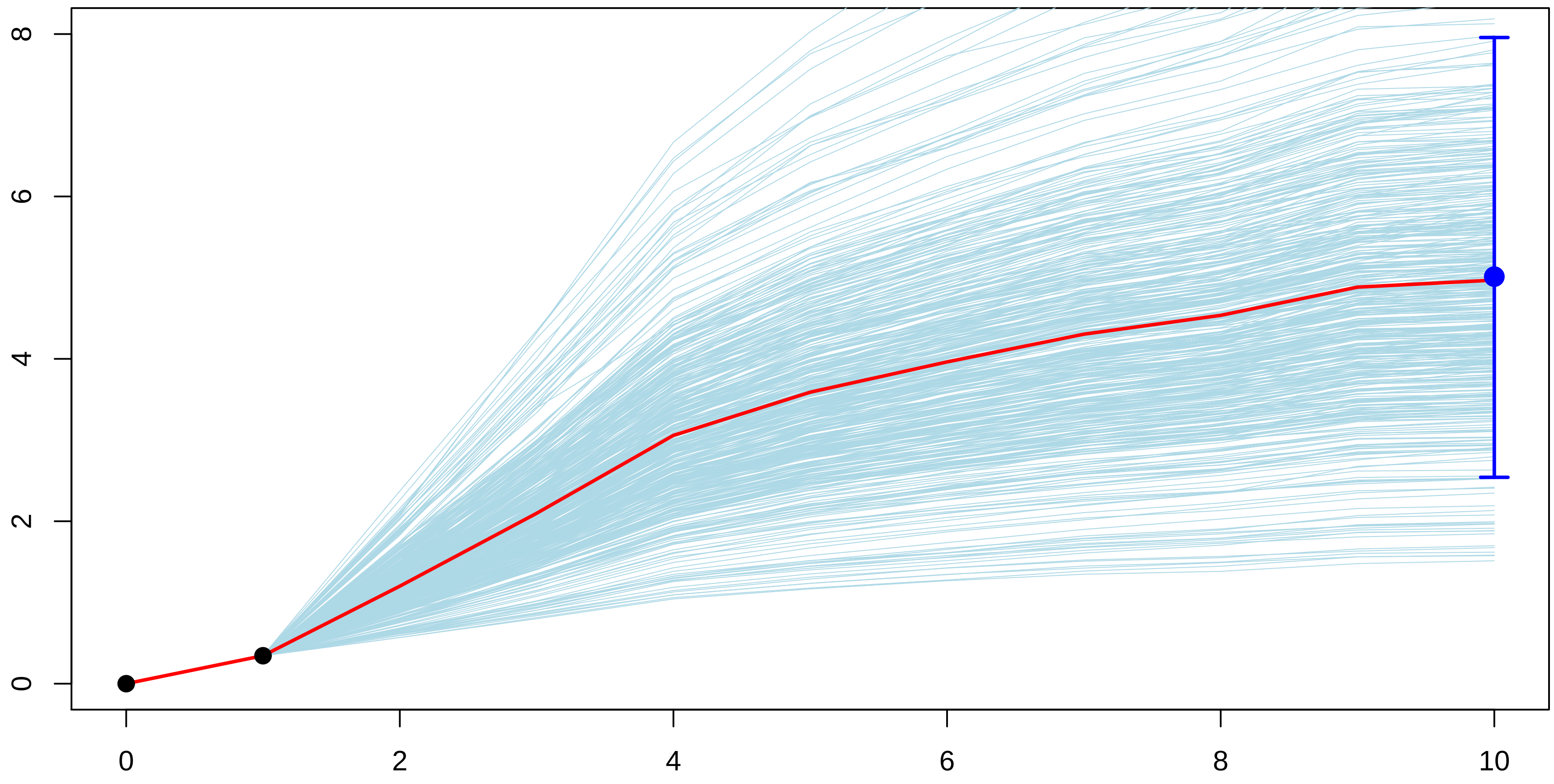
Bootstrap and GLM log-Poisson in triangles



Bootstrap and GLM log-Poisson in triangles

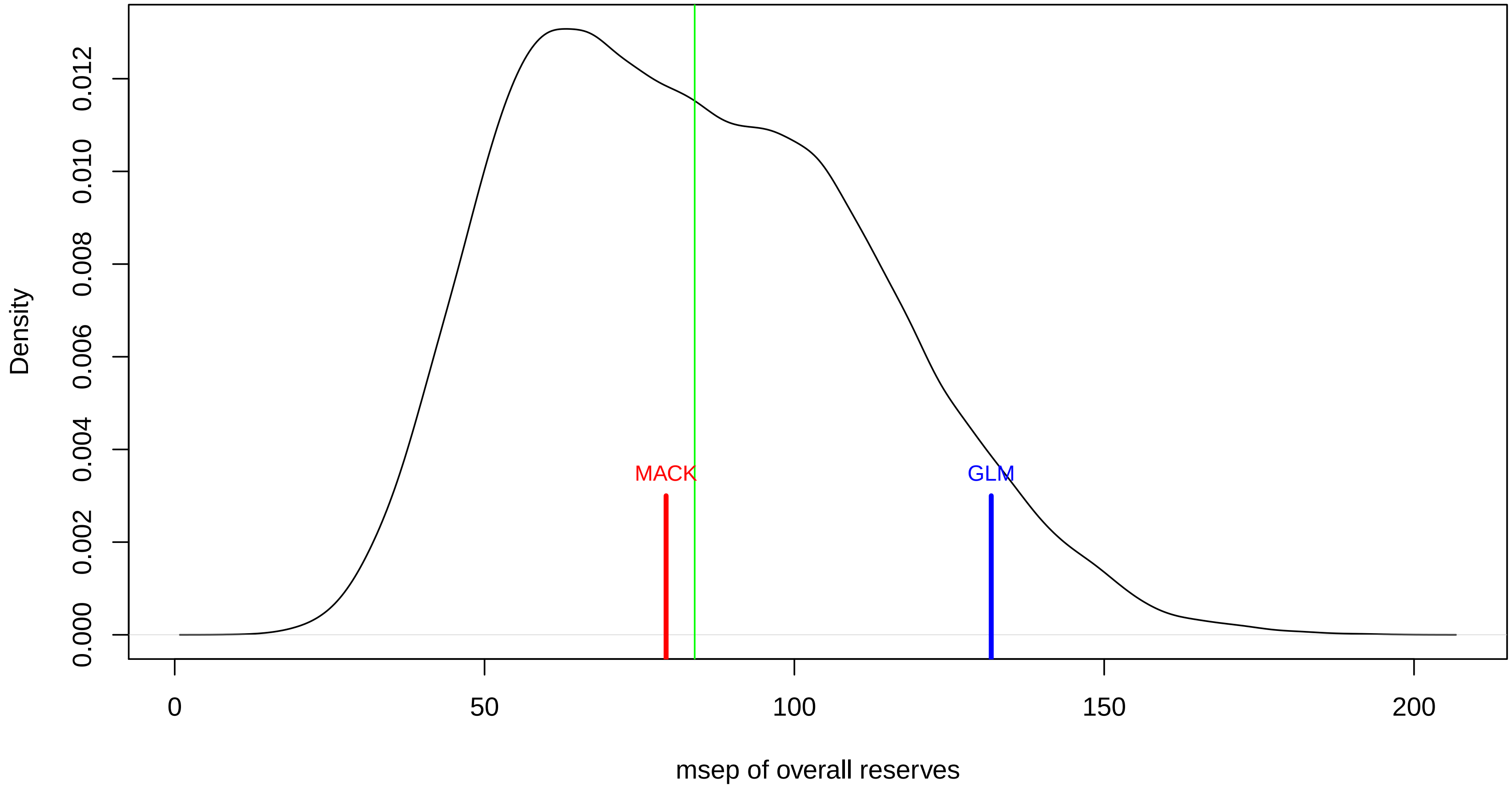


Bootstrap and GLM log-Poisson in triangles



Bootstrap and GLM log-Poisson in triangles

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



Merz & Wüthrich's one year uncertainty

Further, it can be proved that $(\text{CDR}_i(t))_t$'s are non correlated, and thus

$$\text{msepc}_t(\widehat{C}_{i,\infty}^t) = \text{Var}(C_{i,\infty}|\mathcal{F}_t) = \sum_{h\geq 1} \text{Var}(\text{CDR}_i(t+k)|\mathcal{F}_t)$$

which gives

$$\text{msepc}_{t-1}(\text{CDR}_i(t)) = \text{Var}(\text{CDR}_i(t)|\mathcal{F}_{t-1}) = \mathbb{E}(\text{CDR}_i(t)^2|\mathcal{F}_{t-1})$$

Merz & Wüthrich (2008) proved that the one year horizon error can be estimated with a formula similar to Mack (1993)

$$\widehat{\text{msepc}}_{t-1}(\text{CDR}_i(t)) = \widehat{C}_{i,\infty}^2 \left(\widehat{\Gamma}_{i,n} + \widehat{\Delta}_{i,n}\right)$$

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

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

Merz & Wüthrich (2008) mentioned that this term can be approximated as

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

using a simple development of $\prod(1+u_i) \approx 1 + \sum u_i$, but which is valid *only* if u_i is extremely small, i.e.

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

Implementing Merz& Wüthrich's formula

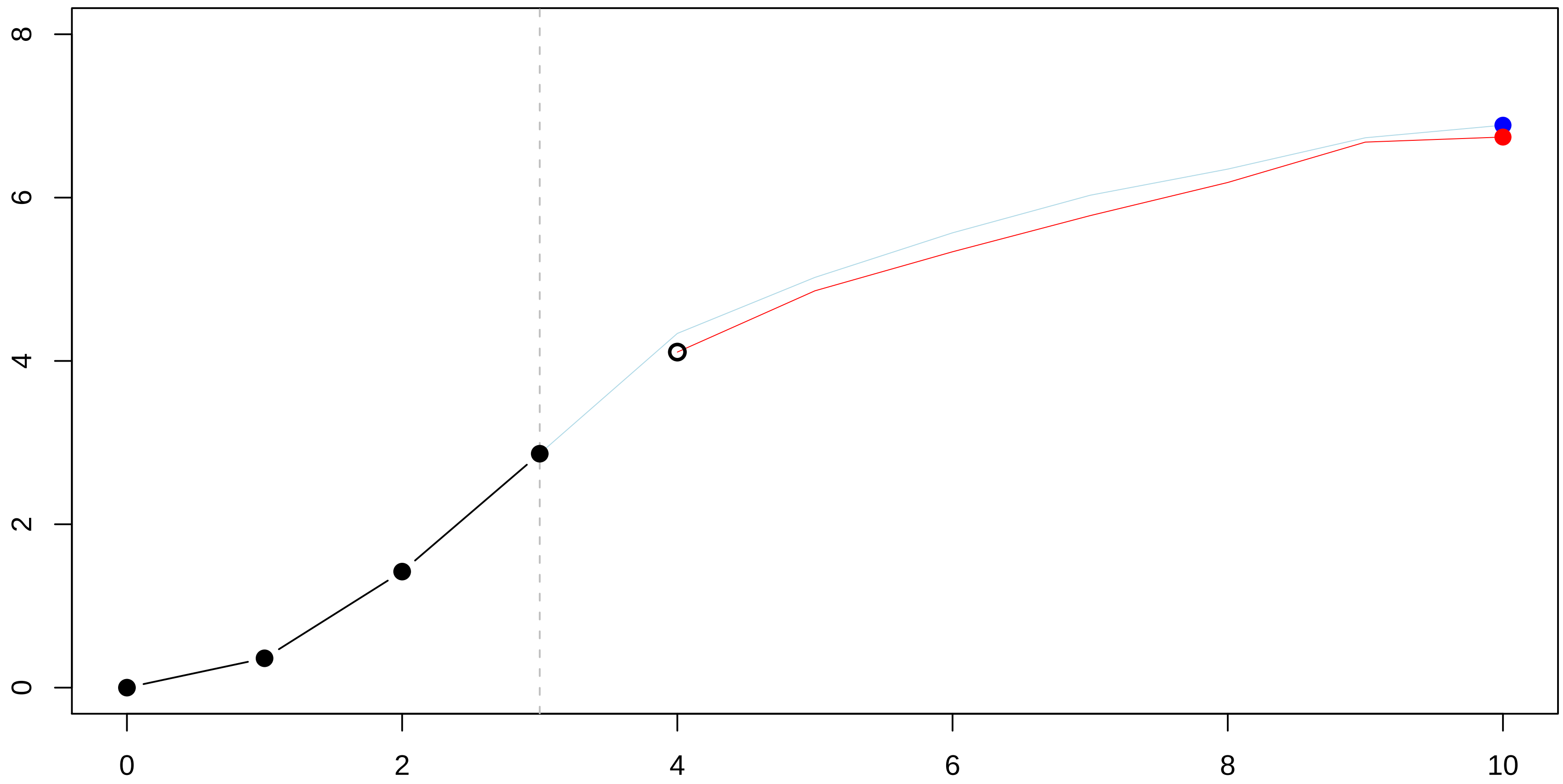
```
> source("http://perso.univ-rennes1.fr/arthur.charpentier/merz-wuthrich-triangle.R")
> MSEP_Mack_MW(PAID,0)
   MSEP Mack MSEP observable approche MSEP observable exacte
1  0.0000000                 0.000000               0.000000
2  0.6393379                 1.424131               1.315292
3  2.5025153                 2.543508               2.543508
4  5.0459004                 4.476698               4.476698
5 31.3319292                30.915407              30.915407
6 68.4489667                60.832875              60.832898
7 79.2954414                72.574735              72.572700
```

Implementing Merz& Wüthrich's formula

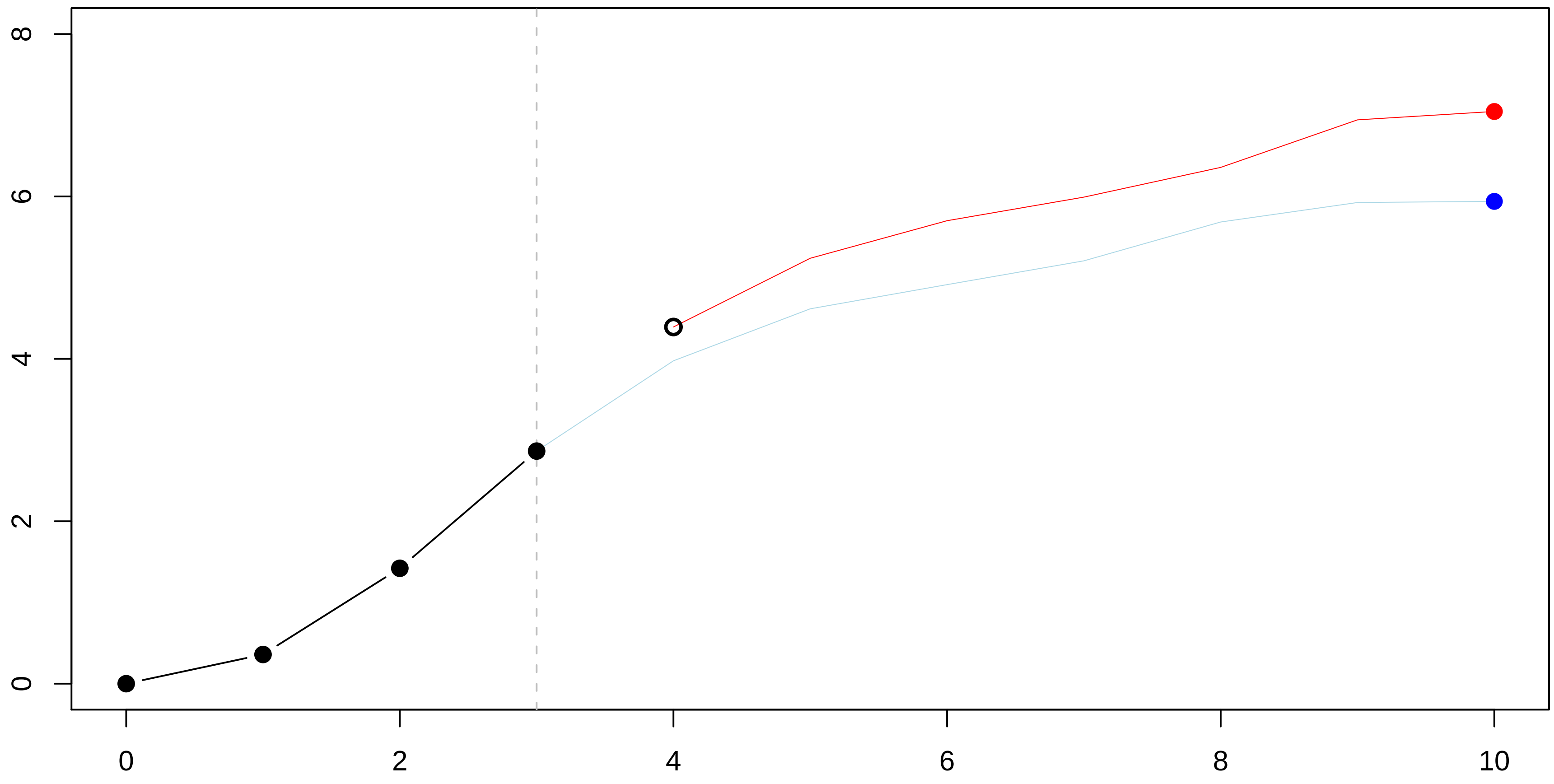
Could Merz& Wüthrich's formula end up with more uncertainty than Mack's

```
> Triangle = read.table("http://perso.univ-rennes1.fr/arthur.charpentier/
+            GAV-triangle.csv",sep=";")/1000000
> MSEP_Mack_MW(Triangle,0)
    MSEP Mack MSEP observable approche MSEP observable exacte
1  0.00000000                0.0000000              0.0000000
2  0.01245974                0.1296922              0.1526059
3  0.20943114                0.2141365              0.2144196
4  0.25800338                0.1980723              0.1987730
5  3.05529740                3.0484895              3.0655251
6 58.42939329               57.0561173             67.3757940
7 58.66964613               57.3015524             67.5861066
```

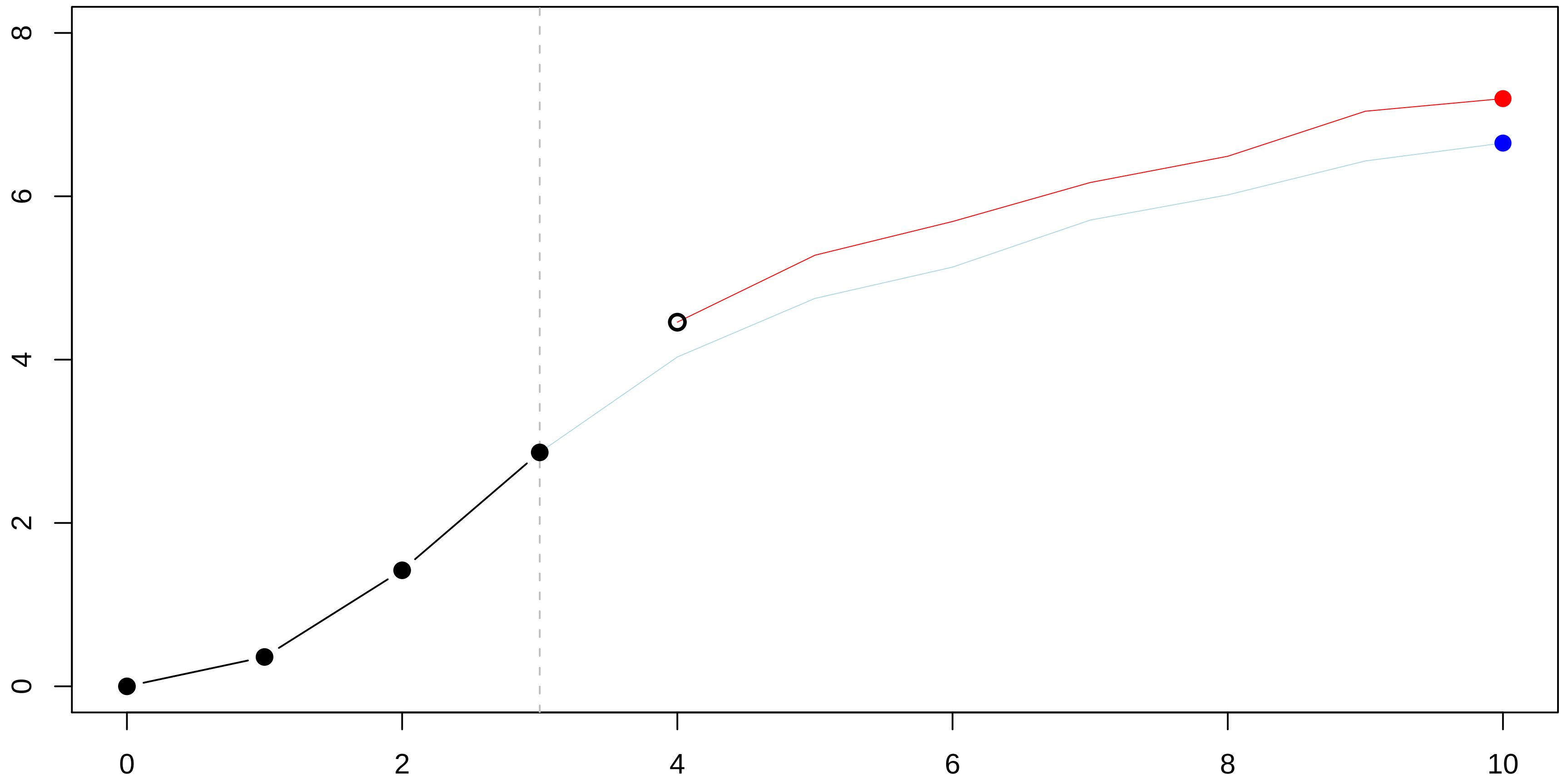
Bootstrap and one year uncertainty



Bootstrap and one year uncertainty

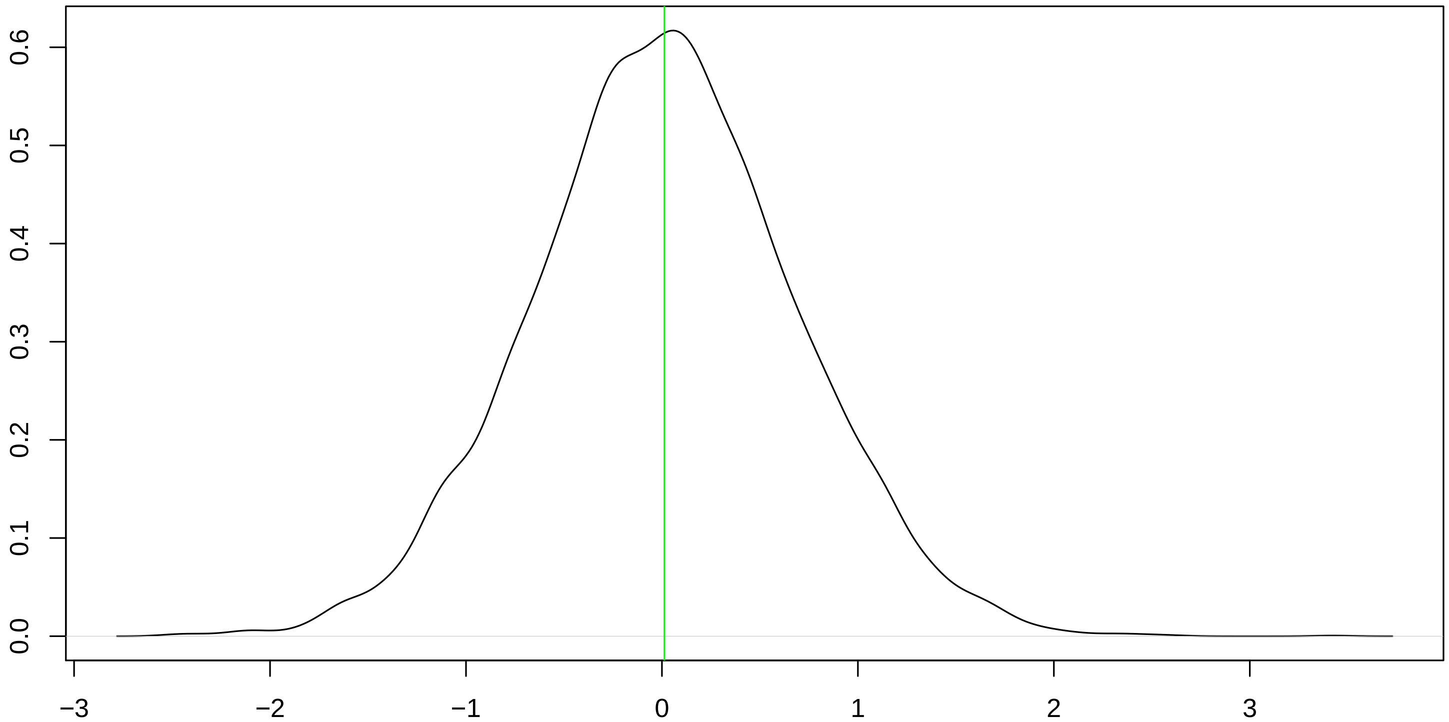


Bootstrap and one year uncertainty



Bootstrap and one year uncertainty

Here, we obtain the following distribution for the one year difference uncertainty



Note that $\mathbb{E}(CRD_i(t)|\mathcal{F}_t) = 0$ (i.e. neither boni nor mali should be expected).

Bootstrap and one year uncertainty

Further

```
> sd(DIFF)
[1] 0.6707964
```

to be compared with

```
> MSEP_Mack_MW(GenIns,0)
    MSEP Mack MSEP observable approche MSEP observable exacte

8  0.87488222                0.62968103             0.62975724
```

$$\prod(1+u_i)\approx 1+\sum u_i$$

$$\widehat{\text{msepc}}_{n-1}(\text{CDR}_i(n)) = \widehat{C}^2_{i,\infty}\left(\widehat{\Gamma}_{i,n}+\widehat{\Delta}_{i,n}\right)$$

$$\text{msep}(\widehat{R}_i) = \mathbb{E}([\widehat{R}_i - R_i]^2 | \mathcal{F}_{n-i})$$

$$\mathbb{E}(X_{i,j}) = \exp(\gamma + \alpha_i + \beta_j)$$

$$\mathbb{E}(X_{i,j}) = \text{Var}(X_{i,j})$$

$$Y \sim \mathcal{P}(\exp[\gamma + \alpha X])$$

$$Y \sim \mathcal{N}(\exp[\gamma + \alpha X], \sigma^2)$$

$$Y \sim \mathcal{N}(\gamma + \alpha X, \sigma^2)$$

$$\widehat{\text{msep}}(\widehat{R}_i) = \widehat{C}_{i,n}^2 \sum_{j=n-i+1}^{n-1} \frac{\widehat{\sigma}_{j+1}^2}{\widehat{\lambda}_{j+1}^2} \frac{1}{\widehat{C}_{i,j}} + \frac{\widehat{\sigma}_{j+1}^2}{\widehat{\lambda}_{j+1}^2} \frac{1}{\widehat{S}_j^n}$$

$$\widehat{S}_j^n = \sum_{k=1}^{n-j} \widehat{C}_{k,j}$$

$$\widehat{\text{msep}}(\widehat{\text{CDR}}_{i+1}) = \widehat{C}_{i,n}^2 \left(\frac{\widehat{\sigma}_{j+1}^2}{\widehat{\lambda}_{j+1}^2} \frac{1}{C_{i,n-i}} + \frac{\widehat{\sigma}_{j+1}^2}{\widehat{\lambda}_{j+1}^2} \frac{1}{\widehat{S}_{n-i}^n} \right)$$
$$+ \widehat{C}_{i,n}^2 \sum_{j=n-i+2}^{n-1} \frac{\widehat{\sigma}_{j+1}^2}{\widehat{\lambda}_{j+1}^2} \frac{1}{\widehat{C}_j} \left(\frac{C_{n-j+1,j}}{\widehat{S}_j^n} \right)^2$$