

Can you see the quality of a financial risk?

Hans Bühlmann, ETH Zürich

On this common ASTIN–AFIR day, I should like to address a topic of equal importance in insurance and finance

Performance measurement

Performance prediction

Performance is measured by appropriate ratios

Insurance: Loss ratios (with or without costs)

Finance: Return rates (on investment, on equity etc.)

Log return rates

Both loss ratios and return rates are used routinely. The more it is astonishing that the ongoing actuarial techniques for dealing with these two quantities have very little in common. Let me make my point even more provocative:

I believe strongly that the thinking developed in insurance to deal with loss ratios should also be used in the finance sector for dealing with returns. Obviously some modifications are appropriate and I shall talk about them. But – and that is my message to the specialists in finance – the basic concepts are there to be used, developed over more than eighty years by casualty actuaries and by members of ASTIN on one side and by statisticians on the other.

1. Insurance and loss ratios

In insurance the customary ratio is

$$\frac{\text{Incurred claims (with or without costs)}}{\text{volume}} = X$$

for given periods (e.g. years, semesters, quarters). Volume may mean

premiums earned
 sums insured
 total salary

depending on traditions in different branches and parts of the world.

As actuaries we are then typically confronted with the following scheme relating to a group of risks all considered to belong to the same collective

| | | | | | | |
|-------------|----------|-----|-------------|-----|-------------|---|
| X_{11} | X_{12} | ... | X_{ij} | ... | X_{1N} | |
| X_{21} | ... | ... | ... | ... | X_{2N} | |
| : | : | | : | | : | |
| : | : | | X_{ij} | | : | $X_{ij} \sim$ loss ratio for period i |
| : | : | | : | | : | risk j |
| X_{n1} | X_{n2} | ... | X_{nj} | ... | X_{nN} | |
| | | | | | | |
| $X_{n+1,1}$ | ... | ... | $X_{n+1,j}$ | ... | $X_{n+1,N}$ | |

Quantities relating to the same individual risk are in the same column.

Rows $1, 2, \dots, n$ are observed,
 row $n + 1$ is to be predicted.

Modelling of individual risk

Individual risk may physically mean

- a group life contract
- a tariff position in fire insurance
- a reinsurance treaty etc.

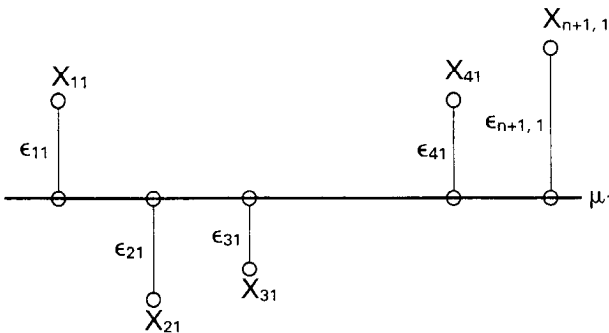
in short, the smallest statistical unit considered in our analysis.

Take one such individual risk (e.g. number 1) and its loss ratio column

$$\begin{array}{c} X_{11} \\ X_{21} \\ \vdots \\ X_{n1} \\ \hline X_{n+1,1} \end{array}$$

They fluctuate around a common theoretical value μ_1

Figure 1



We write
$$X_{i1} = \underset{\substack{\uparrow \\ \text{signal}}}{\mu_1} + \underset{\substack{\uparrow \\ \text{noise}}}{\epsilon_{i1}} \quad i = 1, \dots, n + 1$$

Question: After we have observed n ratios,

$$X_{11}, X_{21}, \dots, X_{n1}$$

How can we decompose them into signal?
noise?

This is, how electrical engineers talk about this problem and they call the decomposition filtering.

We actuaries call the signal the true premium rate and the decompositions name is credibility.

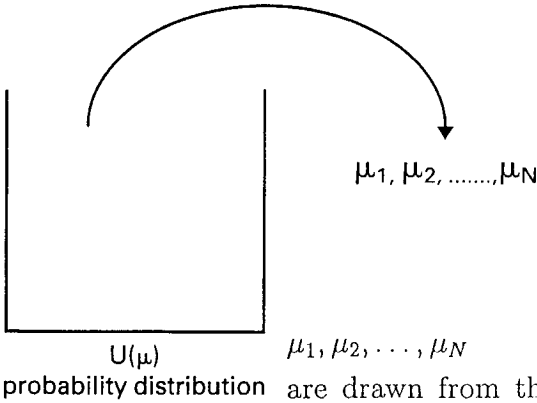
Unfortunately this true premium can not be observed, still it is probably the most important item, like the signal in telecommunication. This remark should justify the title chosen for this lecture.

Modelling of the Collective

Premium rating in insurance has always relied heavily on collective data. They are typically abundant whereas individual data are scarce. What we mean by a collective is typically a matter of definition by convenience. The requirement of a homogeneous collective – still found in older textbooks of insurance – is unnecessary, I would say even counterproductive.

Reverting to our modelling of the individual risk, the best way to model the collective is an urn, that contains all possible individually correct premiums μ

Figure 2



are drawn from this urn in an i.i.d fashion, where the content of the urn is described by the structural function $U(\mu)$.

If you have no data on an individual risk (e.g. number 1) you have no clue to guess its correct premium μ_1 , but as you know the structural distribution your best guess is

$$m = E[\mu] = \int \mu dU(\mu).$$

Filtering = Credibility formula

After you have made the observations $X_{11}, X_{21}, \dots, X_{n1}$ you estimate μ_1 by

$$\begin{aligned} \hat{\mu}_1 &= m + Z [\bar{X}_{.1} - m] && \text{formula as used} \\ & && \text{by electrical engineers} \\ &= Z \bar{X}_{.1} + (1 - Z)m && \text{formula used} \\ & \quad \uparrow && \text{by actuaries} \\ & \text{credibility weight} \end{aligned}$$

2. A quick technical summary

I have presented the most simple case

$$\begin{aligned}
 (1) \quad \text{noise } \varepsilon_{ij} \text{ all i.i.d.} & & E[\varepsilon_{ij}] &= 0 \\
 & & \text{Var}[\varepsilon_{ij}] &= \sigma^2 \\
 & & \text{Var}[\mu] &= \tau^2
 \end{aligned}$$

$$Z = \frac{n}{n + \frac{\sigma^2}{\tau^2}} \quad \begin{array}{l} \text{American actuaries} \\ \text{Whitney 1918 etc.} \end{array}$$

Let me generalize as follows:

You draw the individual distribution ϑ from the collective non not only the individual μ .

Hence we write from here on:

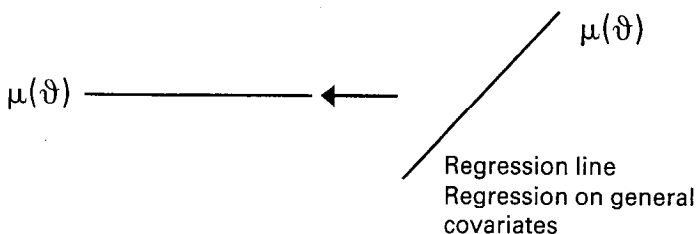
$$\mu(\vartheta) \quad \sigma(\vartheta) \quad \text{etc.}$$

$$\begin{aligned}
 (2) \quad \text{replace either } e_{ij} &\leftarrow \sigma(\vartheta_j)\varepsilon_{ij} \\
 \text{or } e_{ij} &\leftarrow \sigma(\vartheta_j)\frac{\varepsilon_{ij}}{\sqrt{V_{ij}}}
 \end{aligned}$$

The second substitution leads to the most commonly used model $V_{ij} \sim$ volume in period i for risk j

(3) replace: $\mu(\vartheta) \leftarrow$ Regression line or regression on general covariates (Hachemeister 1975).

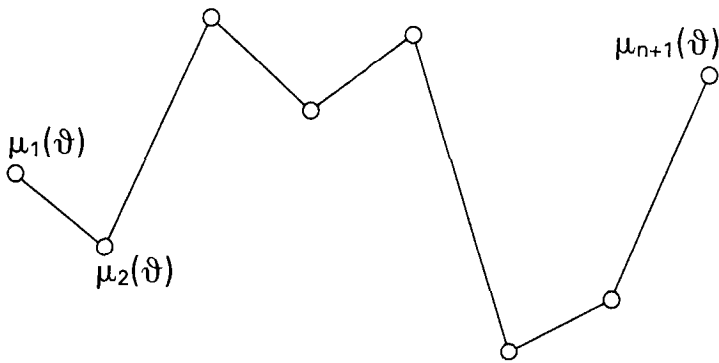
Figure 3



It is the next generalizer which I suggest that to be used in finance to model returns or better log returns.

(4) $\mu(\vartheta)$ becomes a stochastic curve

Figure 4



Professor Cramer 1967 (?), Gerber/Jones 1975, Kremer 1982, Sundt 1981, 1983

3. Finance and returns on investment (log returns)

We use the same schema as for loss ratios in insurance, but we interpret now

$$X_{ij} = \log \frac{S_i^{(j)}}{S_{i-1}^{(j)}}$$

where $S^{(j)}(j = 1, 2, \dots, N)$ denote values of financial instruments chosen from the same collective (e.g. bonds with same rating, stocks with same rating)

| | | | |
|---------------|---------------|---------------|----------------------------|
| X_{11} | X_{12} | X_{1N} | |
| X_{21} | X_{22} | | observed |
| \vdots | | | |
| X_{n1} | X_{n2} | X_{nN} | |
| $X_{n+1,1}$ | $X_{n+1,2}$ | $X_{n+1,N}$ | next period |
| ϑ_1 | ϑ_2 | ϑ_N | quality (never observable) |

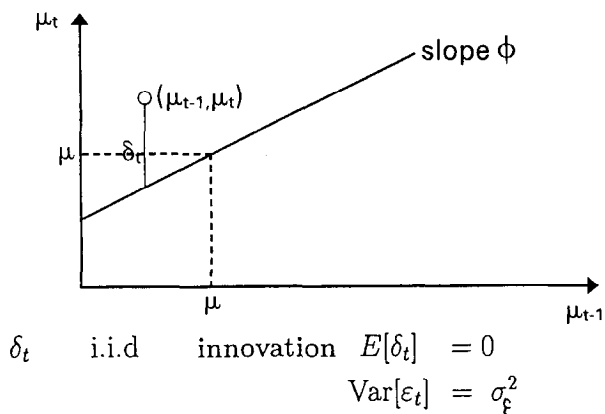
Modelling: (in the spirit of (4))

$$\mu_t(\vartheta) = \underbrace{\mu(\vartheta)}_{\text{long term level}} + \Phi \left[\underbrace{\mu_{t-1}(\vartheta)}_{\text{long term level}} - \underbrace{\mu(\vartheta)}_{\text{long term level}} \right] + \delta_t$$

In mathematical statistics this is called an autoregressive model.

AR (1) on the μ -space (state-space)

Figure 5



Observations

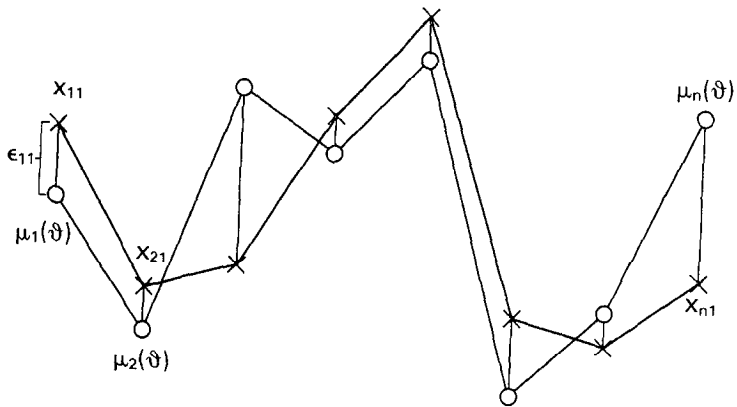
$$X_{tj} = \mu_t(\delta_j) + \begin{array}{c} \varepsilon_{tj} \\ \uparrow \\ \text{noise i.i.d} \end{array} \quad \begin{array}{l} E[\varepsilon_t] = 0 \\ \text{Var}[\varepsilon_t] = \sigma_\varepsilon^2 \end{array}$$

It is instructive to draw both the

μ -curve (red)

X -curve (black)

Figure 6



Problem: You observe the black X -curve
 You want to filter the red μ -curve

Observe: $\mu(\vartheta)$ from collective urn
 Φ fixed (could be generalized)
 $\mu_t(\vartheta)$ depends on collective urn
 + innovations
 δ_t innovations do not depend on ϑ
 (generalization meaningful?)

How to work with this model?

Structural parameters $E[\mu(\vartheta)] = m$
 $\text{Var}[\mu(\vartheta)] = \tau^2$ collective variance
 Φ
 $\text{Var}[\delta_t] = \sigma_\delta^2$ variance of innovation
 $\text{Var}[\varepsilon_t] = \sigma_\varepsilon^2$ variance of noise

estimation Peter Bühlmann
 Hans Andresen (student of Ragnar Norberg)

Difference between innovation and noise

$$\begin{aligned}\mu_t &= \mu + \Phi [\Phi(\mu_{t-2} - \mu) + \delta_{t-1}] + \delta_t \\ &= \mu + \delta_t + \Phi\delta_{t-1} + \Phi^2\delta_{t-2} + \Phi^3\delta_{t-3} + \dots\end{aligned}$$

$$X_t = \mu + (\delta_t + \Phi\delta_{t-1} + \Phi^2\delta_{t-2} + \dots) + \varepsilon_t$$

Noise: deviation only at time t

Innovation: deviation remains active over the whole time range
(discounted by Φ)

Observe that under stationary conditions

$$\text{Var}[\mu_t(\vartheta)] = \text{Var} \left[\mu(\vartheta) + \frac{\sigma_\delta^2}{1 - \Phi^2} \right]$$

and

$$\mu_t(\vartheta) - \mu(\vartheta) \text{ independent of } \mu(\vartheta)$$

4. Applying the Kalman filter

The good news is that the just described model is taylor-made for applying the Kalman-filter (see e.g. Abraham/Ledolter 1983).

As we can do that for each individual risk we use in this section only one index which denotes time

$$\begin{array}{l} \text{State vector} \\ \text{(dimension 2)} \end{array} \quad \begin{pmatrix} \mu(\vartheta) \\ \mu_t(\vartheta) \end{pmatrix} = S_t$$

$$\begin{array}{l} \text{Observation} \\ \text{(dimension 1)} \end{array} \quad X_t = HS_t + \varepsilon_t$$

$$\quad \quad \quad \uparrow$$

$$\quad \quad \quad H = (0, 1)$$

Movement of state vector

$$S_t = AS_t + \begin{pmatrix} 0 \\ \delta_t \end{pmatrix}$$

$$\quad \quad \quad \uparrow$$

$$A = \begin{pmatrix} 1 & 0 \\ 1 - \Phi & \Phi \end{pmatrix}$$

Denote by $S_{t/k} \sim$ best estimate of S_t based on
 observations X_1, X_2, \dots, X_k

$P_{t/k} \sim$ expected square deviation matrix
 between $S_{t/k}$ and S_t

and use the following initial values

$$S_{0/0} = \begin{pmatrix} m \\ m \end{pmatrix}$$

For $P_{0/0}$ we may take two different views.

First case:

We start at time 0 with $S_0 = \begin{pmatrix} \mu(\vartheta) \\ \mu(\vartheta) \end{pmatrix}$

This leads to $P_{0/0} = \begin{pmatrix} \tau^2 & \tau^2 \\ \tau^2 & \tau^2 \end{pmatrix}$

Second case:

We start at time 0 with $S_0 = \begin{pmatrix} \mu(\vartheta) \\ \mu_t(\vartheta) \end{pmatrix}$

assuming the stationary distribution for S_0 .

This leads to $P_{0/0} = \begin{pmatrix} \tau^2 & \tau^2 \\ \tau^2 & \tau^2 + \frac{\sigma_\delta^2}{1-\phi^2} \end{pmatrix}$

The Kalman equations are $(t = 0, 1, 2, \dots)$

Movement

$$S_{t+1/t} = AS_{t/t}$$

$$P_{t+1/t} = AP_{t/t}A' + \begin{pmatrix} 0 & 0 \\ 0 & \sigma_\delta^2 \end{pmatrix}$$

Updating

$$S_{t+1/t} = K_{t+1}(X_t - HS_{t+1/t})$$

$$P_{t+1/t} = P_{t+1/k} - K_{t+1}HP_{t+1/t}$$

where

$$K_{t+1} = P_{t+1/t} H' \underbrace{(H P_{t+1/t} H' + \sigma_\varepsilon^2)^{-1}}_{\text{Scalar}}$$

It is instructive to have a look at K_{t+1} . You find

$$K_{t+1} = \left(\frac{\text{second column}}{\text{lower right element} + \sigma_\varepsilon^2} \right) \text{ of matrix } P_{t+1/t}$$

e.g. assuming first case initial conditions

$$K_1 = \begin{pmatrix} \frac{\tau^2}{\tau^2 + \sigma_\delta^2 + \sigma_\varepsilon^2} \\ \frac{\tau^2 + \sigma_\delta^2}{\tau^2 + \sigma_\delta^2 + \sigma_\varepsilon^2} \end{pmatrix} = \begin{pmatrix} Z_{11} \\ Z_{21} \end{pmatrix}$$

Under second case initial conditions σ_δ^2 has to be replaced by $\sigma_\delta^2(1 + \frac{\phi^2}{1-\phi^2})$

$$Z_{11} \sim \text{credibility of } X_1 \text{ for } \mu(\vartheta)$$

$$Z_{21} \sim \text{credibility of } X_1 \text{ for } \mu_1(\vartheta)$$

5. Outlook

We have discussed a special model to predict returns. The same techniques can be used to predict volatilities. The latter task seems to be even more in practical demand.

The basic idea which I was driving at was that of bringing two approaches together. On one side the understanding of the prior distribution (structural distribution) as a description of a collective

into which the individual risk is embedded. On the other side the time series approach to model the evolution of the individual risk quality. The Kalman filter has turned out to be an appropriate tool that can be used for prediction even in this combined view, which renders the technique useful for practical applications.

Let us therefore hope that many such applications shall be reported about in our future meetings.

Literature

- Abraham / Ledolter "Statistical Methods for Forecasting", (Wiley, 1983)
- Jones, D.A. and Gerber, H.U. "Credibility Formulas of the Updating Type", Transactions Society of Actuaries, vol. XXVII, 1975
- Kremer, E. "Credibility for some Evolutionary Models", Scandinavian Actuarial Journal (1982)
- Sundt, B. "Recursive Credibility Estimation", Scandinavian Actuarial Journal (1981)
- Sundt, B. "Finite Credibility Formulae in Evolutionary Models", Scandinavian Actuarial Journal (1983)
- Whitney, A.W. "The Theory of Experience Rating", Proceedings of Casualty Actuarial Society, vol. 4, 1918

