Statistical approach to operational risk measurement using *R*

Roberto Ugoccioni
Sanpaolo IMI Group
Torino, Italy

ORM UNIT di GRUPPO

useR!

GRUPPO SANPAOLO IMI

Operational risk measurement

- OR: Failures of normal processes (mistakes, frauds, robberies, liabilities, ...)
- Two complementary approaches:
  - historical data (backward-looking)
  - scenario analysis (forward-looking)
- Actuarial method: compose
  - Frequency distribution
  - Severity distribution

Why *R* and how

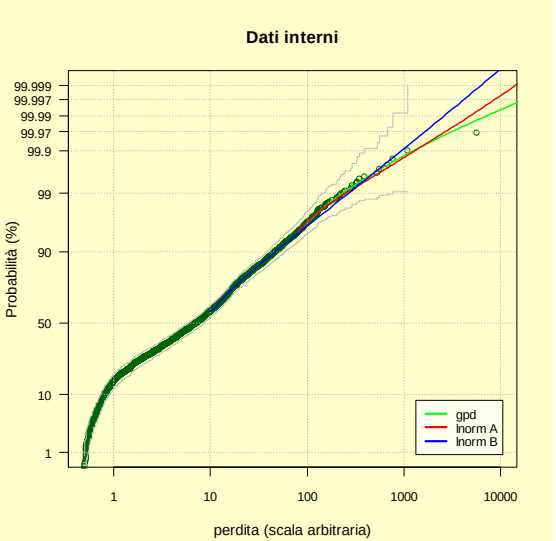
- Why use *R*?
  - no best-practice in the field, few existing tools
  - powerful, complete language
  - flexible framework
- How *R* is used at Sanpaolo IMI:
  - methodological research
  - application prototyping
  - production environment

Historical loss data analysis

- For each risk class (i.i.d.):
  - Fit distributions (maximum likelihood) to internal and external data
  - Choose best fits (GOF tests)
  - Compose internal/external distributions
  - Use FFT to compute 1-year period aggregate, including insurance effects
- Measure rank correlations between risk classes in the data
- Aggregate classes using copulas

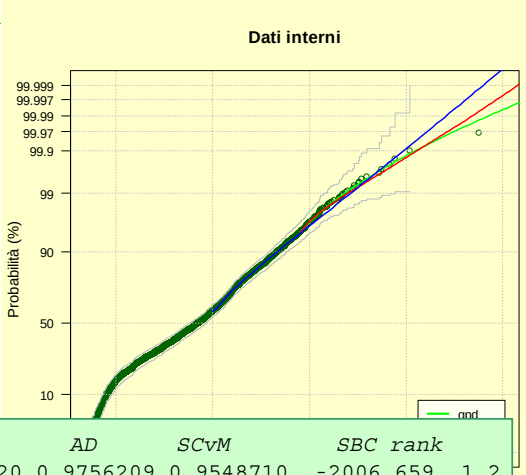
Example loss data analysis

- Fit distributions (maximum likelihood) to internal and external data
- Choose best fits (GOF tests)
- Compose internal/external distributions
- Use FFT to compute 1-year period aggregate, including insurance effects



Example loss data analysis

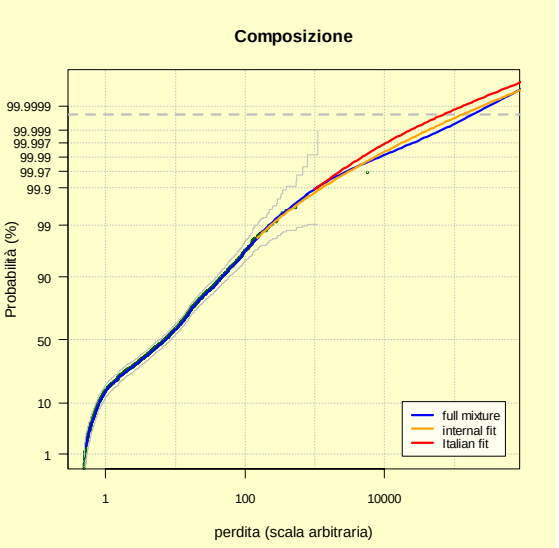
- Fit distributions (maximum likelihood) to internal and external data
- Choose best fits (GOF tests)
- Compose internal/external distributions
- Use FFT to compute 1-year period aggregate, including insurance effects



	chisquare	KS	AD	SCvM	SBC	rank
gpd	0.3517447	0.8666720	0.9756209	0.9548710	-2006.659	1.2
ln.B	0.5970360	0.7966336	0.6693486	0.5158578	-2009.520	1.8
ln.A	0.2199431	0.2291122	0.3275139	0.3912281	-14138.940	3.0

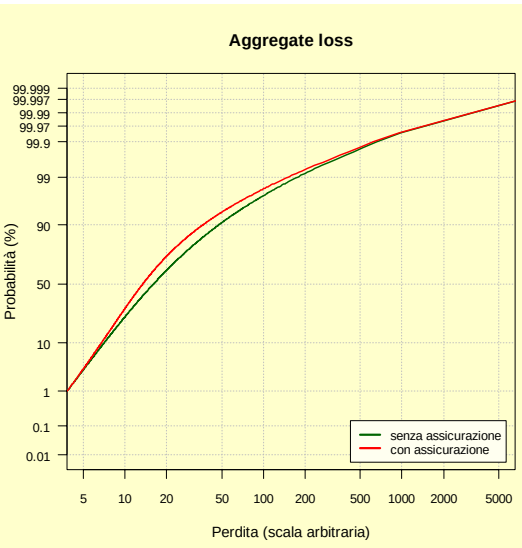
Example loss data analysis

- Fit distributions (maximum likelihood) to internal and external data
- Choose best fits (GOF tests)
- Compose internal/external distributions
- Use FFT to compute 1-year period aggregate, including insurance effects



Example loss data analysis

- Fit distributions (maximum likelihood) to internal and external data
- Choose best fits (GOF tests)
- Compose internal/external distributions
- Use FFT to compute 1-year period aggregate, including insurance effects



Scenario analysis

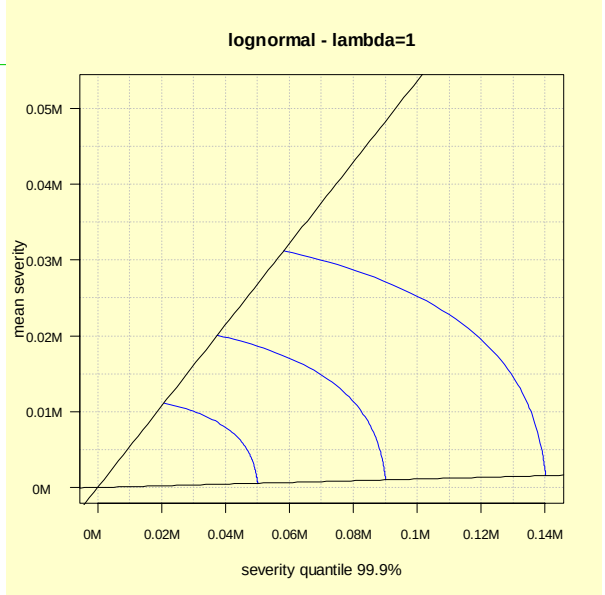
- Interview local management
- For each event type
  - Ask average frequency
  - Ask average loss
  - Ask “worst case” loss (99% quantile)
- Use ranges to guide these answers

Preparing scenario analysis

- Prepare the answer ranges:
  - fix frequency classes
  - determine three possible values for the 1-year aggregate unexpected loss (UL+EL= 99.9% quantile)
  - for each mean frequency, determine points with same UL
  - determine mean loss ranges
  - for each mean loss range, determine worst-case ranges by instersecting with iso-UL curves

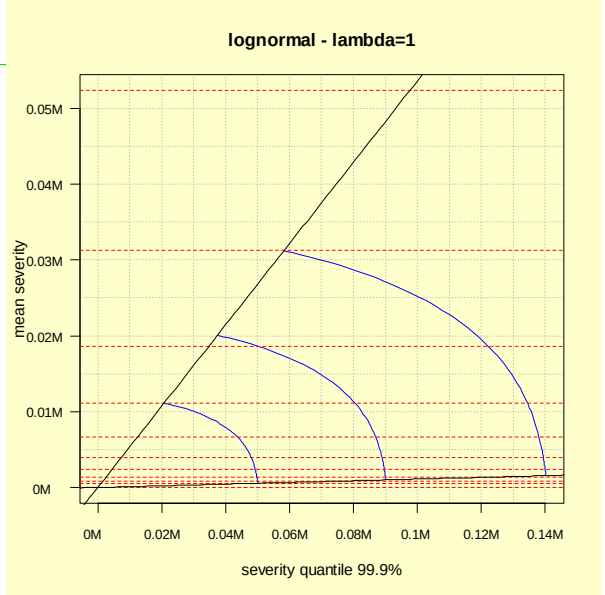
Example scenario construction

- fix frequency classes
- determine three possible values for the 1-year aggregate unexpected loss
- for each mean frequency, determine points with same UL
- determine mean loss ranges
- for each mean loss range, determine worst-case ranges by instersecting with iso-UL curves



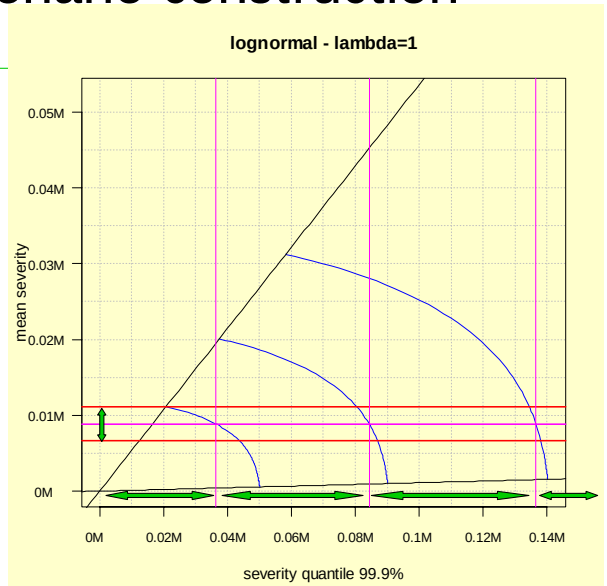
Example scenario construction

- fix frequency classes
- determine three possible values for the 1-year aggregate unexpected loss
- for each frequency class, determine curves with same UL
- determine mean loss ranges
- for each mean loss range, determine worst-case ranges by instersecting with iso-UL curves



Example scenario construction

- fix frequency classes
- determine three possible values for the 1-year aggregate unexpected loss
- for each frequency class, determine curves with same UL
- determine mean loss ranges
- for each mean loss range, determine worst-case ranges by instersecting with iso-UL curves

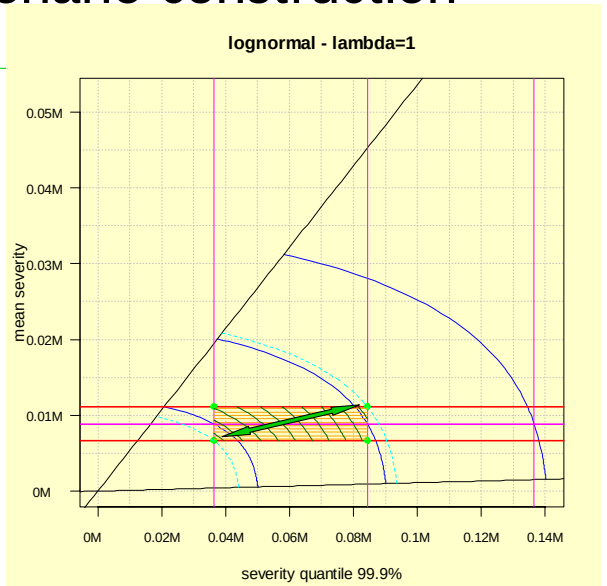


Example scenario construction

```
UL range:
44,005.20 93,386.79

Results:
EL : mean     8,868
     stddev   1,287
UL : mean    67,295
     stddev  11,909
VaR: mean    76,085
     stddev  12,436

rating (%):
   A     B     C     D
 3.4  93.2   3.4   0.0
```



Conclusions

- How did *R* perform?
  - methodological research and application prototyping: flexible tool, powerful language; library of tools developed
  - production environment: needs ad hoc GUI, has little support for compilation on mainframe architectures (e.g. HPUX)
  - memory/performance saturation limit hit when needing to handle very large amounts of data (>10^7 points)