Risk Measure Preserving Approximation of Univariate Monte Carlo Simulation Results with Insurance Applications

Philipp Arbenz, PhD, Actuary SAA

Joint work with William Guevara-Alarcón

This Talk is less Complicated than the Title Suggests

This presentation is about

- an **actuarial** problem,
- implemented with an approach typically seen in **computer science**,
- resulting in an algorithm falling into the area of **non-parametric statistics.**

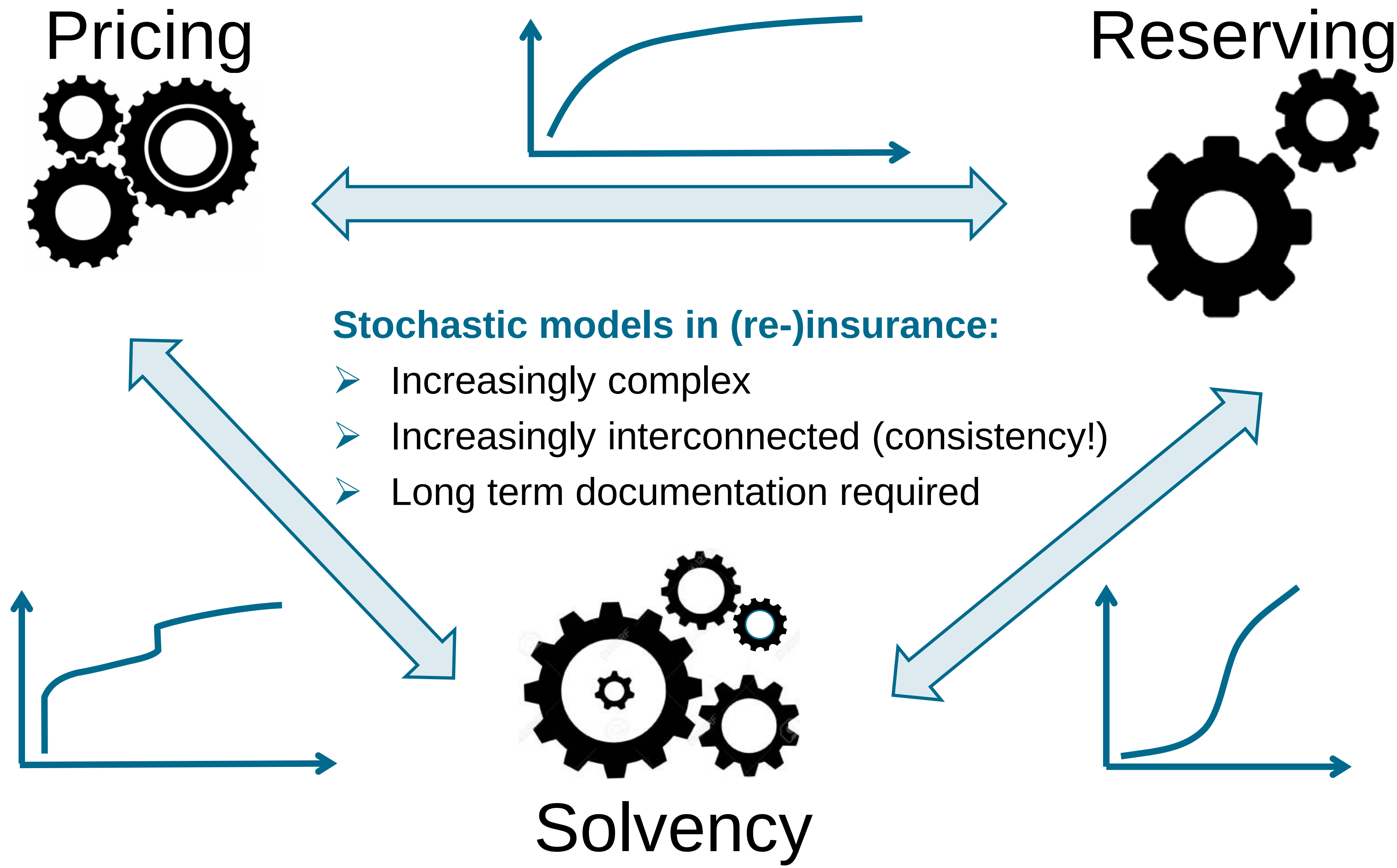
- Title for **actuaries** and **financial mathematicians**:
  "Piecewise linear approximation of loss simulations"

- Title for **computer scientists**:
  "A fast divide-and-conquer compression algorithm for random variables"

- Title for **statisticians**:
  "An efficient non-parametric estimator for empirical sample distributions"

Piecewise Linear Approximation of Univariate Sample Distributions

Motivation:
Increasing Complexity and Interconnection Between Stochastic Models



Stochastic models in (re-)insurance:

- Increasingly complex
- Increasingly interconnected (consistency!)
- Long term documentation required

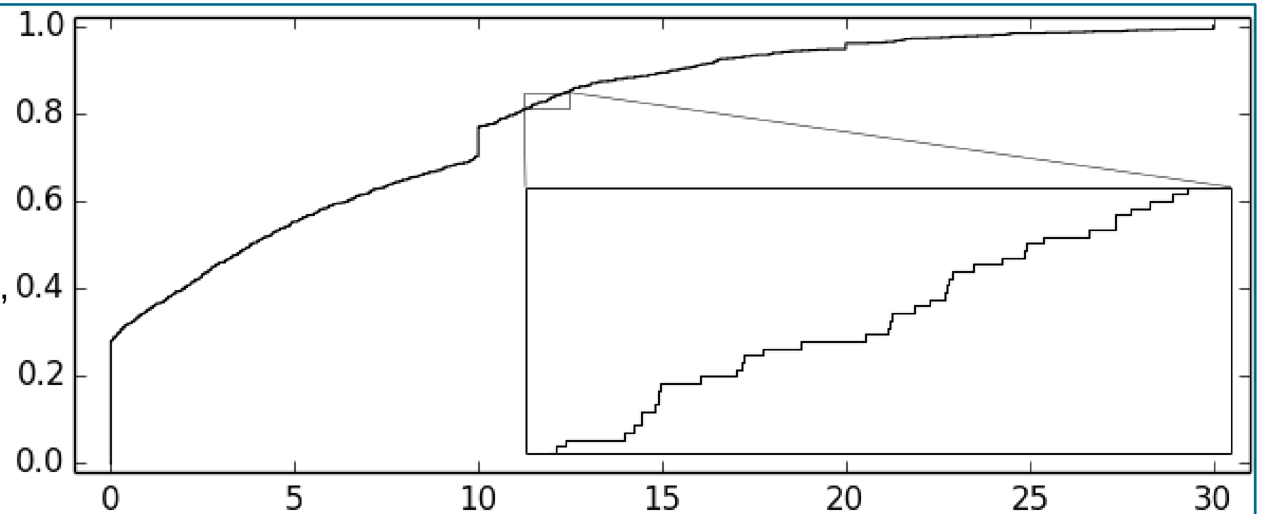
Main Idea

Problem

i. Stochastic model with univariate result distribution
ii. Result distribution not analytic → use Monte Carlo Simulation
iii. Need to **store** result or **transfer** result to other model or system

Example: Aggregate Loss of Excess-of-Loss reinsurance

- *XL Reinsurance Contract:* Limit = 10, Deductible = 12, Aggregate limit 30
- *Frequency:* Poisson($\lambda = 1$)
- *Severity:* Pareto($x_0 = 10, \alpha = 1.5$)



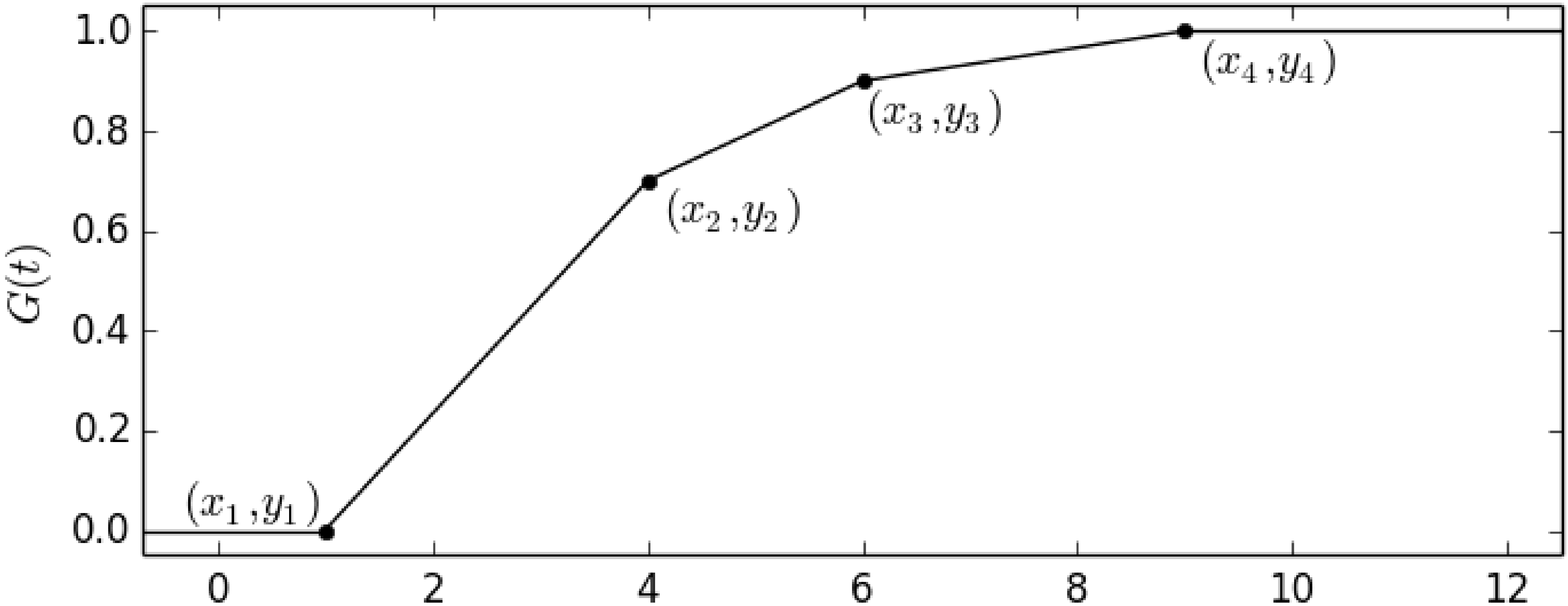
➢ Mean and standard deviation do not capture complex shape of distribution!

➢ Storing full sample requires too much memory! (typical: n = 1’000’000 = 64 Megabytes)

Solution ➢ Approximate sample distribution with piecewise linear distribution!

Piecewise Linear Approximation of Univariate Sample Distributions

Piecewise Linear Distribution (PWL)

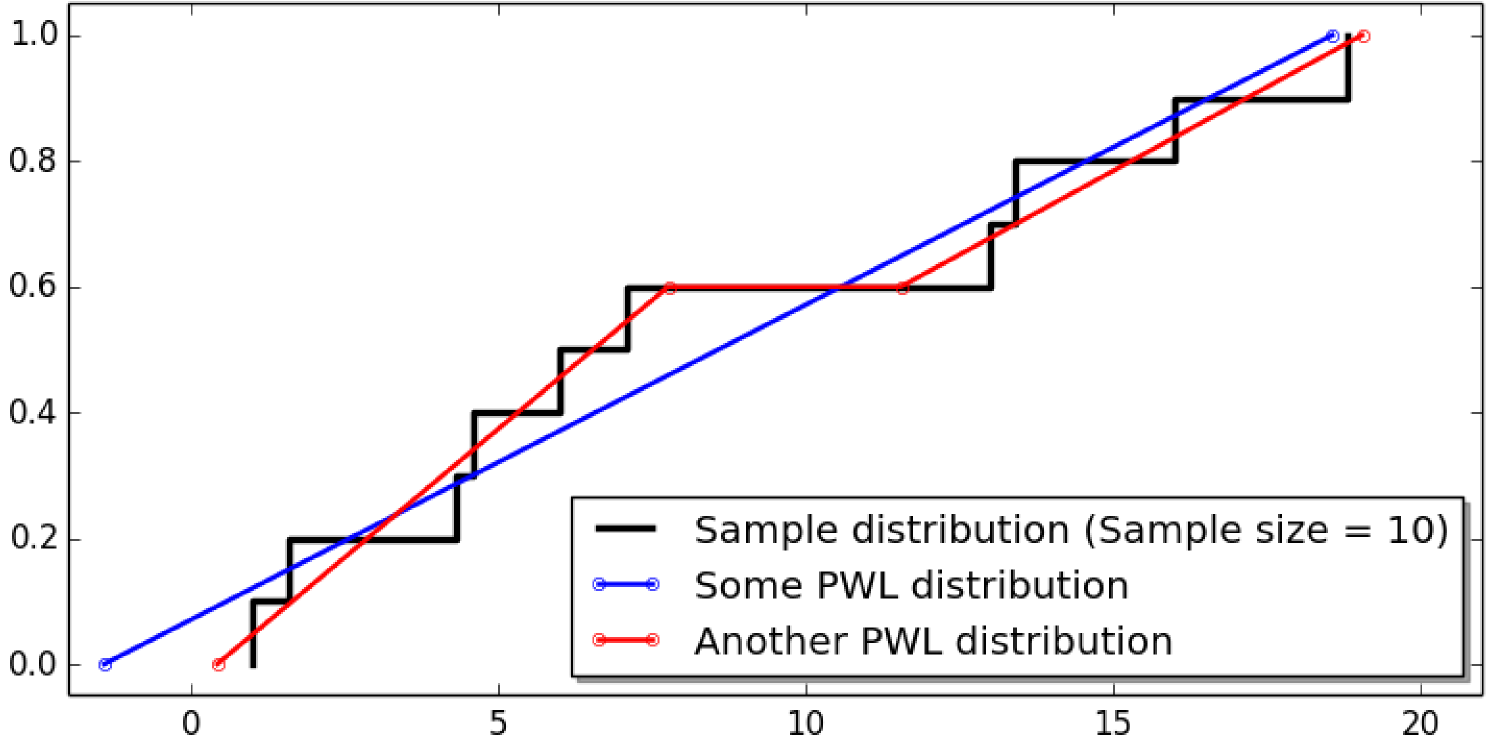


$$G \sim PWL\left(\boldsymbol{x} = (1, 4, 6, 9), \boldsymbol{y} = (0, 0.7, 0.9, 1)\right)$$

- Cumulative distribution function is piecewise linear, not the density!
- PWL can approximate any shape of the sample distribution (more details later…)

Which Piecewise Linear Distribution to Choose as Approximation?

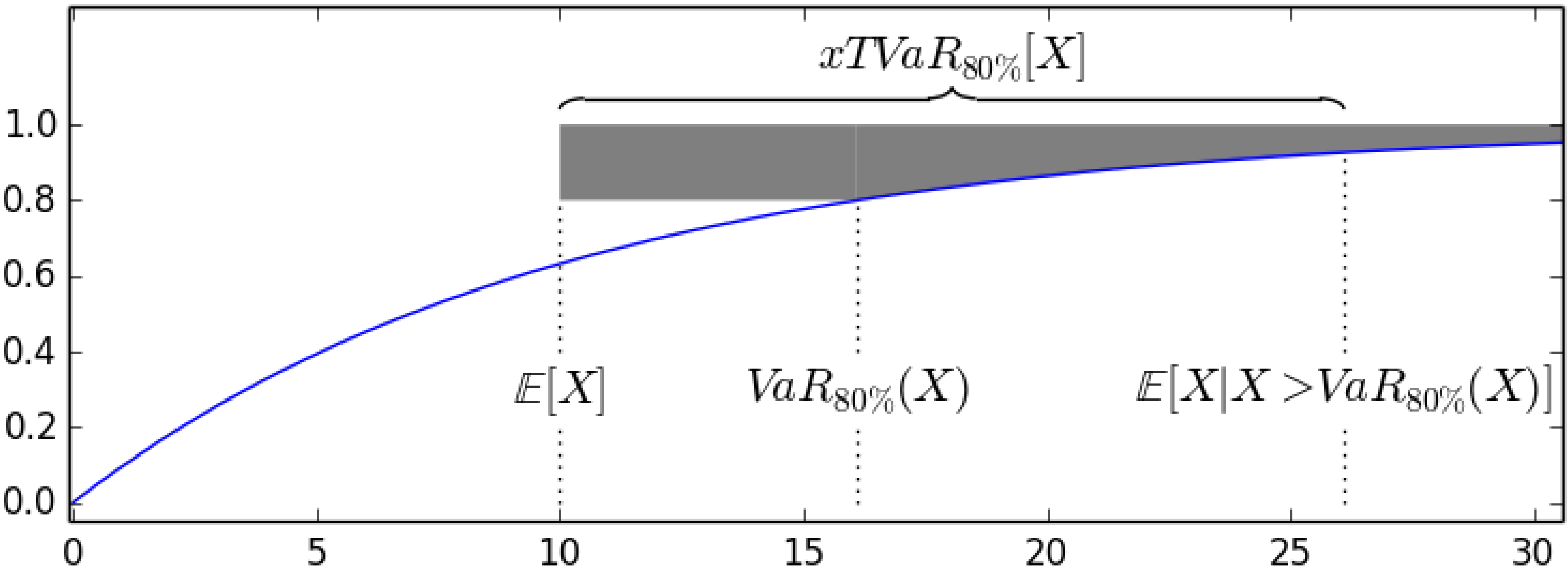
Question:	Which piecewise linear (PWL) distribution to choose as approximation of sample distribution?



Qualitative answer:	Appropriately reflect **riskyness** of sample distribution

Risk Measure xTVaR ("Excess Tail Value at Risk")

Definition: xTVaR	For a random variable $X: \Omega \mapsto \mathbb{R}$ and a quantile threshold $\alpha \in (0,1)$: $$xTVaR_\alpha(X) = \mathbb{E}[\, X - \mathbb{E}[X] \mid X > VaR_\alpha(X)\,]$$



Admissible Piecewise Linear Approximation

Definition: xTVaR	For a random variable $X\colon \Omega \mapsto \mathbb{R}$ and a quantile threshold $\alpha \in (0,1)$: $$xTVaR_\alpha(X) = \mathbb{E}[\, X - \mathbb{E}[X] \mid X > VaR_\alpha(X)\,]$$

Definition:

Admissible Approx-imation

Given

- $F_n(t) = \sum_{i=1}^{n} \mathbb{1}\{X_i \leq t\}$ the sample distribution of $\{X_1, X_2, \dots, X_n\}$,
- $G \sim PWL(\boldsymbol{x}, \boldsymbol{y})$ a piecewise linear distribution,
- $\epsilon > 0$ an accuracy parameter.

Then, G is an **admissible approximation** of F_n if

a) G is *mean invariant*: $\mathbb{E}[G] = \mathbb{E}[F_n]$

b) G is *risk preserving*: For every $0 < \alpha < 1$,

$$|xTVaR_\alpha(G) - xTVaR_\alpha(F_n)| \leq \epsilon \cdot xTVaR_\alpha(F_n)$$

Numerical Approach of Approximation Algorithm: Divide-and-Conquer

Algorithm: Given a univariate sample $\{X_1, X_2, \dots, X_n\}$,

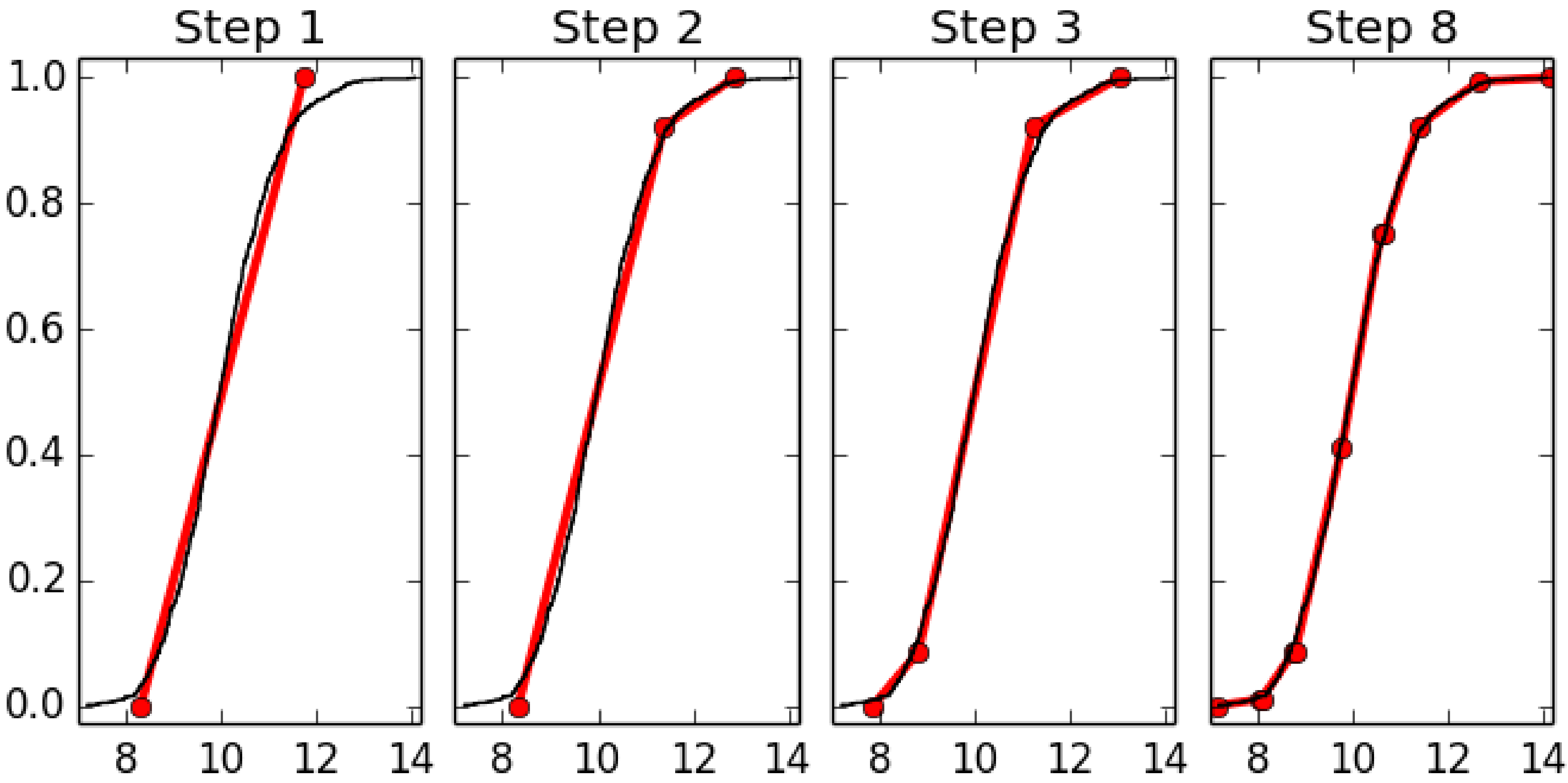
1) Set $\boldsymbol{\epsilon} > \mathbf{0}$ and initialize $\boldsymbol{y} = (0, 1)$
2) Set $\boldsymbol{x}$ of $G \sim PWL(\boldsymbol{x}, \boldsymbol{y})$ such that every linear segment is a least-square regression on F_n^{-1}
3) In case $G \sim PWL(\boldsymbol{x}, \boldsymbol{y})$ is **not** admissible:
   - Insert $\tilde{y}$ into the vector $\boldsymbol{y}$, where

$$\tilde{y} = \underset{0<\alpha<1}{\operatorname{argmax}} \left| xTVaR_\alpha(G) - xTVaR_\alpha(F_n) \right| \cdot (1 - \alpha)$$

   and go back to 2)

Numerical Approach of Approximation Algorithm: Divide-and-Conquer

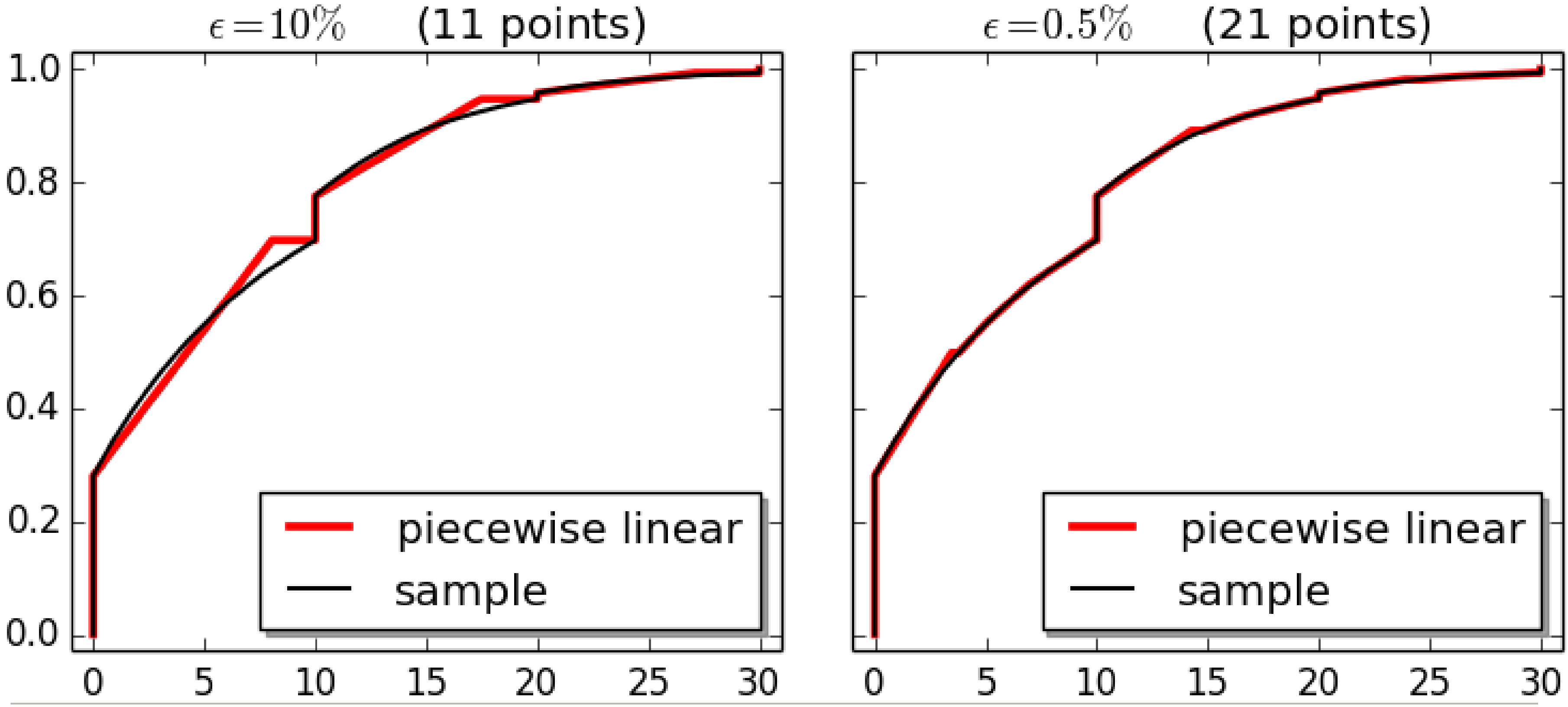
- Sample of Lognormal distribution with mean 10 and standard deviation 1



Piecewise Linear Approximation of Univariate Sample Distributions

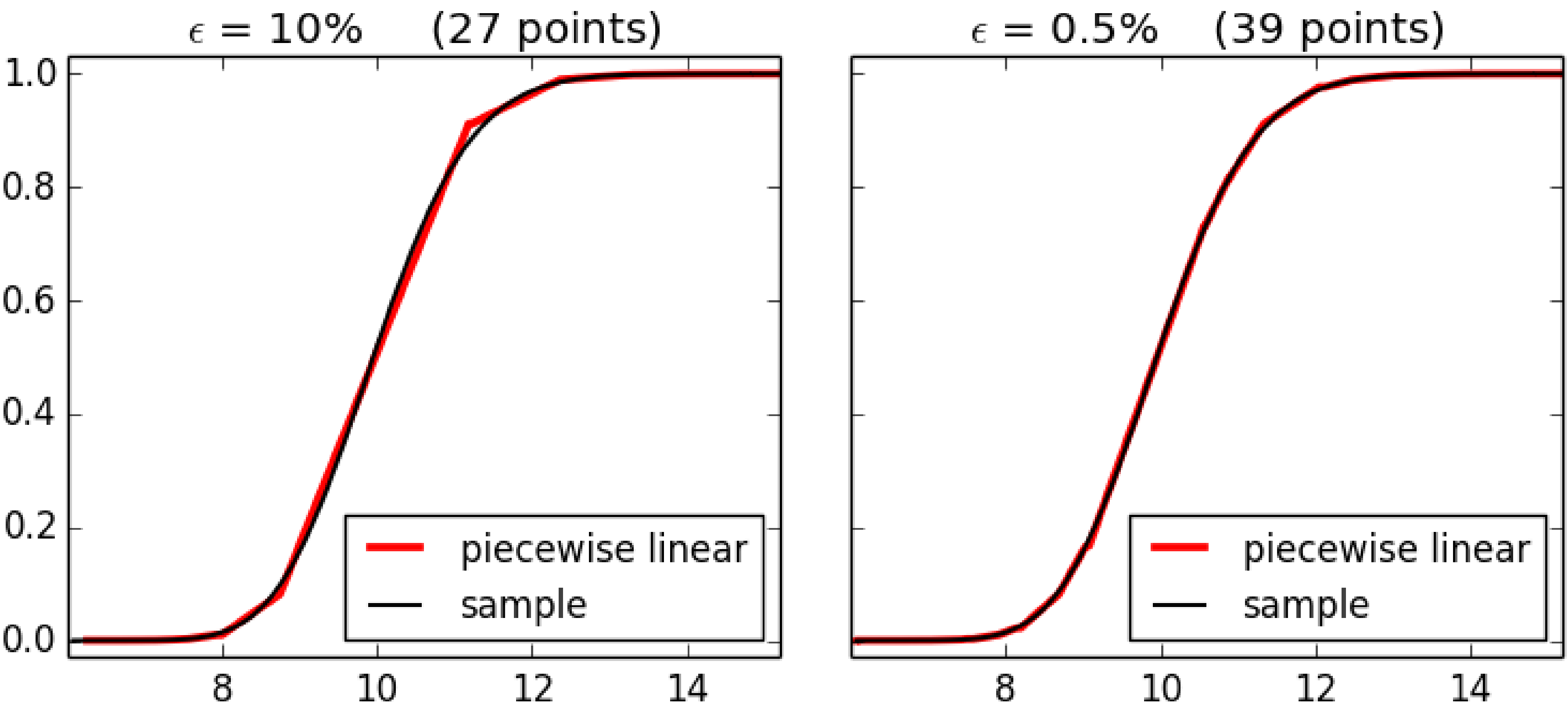
Excess-of-Loss Example: PWL Approximation of Sample Distribution

- **Excess-of-Loss reinsurance:** Limit 10, deductible 12, aggregate limit 30
- **Loss distribution:** Poisson($\lambda = 1$) frequency, Pareto($x_0 = 10, \alpha = 1.5$) severity



Lognormal Example: PWL Approximation of Sample Distribution

- Sample from $X \sim \text{Lognormal}$ with $\mathbb{E}[X] = 10$ and $std(X) = 1$, $n = 1'000'000$



Numerical Compexity and Run Time

Overall **numerical complexity** with sample size $\boldsymbol{n}$ and precision $\boldsymbol{\epsilon}$:

$$\boldsymbol{O(n \log(n/\epsilon))}$$

Example: Excess-of-Loss Sample	Run time in milliseconds		Number of interpolation points	
	$\boldsymbol{n = 100'000}$	$\boldsymbol{n = 1'000'000}$	$\boldsymbol{n = 100'000}$	$\boldsymbol{n = 1'000'000}$
$\boldsymbol{\epsilon = 0.1}$	14	155	11	11
$\boldsymbol{\epsilon = 0.01}$	19	171	17	17
$\boldsymbol{\epsilon = 0.001}$	25	192	27	28

Number of interpolation points: $O(\log(1/\epsilon))$

Practical Experience with PWL Approximation of the Reinsurance SCOR

SCOR (and Converium) experience in last 10 years:

- Piecewise linear approximation is used in treaty pricing system since 2005
- Algorithm has same purpose (PWL approximation), but with a different mathematical foundation (Invented by C. Hummel)

➢ **Realized memory efficiency gains:** 6GB instead of 150TB over last 10 years

150 TB vs 6 GB

(compression ratio 25'000)

➢ **Realized infrastructure efficiency gains:** Simulations done on the actuaries computers and compressed distributions transferred to server later. Scalable!

Free and Open Source Implementations Available

python

```
from compressor import PWLcompressor
sample    = [1, 1.6, 4.3, 4.6, 6, 7.1, 13, 13.4, 16, 18.8]
pwlapprox = PWLcompressor(sample, Accuracy = 0.01)
```

C++

```
#include <vector>
#include "compressor.h"
int main() {
  std::vector<double> sample ={1,1.6,4.3,4.6,6,7.1,13,13.4,16,18.8};
  PWLCompressor pwlapprox = PWLCompressor(sample, 0.01);  }
```

R

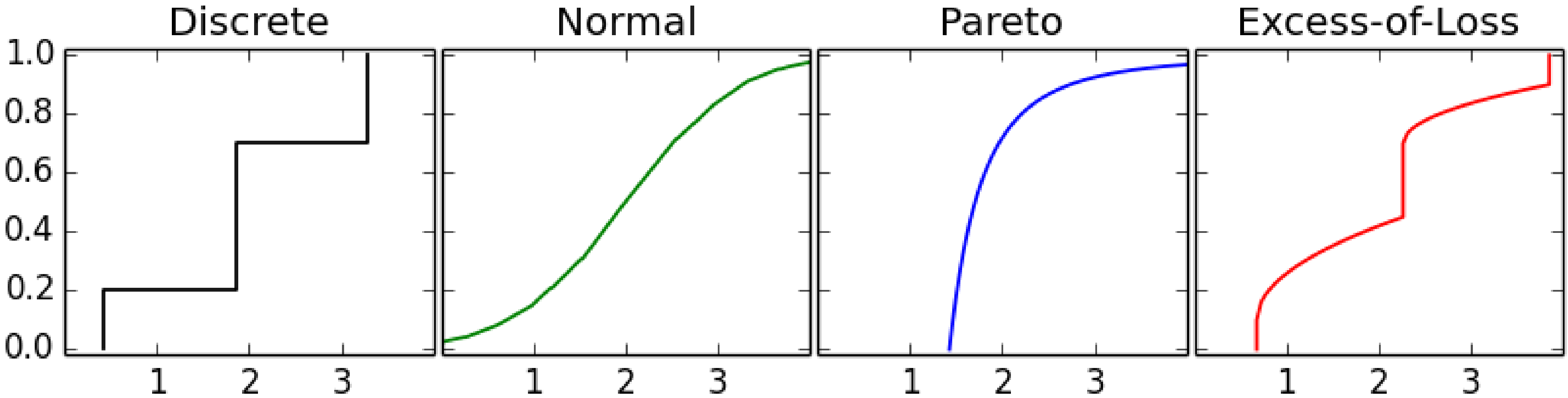
```
library(Rcpp)
dyn.load( paste("libCompressor",sep="",.Platform$dynlib.ext) )
PWLCompressor <- function(sample,acc).Call("Compressor",sample,acc)
sample        <- c(1,1.6,4.3,4.6,6,7.1,13,13.4,16,18.8)
pwlapprox     <- PWLCompressor(sample, 0.01)
```

- Code available on authors homepage, under *MIT license*
- Implementations come with a complete set of test cases

Piecewise Linear Approximation of Univariate Sample Distributions

Desirable Properties of Piecewise Linear Distributions: Shape Preserving, Efficient

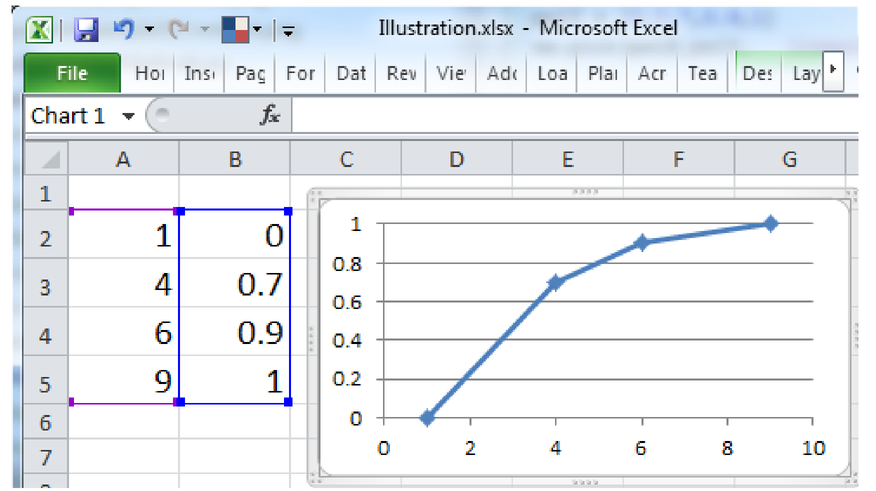
- **Shape:** Can represent any sample distribution shape with arbitrary precision
  - As powerful as approximation with higher degree polynomials



- **Statistics:** Key statistics can be calculated analytically and efficiently
  - $\mathbb{P}[X \leq x],\ VaR_{\alpha}(X)$
  - $\mathbb{E}[X], \mathbb{E}[X^k],\ std(X)$
  - $TVaR_{\alpha}(X),\ xTVaR_{\alpha}(X)$

Desirable Properties of Piecewise Linear Distributions: Intuitive, IT System Independent, Parsimonious

- **PWL parameters (x, y):**
  - Efficient to store and transfer
  - Long term storage is no problem (minimum record keeping periods!)
  - Intuitive: Easy to visualize and to explain to non-technical stakeholders



- **Parsimonious:** Number of interpolation points depends on complexity of sample distribution shape
  - Constant distribution: 2 points
  - Complicated distributions with $\epsilon = 0.001$: usually less than 100 points

Comparison to Alternative Approaches

	Sample distribution shape preserving	Memory & bandwidth efficient	IT system independent
Store only sample statistics	✗	✓	✓
Store full sample	✓	✗	✓
Store software and random seeds	✓	✗	✗
Piecewise linear approximation	✓	✓	✓

Piecewise Linear Approximation of Univariate Sample Distributions

Conclusion

Current actuarial environment :

- **Number** and **usage** of stochastic models increases
- Requirements for **consistent** stochastic models increases
- **Interactions** between stochastic models increases
- Need to **store** model results for reporting and audit purposes increases

Piecewise linear approximation:

- Algorithm allows to approximate univariate sample distributions
- Preserves sample distribution shape in a risk invariant manner
- Open source implementation available
- Algorithm is efficient in terms of run time, memory, and bandwidth requirements

Thank you for your attention!