Risk Theory Calculations Using R and **actuar**

Vincent Goulet, Ph.D.

École d'actuariat, Université Laval
Québec, Canada

Actuarial Risk Modeling Process

1. Model costs process at the individual level
   $\Rightarrow$ Modeling of loss distributions
2. Aggregate risks at the collective level
   $\Rightarrow$ Risk theory
3. Determine revenue streams
   $\Rightarrow$ Ratemaking (including Credibility Theory)
4. Evaluate solvability of insurance portfolio
   $\Rightarrow$ Ruin theory

Collective Risk Model

- Let

$$\begin{aligned} S &: \text{aggregate claim amount} \\ N &: \text{number of claims (frequency)} \\ C_j &: \text{amount of claim } j \text{ (severity)} \end{aligned}$$

- We have the random sum

$$S = C_1 + \cdots + C_N$$

- We want to find

$$\begin{aligned} F_S(x) &= \Pr[S \leq x] \\ &= \sum_{n=0}^{\infty} \Pr[S \leq x | N = n] \Pr[N = n] \\ &= \sum_{n=0}^{\infty} F_C^{*n}(x) \Pr[N = n] \end{aligned}$$

Collective Risk Model

- Let

$$\begin{aligned} S &: \text{aggregate claim amount} \\ N &: \text{number of claims (frequency)} \\ C_j &: \text{amount of claim } j \text{ (severity)} \end{aligned}$$

- We have the random sum

$$S = C_1 + \cdots + C_N$$

- We want to find

$$\begin{aligned} F_S(x) &= \Pr[S \leq x] \\ &= \sum_{n=0}^{\infty} \Pr[S \leq x | N = n] \Pr[N = n] \\ &= \sum_{n=0}^{\infty} F_C^{*n}(x) \Pr[N = n] \end{aligned}$$

Collective Risk Model

- Let

$$
\begin{aligned}
S &: \text{aggregate claim amount} \\
N &: \text{number of claims (frequency)} \\
C_j &: \text{amount of claim } j \text{ (severity)}
\end{aligned}
$$

- We have the random sum

$$S = C_1 + \cdots + C_N$$

- We want to find

$$
\begin{aligned}
F_S(x) &= \Pr[S \leq x] \\
&= \sum_{n=0}^{\infty} \Pr[S \leq x | N = n] \Pr[N = n] \\
&= \sum_{n=0}^{\infty} F_C^{*n}(x) \Pr[N = n]
\end{aligned}
$$

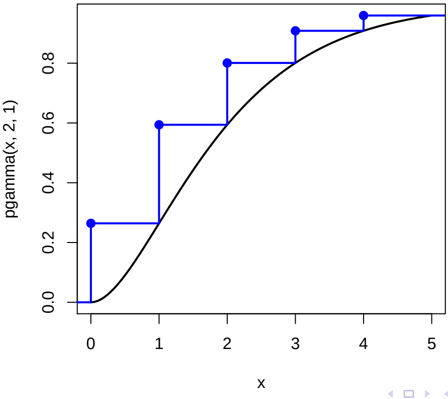
Aggregate Claim Amount Distribution

- Function `aggregateDist()` supports five methods
- Main one is the recursive method (Panjer algorithm):

$$f_S(x) = \frac{1}{1 - af_C(0)} \Bigg[(p_1 - (a+b)p_0) f_C(x) + \sum_{y=1}^{\min(x,m)} (a + by/x) f_C(y) f_S(x-y) \Bigg]$$

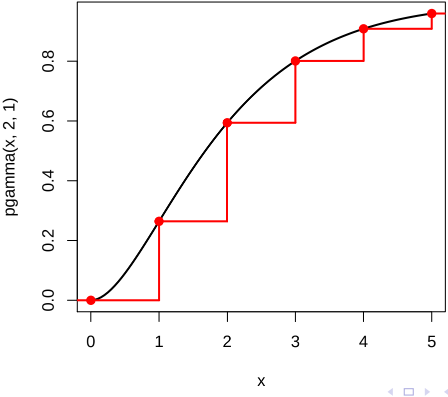
Discretization of Continuous Distributions

```
> discretize(pgamma(x, 2, 1), from = 0, to = 5,
+                method = "upper")
```



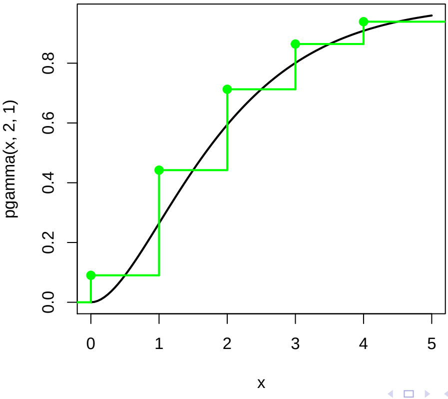
Discretization of Continuous Distributions

```
> discretize(pgamma(x, 2, 1), from = 0, to = 5,
+            method = "lower")
```



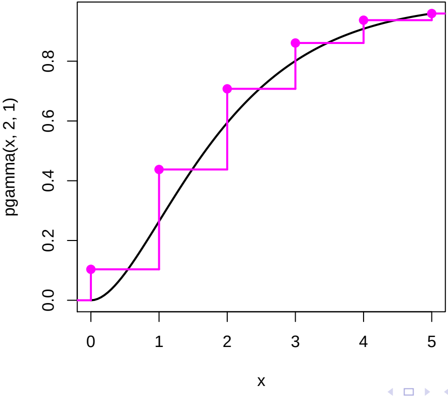
Discretization of Continuous Distributions

```
> discretize(pgamma(x, 2, 1), from = 0, to = 5,
+            method = "rounding")
```



Discretization of Continuous Distributions

```
> discretize(pgamma(x, 2, 1), from = 0, to = 5,
+            method = "unbiased",
+            lev = levgamma(x, 2, 1))
```



Example

Assume

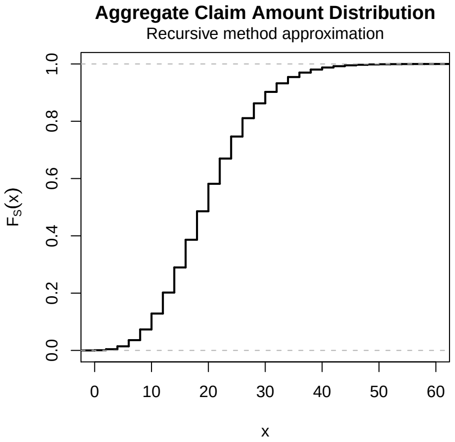
$$N \sim \text{Poisson}(10)$$
$$C \sim \text{Gamma}(2, 1)$$

```
> fx <- discretize(pgamma(x, 2, 1), from = 0,
+                  to = 22, step = 2,
+                  method = "unbiased",
+                  lev = levgamma(x, 2, 1))

> Fs <- aggregateDist("recursive",
+                   model.freq = "poisson",
+                   model.sev = fx,
+                   lambda = 10, x.scale = 2)
```

Example (continued)

```
> plot(Fs)
```



Example (continued)

```
> summary(Fs)

Aggregate Claim Amount Empirical CDF:
    Min.  1st Qu.   Median     Mean  3rd Qu.
 0.00000 12.00000 18.00000 19.99996 24.00000
    Max.
74.00000

> knots(Fs)

 [1]  0  2  4  6  8 10 12 14 16 18 20 22 24
[14] 26 28 30 32 34 36 38 40 42 44 46 48 50
[27] 52 54 56 58 60 62 64 66 68 70 72 74

> Fs(c(10, 15, 20, 70))

[1] 0.1287553 0.2896586 0.5817149 0.9999979
```

Example (continued)

```
> mean(Fs)

[1] 19.99996

> VaR(Fs)

90% 95% 99%
 28  32  40

> CTE(Fs)

     90%      95%      99%
34.24647 37.76648 45.09963
```

Long Term Risk Analysis

- Study evolution of the surplus of the insurance company over many periods of time
- Quantity of interest: probability that surplus becomes negative
- Technical ruin of the insurance company ensues
- Equivalent idea in other fields

Continuous Time Ruin Model

- Let

$$U(t): \text{surplus at time } t$$
$$c(t): \text{premiums collected through time } t$$
$$S(t): \text{aggregate claims paid through time } t$$

- If u is the initial surplus at time $t = 0$, then we have

$$U(t) = u + c(t) - S(t)$$

- We want

$$\psi(u) = \Pr[U(t) < 0 \text{ for some } t \geq 0]$$

Continuous Time Ruin Model

- Let

$$\begin{aligned} U(t) &: \text{surplus at time } t \\ c(t) &: \text{premiums collected through time } t \\ S(t) &: \text{aggregate claims paid through time } t \end{aligned}$$

- If u is the initial surplus at time $t = 0$, then we have

$$U(t) = u + c(t) - S(t)$$

- We want

$$\psi(u) = \Pr[U(t) < 0 \text{ for some } t \geq 0]$$

Continuous Time Ruin Model

- Let

$$U(t): \text{surplus at time } t$$
$$c(t): \text{premiums collected through time } t$$
$$S(t): \text{aggregate claims paid through time } t$$

- If u is the initial surplus at time $t = 0$, then we have

$$U(t) = u + c(t) - S(t)$$

- We want

$$\psi(u) = \Pr[U(t) < 0 \text{ for some } t \geq 0]$$

Ruin Probabilities

- If $W_j \sim$ Exponential(λ) and $C_j \sim$ Exponential(β), then

$$\psi(u) = \frac{\lambda}{c\beta} e^{-(\beta - \lambda/c)u}$$

- Most common distributions for claim amounts and waiting times:
  - mixtures of exponentials
  - mixtures of Erlang
  - phase-type
- In most cases `ruin()` computes probabilities with `pphtype()`

Example

Mixture of two exponentials for claims, exponential interarrival times

```
> psi <- ruin(claims = "exponential",
+             par.claims = list(rate = c(3, 7),
+                               weights = 0.5),
+             wait = "exponential",
+             par.wait = list(rate = 3),
+             premium.rate = 1)
```

```
> u <- 0:10
> psi(u)
```

```
 [1] 7.142857e-01 2.523310e-01 9.280151e-02
 [4] 3.413970e-02 1.255930e-02 4.620307e-03
 [7] 1.699716e-03 6.252905e-04 2.300315e-04
[10] 8.462387e-05 3.113138e-05
```

Example

Mixture of two exponentials for claims, exponential interarrival times

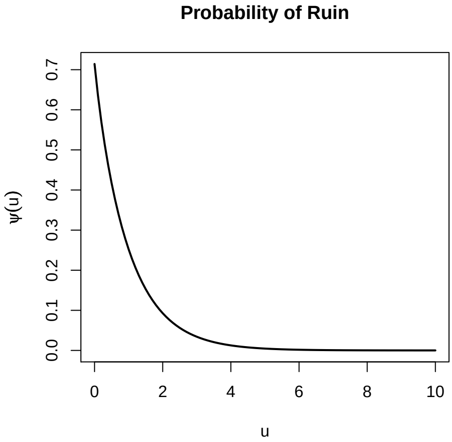
```
> psi <- ruin(claims = "exponential",
+             par.claims = list(rate = c(3, 7),
+                                weights = 0.5),
+             wait = "exponential",
+             par.wait = list(rate = 3),
+             premium.rate = 1)

> u <- 0:10
> psi(u)

 [1] 7.142857e-01 2.523310e-01 9.280151e-02
 [4] 3.413970e-02 1.255930e-02 4.620307e-03
 [7] 1.699716e-03 6.252905e-04 2.300315e-04
[10] 8.462387e-05 3.113138e-05
```

Example (continued)

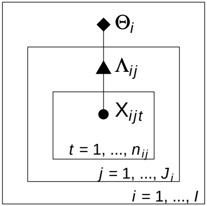
```
> plot(psi, from = 0, to = 10)
```

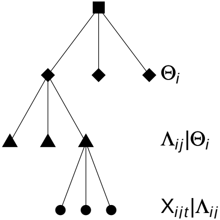


Simulation of Compound Hierarchical Models

- You want to simulate data from this model?

$$
\begin{aligned}
X_{ijt}|\Lambda_{ij}, \Theta_i &\sim \text{Poisson}(\Lambda_{ij}), & t &= 1, \ldots, n_{ij} \\
\Lambda_{ij}|\Theta_i &\sim \text{Gamma}(3, \Theta_i), & j &= 1, \ldots, J_i \\
\Theta_i &\sim \text{Gamma}(2, 2), & i &= 1, \ldots, I,
\end{aligned}
$$





- Or from this one?

$$S_{ijt} = C_{ijt1} + \cdots + C_{ijtN_{ijt}},$$

with

$$\begin{aligned} N_{ijt}|\Lambda_{ij}, \Phi_i &\sim \text{Poisson}(w_{ijt}\Lambda_{ij}) \\ \Lambda_{ij}|\Phi_i &\sim \text{Gamma}(\Phi_i, 1) \\ \Phi_i &\sim \text{Exponential}(2) \end{aligned}$$

$$\begin{aligned} C_{ijtu}|\Theta_{ij}, \Psi_i &\sim \text{Lognormal}(\Theta_{ij}, 1) \\ \Theta_{ij}|\Psi_i &\sim N(\Psi_i, 1) \\ \Psi_i &\sim N(2, 0.1) \end{aligned}$$

- Using only R syntax (i.e. without reverting to BUGS)?

- Then read this fine paper:

  Goulet, V., Pouliot, L.-P. (2008), *Simulation of Compound Hierarchical Models in R*, North American Actuarial Journal, **12**, 401–412.

More Information

- Project's web site

  `http://www.actuar-project.org`

- Package vignettes

`actuar`	`Introduction to actuar`
`coverage`	`Complete formulas used by coverage`
`credibility`	`Risk theory features`
`lossdist`	`Loss distributions modeling features`
`risk`	`Risk theory features`

- Demo files