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
   ⇒ Modeling of loss distributions
2. Aggregate risks at the collective level
   ⇒ Risk theory
3. Determine revenue streams
   ⇒ Ratemaking (including Credibility Theory)
4. Evaluate solvability of insurance portfolio
   ⇒ Ruin theory
Collective Risk Model

- Let
  
  \( S : \text{aggregate claim amount} \)
  
  \( N : \text{number of claims (frequency)} \)
  
  \( C_j : \text{amount of claim } j \text{ (severity)} \)

- We have the random sum
  
  \[ S = C_1 + \cdots + C_N \]

- We want to find
  
  \[ F_S(x) = \Pr[S \leq x] \]
  
  \[ = \sum_{n=0}^{\infty} \Pr[S \leq x|N = n] \Pr[N = n] \]
  
  \[ = \sum_{n=0}^{\infty} F_C^n(x) \Pr[N = n] \]
Collective Risk Model

- Let
  
  \( S : \) aggregate claim amount
  
  \( N : \) number of claims (frequency)
  
  \( C_j : \) amount of claim \( j \) (severity)

- We have the random sum
  
  \( S = C_1 + \cdots + C_N \)

- We want to find
  
  \[ F_S(x) = \Pr[S \leq x] = \sum_{n=0}^{\infty} \Pr[S \leq x|N = n] \Pr[N = n] = \sum_{n=0}^{\infty} F_C^n(x) \Pr[N = n] \]
Let

$S : \text{aggregate claim amount}$

$N : \text{number of claims (frequency)}$

$C_j : \text{amount of claim } j \text{ (severity)}$

We have the random sum

$S = C_1 + \cdots + C_N$

We want to find

$F_S(x) = \Pr[S \leq x]$

$= \sum_{n=0}^{\infty} \Pr[S \leq x | N = n] \Pr[N = n]$

$= \sum_{n=0}^{\infty} F_{C_n}^*(x) \Pr[N = n]$
Function `aggregateDist()` supports five methods
- Main one is the recursive method (Panjer algorithm):

\[
f_S(x) = \frac{1}{1 - af_C(0)} \left[ (p_1 - (a + b)p_0)f_C(x) \right. \\
\left. + \sum_{y=1}^{\min(x,m)} (a + by/x)f_C(y)f_S(x - y) \right]
\]
Discretization of Continuous Distributions

```r
> discretize(pgamma(x, 2, 1), from = 0, to = 5, 
  +     method = "upper")
```
Discretization of Continuous Distributions

\[ > \text{discretize}(\text{pgamma}(x, 2, 1), \text{from} = 0, \text{to} = 5, + \text{method} = \"lower\") \]
> discretize(pgamma(x, 2, 1), from = 0, to = 5, + method = "rounding")
> 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)
> plot(Fs)

Aggregate Claim Amount Distribution
Recursive method approximation

\[ F_S(x) \]
> summary(Fs)

Aggregate Claim Amount Empirical CDF:

<table>
<thead>
<tr>
<th></th>
<th>Min.</th>
<th>1st Qu.</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Qu.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.00000</td>
<td>12.00000</td>
<td>18.00000</td>
<td>19.99996</td>
<td>24.00000</td>
</tr>
<tr>
<td>Max.</td>
<td>74.00000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

- 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

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

- 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 $W_j \sim \text{Exponential}(\lambda)$ and $C_j \sim \text{Exponential}(\beta)$, 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, \text{ruin()} computes probabilities with \text{pphtype()}
Example

Mixture of two exponentials for claims, exponential interarrival times

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

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

```
[1] 7.142857e-01 2.523310e-01 9.280151e-02
[10] 8.462387e-05 3.113138e-05
```
Example

Mixture of two exponentials for claims, exponential interarrival times

```r
> 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
[10]  8.462387e-05  3.113138e-05
```
> plot(psi, from = 0, to = 10)
You want to simulate data from this model?

\[ X_{ijt} | \Lambda_{ij}, \Theta_i \sim \text{Poisson}(\Lambda_{ij}), \quad t = 1, \ldots, n_{ij} \]
\[ \Lambda_{ij} | \Theta_i \sim \text{Gamma}(3, \Theta_i), \quad j = 1, \ldots, J_i \]
\[ \Theta_i \sim \text{Gamma}(2, 2), \quad i = 1, \ldots, I, \]
Or from this one?

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

with

\[ N_{ijt}|\Lambda_{ij}, \Phi_i \sim \text{Poisson}(\lambda_{ijjt}) \]
\[ \Lambda_{ij}|\Phi_i \sim \text{Gamma}(\Phi_i, 1) \]
\[ \Phi_i \sim \text{Exponential}(2) \]

\[ C_{ijtu}|\Theta_{ij}, \Psi_i \sim \text{Lognormal}(\Theta_{ij}, 1) \]
\[ \Theta_{ij}|\Psi_i \sim N(\Psi_i, 1) \]
\[ \Psi_i \sim N(2, 0.1) \]
Using only R syntax (i.e. without reverting to BUGS)?
Then read this fine paper:

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