MATH 4281 Risk Theory–Ruin and Credibility

Intro to the course and start of Module 1

January 12, 2021

Introduction to this course

The story so far...

For those who took Math 4280 you essentially studied the following problem:

How to compute $\rho[L]$? = ES[L]?

Where:

- ρ is some (potentially) coherent risk measure e.g. ES/TCE/CVaR, etc...
- L is some random variable representing a loss.

Some Questions

Q1: What do you do when L is equal to a sum of smaller RVs?

Q2: How do you introduce **time** to this model?

Q3: How do I estimate the parameters of the model for L...if I don't have a nice heterogeneous sample?

Some Questions

Q1: What do you do when L is equal to a sum of smaller RVs?
$\Rightarrow$ Module 1: Aggregate Loss Models

Q2: How do you introduce **time** to this model?
$\Rightarrow$ Module 2: Ruin Theory

Q3: How do I estimate the parameters of the model for L...if I don't have a nice heterogeneous sample?
$\Rightarrow$ Module 3: Credibility

Start of Module 1: Intro to Aggregate Loss Models

Models for aggregate losses

A portfolio of insurance contracts or an insurance contract will potentially experience a sequence of losses:

$$X_1, X_2, X_3, \ldots$$

We are interested in the aggregate sum S of these losses over a certain period of time.

Assumptions going forward

- How do losses relate to each other?
  - $\hookrightarrow$ Assume independent X_i's.
- When do these losses occur?
  - $\hookrightarrow$ Assume no time value of money i.e. short term models
- How many losses will occur?
  - $\hookrightarrow$ if deterministic $(n) \longrightarrow$ individual risk model
  - $\hookrightarrow$ if random $(N) \longrightarrow$ collective risk model

Definition: The Individual Risk Model

The Individual Risk Model[1]

$$S = X_1 + \cdots + X_n = \sum_{i=1}^{n} X_i,$$

- The random variables X_i, $i = 1, 2, ..., n$, are assumed to be independent
- BUT they are not assumed to be identically distributed.
- Typically the X_i's have mass at 0 (representing no loss/payment).

[1] Refer to Def. 9.2 in the loss model textbook

Definition: The Collective Risk Model

In the Collective Risk Model, aggregate losses become

$$S = X_1 + \ldots + X_N = \sum_{i=1}^{N} X_i.$$

This is a random sum. We make the following assumptions:

- N is the number of claims
- X_i is the amount of the ith claim
- the X_i's are iid with
  - CDF $F(x)$
  - P(D/M)F $f(x)$
  - Moments exist and are finite!
- the X_i's and N are mutually independent

How to compute S?

We will study methods to get probabilities about S:

1. If possible we will get the true distribution of S via:
    - Convolutions
    - Method of generating functions

2. Otherwise we will approximate with the help of the moments of S

Examples of The IRM vs CRM

A group life insurance contract where each employee has a different age, gender, and death benefit[2].

A: IRM

why: · we know "n"

· X_i's not identical

[2] ex 9.2 of the loss models book

Examples of The IRM vs CRM

A reinsurance contract that pays when the annual total medical malpractice costs at a certain hospital exceeds a given amount.

A: CRM

why: N - random # of claims

X_i's are similar → i.i.d justified!

Examples of The IRM vs CRM

A dental policy on an individual pays for at most two checkups per year per family member. A single contract covers any size family at the same price.

A CRM

why? N random → # of family members claiming + # of checkups

An example for the collective Risk model

An insurable event has a 10% probability of occurring and when it occurs results in a loss of \$5,000. Market research has indicated that consumers will pay at most \$550 to purchase insurance against this event. How many policies must a company sell in order to have a 95% chance of making money (ignoring expenses)?[3]

n = # of policies

C = # of claims @ 5,000 \$

$0.95 \leq P(5000C < 550n)$

[3] ex 9.1 of the loss models book

An example for the collective Risk model

$$0.95 \le P(5000C \le 550n)$$

$$= P(C < 0.11n)$$

$$= P\Big(\underbrace{\frac{C - np}{\sqrt{np(1-p)}}}_{\text{std normal}} \le \frac{1}{\sqrt{n}}(0.0\dot{3})\Big)$$

$C \sim Bin(n, 0.1)$

as $n \to \infty$

$C \sim N(np, np(1-p))$

$$\Phi^{-1}(0.95) \le \frac{1}{\sqrt{n}}\, 0.0\dot{3}$$

$$\Rightarrow \quad 2{,}400 \gtrsim n$$

An example for the collective Risk model

Generating Functions and Convolutions

Probability Generating functions

Definition (PGF)

Given a discrete RV X. We define the Probability Generating Function(PGF) $p_X(t)$ as:

$$p_X\left(t\right) = E\left[t^X\right]$$

i.e:

$$p_X\left(t\right) = Pr[X = x_0]t^{x_0} + \Pr[X = x_1]t^{x_1} + \Pr[X = x_2]t^{x_2} + \dots$$

Properties

- There is a 1-1 relation between a distribution and its PGF.
- If X is an integer-valued random variable, then the PGF is

$$p_X(t) = E\left[t^X\right] = \sum_{n=0}^{\infty} \Pr[X = n] t^n,$$

which is in fact the Taylor series of $p_X(t)$:

$$\Pr[X = n] = \frac{\frac{d^n}{dt^n} p_X(t)\,|_{t=0}}{n!}$$

Recover Probability

.

- If $X_i, i = 1, \cdots, n$ are independent, then:

$$p_{X_1+\cdots+X_n}(t) = p_{X_1}(t)\, p_{X_2}(t) \cdots p_{X_n}(t)$$

Moment Generating functions

Definition (MGF)

For a continous random variable X we define the Moment Generating Function (MGF) as:

$$M_X(t) = E\left(e^{tX}\right)$$

Properties

- There is a 1-1 relation between a distribution and its MGF.
- Taylor Expansion:

$$M_X(t) = 1 + E[X]t + E[X^2]\frac{t^2}{2} + E[X^3]\frac{t^3}{6} + \ldots + E[X^k]\frac{t^k}{k!} + \ldots$$

Recover moments

and thus

$$E[X^k] = \left.\frac{d^k}{dt^k} m_X(t)\right|_{t=0}$$

- If $X_i, i = 1, \cdots, n$ are independent then:

(★)

$$M_{X_1+\cdots+X_n}(t) = M_{X_1}(t)\, M_{X_2}(t) \cdots M_{X_n}(t)$$

$S = ?$ know!

Why do we care?

- As we said there is a 1-1 relation between a distribution and its MGF or PGF.
- Sometimes, $m_S(t)$ or $p_S(t)$ can be recognised: this is the case for infinitely divisible distributions (Normal, Poisson, Inverse Gaussian, ...) and certain other distributions (Binomial, Negative binomial)
- Otherwise, $m_S(t)$ or $p_S(t)$ can be expanded numerically to get moments and/or probabilities

Example Consider a portfolio of 10 contracts. The losses X_i's for these contracts are i.i.d. Poisson RVs with parameter 100. Determine the distribution of S.

$X_i \sim Poi(\underbrace{100}_{\text{"}\lambda\text{"}})$ $\qquad S = \sum_{i=1} X_i \; (=?)$

$$p_S(t) = \underbrace{p_{X_1}(t)\cdots p_{X_{10}}(t)}_{\text{same}} = \big(\underbrace{p_{X_i}(t)}_{\text{need to know}}\big)^{10}$$

$$P_{X_i}(t) = E[t^X] = \sum_{i=1}^{\infty} \left(\frac{\lambda^i e^{-\lambda}}{i!} \right) t^i$$

$$= e^{-\lambda} \underbrace{\sum_{i=1}^{\infty} \frac{(\lambda t)^i}{i!}}_{= e^{\lambda t}} = e^{\lambda(t-1)}$$

$$P_S(t) = \left(P_{X_i}(t)\right)^{10} = \left(e^{100(t-1)}\right)^{10} = \underbrace{e^{1000(t-1)}}_{= \text{PGF of Poi}(1000)}$$

$$\Rightarrow S \sim \text{Poi}(1000)$$

Example Consider three independent RVs X_1, X_2, X_3. For $i = 1, 2, 3$, X_i has an exponential distribution and $E[X_i] = 1/i$. Derive the PDF of $S = X_1 + X_2 + X_3$ by recognition of the MGF of S.

MGF: $E[e^{tX_i}] = \frac{i}{i-t}$ for $t<i$ (Exercise)

$$M_S(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdot M_{X_3}(t)$$
$$= \left(\frac{1}{1-t}\right) \cdot \left(\frac{2}{2-t}\right) \cdot \left(\frac{3}{3-t}\right)$$

$$M_S(t) = \left(\frac{1}{1-t}\right)\left(\frac{2}{2-t}\right)\left(\frac{3}{3-t}\right)$$

$$= 3\left(\frac{1}{1-t}\right) - 3\left(\frac{2}{2-t}\right) + \left(\frac{3}{3-t}\right)$$

by Partial Fractions

Note

$f_X(x) = g(x) + h(x)$ $\qquad E[e^{tX}] = \int e^{tx}(g+h)\,dx = \int + \int$

$$f_S(x) = 3\left(e^{-x}\right) - 3\left(2e^{-2x}\right) + \left(3e^{-3x}\right)$$

Convolutions

- The operation of computing the distribution of the sum of two independent random variables is called a convolution. It is denoted by:

$$F_{X+Y} = F_X * F_Y$$

- The result can then be convoluted with the distribution of another random variable:

$$F_{X+Y+Z} = F_Z * F_{X+Y}$$

- And so on...(as we will see for *n*-fold convolutions)

Formulas

Continuous case:

- CDF: $F_{X+Y}(s) = \int_{-\infty}^{\infty} F_Y(s-x) f_X(x)\, dx$
- PDF: $f_{X+Y}(s) = \underbrace{\int_{-\infty}^{\infty} f_Y(s-x) f_X(x)\, dx}_{\text{pdf of } X+Y} = f_Y \circledast f_X$

Discrete case:

- CDF: $F_{X+Y}(s) = \sum_x F_Y(s-x) f_X(x)$
- PMF: $f_{X+Y}(s) = \sum_x f_Y(s-x) f_X(x)$

n-fold convolutions

For i.i.d. continuous random variables X_i with a common CDF $F_X(x)$, the n-fold convolution of $F_X(x)$ is denoted by $F_X^{*n}(x)$:

$$\begin{aligned} F_X^{*k}(x) &= \int_{-\infty}^{\infty} F_X^{*(k-1)}(x-y)\, f_X(y) dy \\ &= \int_0^x F_X^{*(k-1)}(x-y)\, f_X(y) dy \quad \text{if positive support} \end{aligned}$$

for $k = 1, 2, \ldots$ where:

$$F_X^{*0}(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

n-fold convolutions (Cont.)

Continuous case PDF for $k = 1, 2, \ldots$:

$$\begin{aligned} f_X^{*k}(x) &= \int_{-\infty}^{\infty} f_X^{*(k-1)}(x-y)\, f_X(y)dy \\ &= \int_0^x f_X^{*(k-1)}(x-y)\, f_X(y)dy \quad \text{if positive support} \end{aligned}$$

n-fold convolutions (Cont.)

Discrete case:

- CDF:

$$F_X^{*k}(x) = \sum_{y=0}^{x} F_X^{*(k-1)}(x-y)\, f_X(y) \text{ for } x = 0, 1, \ldots, \;\; k = 2, 3, \ldots$$

- PMF:

$$f_X^{*k}(x) = \sum_{y=0}^{x} f_X^{*(k-1)}(x-y)\, f_X(y) \text{ for } x = 0, 1, \ldots, \;\; k = 2, 3, \ldots$$

Why?

Exercise: Show the convolution gives the distribution for the sum.

Hint: Law of total prob.