MATH 4281 Risk Theory–Ruin and Credibility

Module 1 (cont.)

January 14, 2021

1. Generating Functions and Convolutions (cont.)

2. Frequency and Severity in the IRM

Generating Functions and Convolutions (cont.)

An Example Exercise of a Convolution

Requested last class. Consider 3 independent discrete RVs with PMFs:

$$\begin{aligned} f_1(x) &= \frac{1}{4}, \frac{1}{2}, \frac{1}{4} \text{ for } x = 0, 1, 2 \\ f_2(x) &= \frac{1}{2}, \frac{1}{2} \text{ for } x = 0, 2 \\ f_3(x) &= \frac{1}{4}, \frac{1}{2}, \frac{1}{4} \text{ for } x = 0, 2, 4 \end{aligned}$$

Complete the following table for the PMF f_{1+2+3} and the CDF F_{1+2+3} of the sum of the three random variables.

	Given		(1)	Gave	(2)	(3)
x	$f_1(x)$	$f_2(x)$	$f_{1+2}(x)$	$f_3(x)$	$f_{1+2+3}(x)$	$F_{1+2+3}(x)$
0	1/4	1/2	1/8	1/4	1/32	1/32
1	1/2	0	–	0	–	3/32
2	1/4	1/2	–	1/2	–	7/32
3	0	0	–	0	–	–
4	0	0	–	1/4	–	–
5	0	0	0	0	–	–
6	0	0	0	0	–	–
7	0	0	0	0	–	–
8	0	0	0	0	–	–

E.g. $f_{1+2}(0) = f_1(0)f_2(0) = \left(\frac{1}{4}\right)\left(\frac{1}{2}\right) = \frac{1}{8}$ as given.

$$f_{1+2}(1) = \sum_{y=0}^{1} f_1(x-y) f_2(y) = f_1(1-0)f_2(0) + f_1(1-1)f_2(1)$$
$$= (1/2)(1/2) + (1/4)(0) = 1/4 = 2/8$$

Another Example Exercise of a Convolution

Consider independent $X, Y \sim \mathcal{U}[0,1]$. Find the pdf of $X+Y$:

$Z = X + Y$ $\quad$ $$f_Z(z) = \int_{-\infty}^{\infty} \underbrace{f_X(z-y) f_Y(y)}_{= 1 \text{ or } 0} \, dy$$

① $0 < y < 1 \Rightarrow f_Y = 1$

$0 < z - y < 1 \Rightarrow 0 < y < z \Rightarrow f_X(z-y) = 1$

$$f_Z(z) = \int_0^z f_X(z-y) f_Y(y) \, dy = \int_0^z dy = z$$

(2) Note $0 \le X + Y \le 2$ so consider $z > 1$

$z > 1$; $\underbrace{z - y < 1}_{\text{in order for } f_X(x-z) = 1}$ $\Rightarrow$ $\underbrace{z - 1 < y < 1}_{f_Y(y) = 1}$

$$f_Z(z) = \int_{z-1}^{1} f_X(z - y) f(y)\, dy = \int_{z-1}^{1} (1)\, dy = y \Big|_{z-1}^{1} = 1 - (z - 1)$$

$$= 2 - z$$

$$\boxed{f_Z(z) = \begin{cases} z & z \in [0, 1] \\ 2 - z & z \in (1, 2] \end{cases}}$$

The Normal MGF

A quick review of how to handle some kinds of Gaussian integrals. Note that[1]:

$$tx - \frac{(x-\mu)^2}{2\sigma^2} = -\frac{(x-(\mu+\sigma^2 t))^2}{2\sigma^2} + \mu t + \frac{\sigma^2 t^2}{2}$$

Then clearly for $X \sim \mathcal{N}(\mu, \sigma^2)$:

$$\underbrace{E[e^{tX}]}_{= \int e^{tx} e^{-(x-\mu)^2/2\sigma^2} dx} = e^{\mu t + \frac{\sigma^2 t^2}{2}} \left(\underbrace{\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{(x-(\mu+\sigma^2 t))^2}{2\sigma^2}} dx}_{=1} \right) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

[1] complete the square

Another IRM example

Example: Consider a portfolio of 10 contracts. The losses X_i's for these contracts are i.i.d. Normal RVs with mean 100 and variance 100. Determine the distribution of S.

$$M_S(t) = \prod_{i=1}^{10} M_{X_i}(t) = \left(e^{\mu t + \sigma^2 t^2/2} \right)^{10}$$

$$= \underbrace{e^{[10\mu]t + 5\sigma^2 t^2}}_{\text{Also Normal!}}$$

Normal Approximations for the distribution of the Sum

- Assume $X_1, \cdots, X_n$ are independent and $S = X_1 + \cdots + X_n$.
- Then $E[S] = \sum_{i=1}^{n} E[X_i]$, $Var[S] = \sum_{i=1}^{n} Var[X_i]$
- When n is large (at least 30), the distribution of $\boxed{\frac{S-E[S]}{\sqrt{Var(S)}}}$ can be approximated by the standard normal distribution.

Theoretic Foundation of Normal Approximations

- The central limit theorem[2]:

$$\frac{S - E[S]}{\sqrt{Var(S)}} \xrightarrow{d} N(0,1)$$

- Q: why the "d" above the arrow?
- Q: How could this apply to the normal approximation to the binomial I used yesterday?
- Q: How to prove the CLT via using MGFs

[2] Theorem 3.7 of the loss models textbook

A proof of the CLT

$$\frac{S - E[S]}{\sqrt{Var(S)}} = \frac{1}{\sqrt{n}} \frac{\sum X_i - \mu}{\sigma} = \frac{\left(\sum \left(\frac{X_i - \mu}{\sigma}\right)\right)}{\sqrt{n}}$$

$$M_{\frac{\sum\left(\frac{X_i-\mu}{\sigma}\right)}{\sqrt{n}}}(t) = E\left[\exp\left\{\frac{t}{\sqrt{n}}\left(\sum \frac{X_i-\mu}{\sigma}\right)\right\}\right]$$

A proof of the CLT

X_i's i.i.d

$$e^x = \lim \left(1 + \frac{x}{n}\right)^n$$

$$M_S(t) = \left(M_{\left(\frac{X_i - \mu}{\sigma}\right)}\left(\frac{t}{\sqrt{n}}\right) \right)^n$$

taylor expand

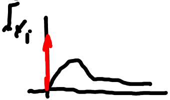
$$= \left(1 + \underbrace{E\left[\frac{X_i - \mu}{\sigma}\right]}_{0}\left(\frac{t}{\sqrt{n}}\right) + \frac{1}{2}\underbrace{\mathrm{Var}\left(\frac{X_i - \mu}{\sigma}\right)}_{1}\left(\frac{t}{\sqrt{n}}\right)^2 + \dots \right)^n$$

$$= \left(1 + \frac{t^2/2}{n} + \dots\right)^n \longrightarrow e^{t^2/2}$$ MGF of $N(0,1)$

Frequency and Severity in the IRM

A Problem Unique to the IRM

- In the CRM we call N the "frequency distribution" and X_i the "severity".
- Recall in the IRM N is fixed at n, some number we know a priori.
- But not every individual is always claiming coverage, in fact, the opposite is true.
  $\Rightarrow$ Must be a big mass of probability at $x = 0$!



- How to handle this?

A Problem Unique to the IRM

For example consider a individual loss like so:

$$\begin{cases} Pr\left(X=0\right)=1/2, \\ f_X(x)=\frac{1}{2}\beta e^{-\beta x}, \text{ for } \beta=0.1, \quad x>0 \end{cases}$$

- Q: How easily can we take convolutions?
- Q: How easily can we take *n-fold* convolutions?
- Q: Mean? Var? MGFs?

How to Separate Frequency from Severity

One approach is to define $X = IB$, where:

- I is an *indicator* of claim with

$$\Pr[I = 1] = q \text{ and } \Pr[I = 0] = 1 - q$$

- B is the claim amount given $I = 1$ (i.e. given a claim occurs).

The distribution function:

Assume $\Pr[I=1]=q$ and $\Pr[X<0]=0$, then for $x \geq 0$:

Law of total Prob.

$$\begin{aligned}
\Pr[X \leq x] &= \underbrace{\Pr[X \leq x | I=0]} \Pr[I=0] + \Pr[X \leq x | I=1] \Pr[I=1] \\
&= (1)(1-q) + (q) \Pr[(1)B \leq x | I=1] \\
&= 1 - q + q \Pr[B \leq x]
\end{aligned}$$

$$\begin{aligned}
P(X \leq x \mid I=0) &= P(IB \leq x \mid I=0) \\
&= P(0 \leq x) \\
&= 1
\end{aligned}$$

Moments

- The Mean[3]:

$$E[X] = E[E[X|I]] = E[X|I = 1]\Pr[I = 1] = qE(B),$$

- Variance[4]:

$$\begin{aligned} Var(X) &= Var(E[X|I]) + E[Var(X|I)] \\ &= [E(B)]^2\, Var(I) + qVar(B) \\ &= q(1-q)(E[B])^2 + qVar(B) \end{aligned}$$

after noting that $E[X|I] = I \cdot E[B],\ Var(X|I) = I^2 \cdot Var(B)$.

[3] Recall the "Tower Property"

[4] The first line makes use of the "Law of Total Variance"

Generating Functions

- MGF:

$$\begin{aligned} M_X(t) &= E[e^{tX}|I=0]\Pr(I=0) + E[e^{tX}|I=1]\Pr(I=1) \\ &= 1-q+E[e^{tB}]q = 1-q+M_B(t)q \end{aligned}$$

- PGF:

$$\begin{aligned} P_X(t) &= E[t^X|I=0]\Pr(I=0) + E[t^X|I=1]\Pr(I=1) \\ &= 1-q+P_B(t)q \end{aligned}$$

Aggregate loss: $S = \sum_{i=1}^{n} X_i$

- Each X_i is separated by $X_i = I_i B_i$, for $i = 1, 2, \ldots, n$
- Mean: $E[S] = \sum_{i=1}^{n} q_i \mu_i$, where $q_i = \Pr(I_i = 1)$ and $\mu_i = \mathrm{E}[B_i]$
- Variance

$$Var(S) = \sum_{i=1}^{n} [q_i \sigma_i^2 + q_i(1 - q_i)\mu_i^2]$$

  where $\sigma_i^2 = \mathrm{Var}(B_i)$
- MGF:

$$M_S(t) = \prod_{i=1}^{n} [1 - q_i + M_{B_i}(t) q_i]$$

- What is the PGF? (Exercise)

A Familiar Example

Suppose claim amount X is distributed as:

$$\begin{cases} P(X=0) = 1/2, \\ f_X(x) = \frac{1}{2}\beta e^{-\beta x}, \text{ for } \beta = 0.1, \quad x > 0 \end{cases}$$

1. Find the expected value of X.
2. Find I and B such that $X = IB$.

A Familiar Example

A Familiar Example

Another Example

Example In an insurance portfolio, there are 15 insured. Ten of the insured persons have 0.1 probability of making a claim, and the other 5 have a 0.2 probability. All claims are independent and follow $Exp(\lambda)$ (Note: $1/\lambda$ is the mean). What is the MGF of the aggregate claims distribution?