ACTSC 431 - Loss Model I

CLASSNOTES FOR FALL 2018

bv

Johnson Ng

BMath (Hons), Pure Mathematics major, Actuarial Science Minor University of Waterloo

Table of Contents

Lis	st of	Definit	ions	5
Li	st of	Theore	ems	7
1	Lect	ture 1 S	Sep 06	9
	1.1	Introd	luction and Overview	9
		1.1.1	Individual Risk Model	10
		1.1.2	Collective Risk Model	11
2	Lect	ture 2 S	Sep 11th	13
	2.1	Revie	w of Probability Theory	13
		2.1.1	Discrete Random Variables	13
		2.1.2	Continuous Random Variables	15
3	Lect	ture 3 S	Sep 13th	19
	3.1	Revie	w of Probability Theory (Continued)	19
		3.1.1	Continuous Random Variables (Continued)	19
		3.1.2	Mixed Random Variable	19
	3.2	Distri	butional Quantities and Risk Measures	21
		3.2.1	Distributional Quantities	21
4	Lect	ture 4 S	Gep 18th	25
	4.1	Distri	butional Quantities and Risk Measures (Continued)	25
		4.1.1	Risk Measures	25
			4.1.1.1 Value-At-Risk	26
			4.1.1.2 Tail-Value-at-Risk	28
5	Lect	ture 5 S	Sep 20th	31
	5.1	Distrb	outional Quantities and Risk Measures (Continued	
		2)		31
		5.1.1	Risk Measures (Continued)	31
	F 2	Sovori	ity Distributions - Croating Soverity Distributions	2.4

4 TABLE OF CONTENTS - TABLE OF CONTENTS

A Additional Material	3	37		
A.1 Individual Risk Model: An Alternate View .		37		
A.2 Coherent Risk Measure	3	38		
Dibli-consulor		ļ1		
Bibliography				
List of Symbols and Abbreviations	4	13		
Index	4	l 5		

List of Definitions

1	Liability Risk	9
2	Individual Risk Model	10
3	Collective Risk Model	11
4	Severity Distribution	11
5	Frequency Distribution	12
6	Aggrement Payment / Loss	12
7	Random Variable	13
8	Discrete Random Variable	13
9	Probability Generating Function	14
10	Continuous Random Variable	15
11	Moment Generating Function	16
12	Hazard Rate Function	17
13	Mixed Random Variable	19
14	Deductibles	20
15	Central Moment	22
16	Quantiles	23
17	Risk Measure	25
18	Premium Principle	25
19	Value-at-Risk (VaR)	26
20	Tail-Value-at-Risk (TVaR)	28
21	Tail-Value-at-Risk (TVaR)	29

6 TABLE OF CONTENTS - TABLE OF CONTENTS

22	Mean Excess Loss
23	Severity Distribution
24	Empirical Distribution Function 35
25	Coherent Risk Measure

List of Theorems

ð	Proposition 1	Relation of $TVaR_p(X)$ and $e_X(d)$		31
ð	Proposition 2	Expection from Survival Function		32
ð	Proposition 3	Expected Deductible		33
6	Proposition 4	An Expression for Mean Excess Value .		34

1 Lecture 1 Sep 06

1.1 Introduction and Overview

Course Objective In Loss Model I, the focus of our study is to learn the basic methods which are used by insurers to quantify risk from mathematical/statistical models, in order for insurers to make various decisions¹. By quantifying risk, it helps us monitor underlying risks so that not only are we aware of them, but also so that we can take actions or preventive measures against them.

Our main interest of this course is:

- to quantify and seek protection against the loss of funds due either to too many claims or a few large claims;
- to reduce adverse financial impact of random events that prevent the realization of reasonable expectations.

The main model that shall be the focus of this course is **models** for liability risk.

Definition 1 (Liability Risk)

A *liability risk* is a risk that insurance companies assume by selling insurance contracts.

In particular, the liability that we shall focus on is **insurance** claims.

We are Interested in modelling the total amount of claims, i.e. the **aggregate claim amount**, of a group fo insurance policies over a

¹ e.g. setting premiums, control expenses, deciding for reinsurance, etc.

Many of the models that we shall see later in the course are also applied for other types of risks, e.g. investment risk, credit risk, liquidity risk, and operational risk. given period of time. In the actuarial literature, there are two main approaches that have been proposed to model the aggrement claim amount of an insurance portfolio, namely:

- individual risk model;
- collective risk model.

1.1.1 Individual Risk Model

Definition 2 (Individual Risk Model)

In an individual risk model, the aggregate claim is modeled by

$$S = \sum_{i=1}^{n} Z_i$$

where n is a deterministic² integer that represents the total number of insurance policies, and Z_i is a random variable for the potential loss of the ith insurance policy.

² i.e. fixed

66 Note

Since a policy may or may not incur a loss³, we have that

$$P(Z_i=0)>0.$$

Thus, in an individual risk model, we may also express the aggregate claim amount as

$$S = \sum_{i=1}^{n} X_i I_i$$

where I_i is the indicator function about the claimant of policy i, while X_i represents the size of the claim(s) for the i^{th} policy provided that there is a claim.⁴

³ Since a claim may or may not be made!

However, in an individual risk model, according to Dhaene and Vyncke $(2010)^5$,

A third type of error that may arise when computing aggregate claims follows from the fact that the assumption of mutual independency of the individual claim amounts may be violated in practice.

⁴ This is actually incorrect, despite being in the recommended textbook. See Appendix A.1.

⁵ Dhaene, J. and Vyncke, D. (2010). The individual risk model. https://www. researchgate.net/publication/ 228232062_The_Individual_Risk_ Model

Due to complications such as this, the individual risk model will not be the focus of our studies.

Collective Risk Model 1.1.2

Definition 3 (Collective Risk Model)

In a collective risk model, the aggregate claim is modeled by

$$S = \sum_{i=1}^{N} X_i,$$

where N is a non-negative integer-valued random variable that denotes the number of claims among a given set of policies, while X_i denotes the size of the ith policy.

66 Note

In a collective risk model, we need to determine:

- the distribution of the total number of claims for the entire portfolio, i.e. the distribution of N; and
- the distribution of the loss amount per claim, i.e. the distribution of X_i .

In this course, the primary focus of our studies will be on collective risk models.

Terminologies To end today's lecture, the following terminologies are introduced:

Definition 4 (Severity Distribution)

The severity distribution is the distribution of the loss amount of the amount paid by the insurer on a given loss/claim.

Definition 5 (Frequency Distribution)

The *frequency distribution* is the distributino fo the number of losses/claims paid by the insurer over a given period of time.

66 Note

The frequency distribution is typically a discrete distribution.

Definition 6 (Aggrement Payment / Loss)

The aggregate payment (loss) is the total amout of all claim payments (losses) over a given period of time.

66 Note

There is a distinction between an aggregate payment and an aggregate loss, since an aggregate payment is "essentially" an aggregate loss after certain claim adjustments, such as deductibles, limits, and coinsurance.

2 Lecture 2 Sep 11th

2.1 Review of Probability Theory

Firstly, we shall review the definition of a random variable.

Definition 7 (Random Variable)

Let Ω be a sample space and \mathcal{F} its σ -algebra¹. A **random variable** (rv) $X:\Omega\to(\Omega,\mathcal{F})$ is a function from a possible set of outcomes to a measurable space (Ω,\mathcal{F}) . Within the context of our interest, X is real-valued, i.e. $(\Omega,\mathcal{F})=\mathbb{R}$.

 $^{\scriptscriptstyle 1}$ For definitions of Ω and ${\cal F}$, see notes on STAT330.

2.1.1 Discrete Random Variables

Definition 8 (Discrete Random Variable)

A discrete random variable (drv) is an rv X that takes only countable (finite) real values.

66 Note

Let X be a drv.

• The probability mass function (pmf) of X is: for $i \in \mathbb{N}$,

$$p(x_i) = P(X = x_i)$$

• The cumulative distribution function (cdf) of X is

$$F(x) = P(X \le x) = \sum_{x_i \le x} p(x_i).$$

• The kth moment of X is²

$$E[X^k] = \sum_{i \in \mathbb{N}} x_i^k p(x_i)$$

if $E[X^k]$ is finite.

• Some commonly seen/introduced discrete distributions are: Poisson, Binomial, Negative Binomial

² This implicitly uses the Law of the Unconcious Statistician.

Example 2.1.1

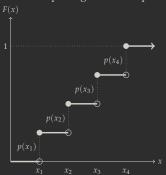
Let X take values from $\{x_1, x_2, x_3, x_4\}$, and

$$p(x_i) = P(X = x_i)$$
 for $i = 1, 2, 3, 4$.

The cdf of X is

$$F(x) = \begin{cases} 0 & x < x_1 \\ p(x_1) & x_1 \le x < x_2 \\ p(x_1) + p(x_2) & x_2 \le x < x_3 \\ 1 - p(x_4) & x_3 \le x < x_4 \\ 1 & x \ge x_4 \end{cases}$$

It is recommended to visualize the cdf first before putting it down in pencil.



66 Note

- It is important that we stress the need for showing right continuity in the graph.
- *Note that the cdf always sums to* 1.
- The "jumps" at x_i correspond to $p(x_i)$, for i = 1, 2, 3, 4.

Definition 9 (Probability Generating Function)

Suppose a drv X only takes non-negative integer values. The proba-

bility generating function (pgf) of X is defined as

$$G(z) = E\left[z^X\right] = \sum_{k=1}^{\infty} z^k p(k)$$

where we note that if $\max X = n$, then p(m) = 0 for all m > n.

66 Note

- The pgf uniquely identifies the distribution of the drv³.
- To get the probability for $k \in \{0, 1, 2, ...\}$, we simply need to do

$$p(k) = \frac{1}{k!} G^{(k)}(x) \Big|_{x=0}.$$

³ This was given as is without proof, and I cannot find any resources that proves this.

Example 2.1.2 (Lecture Slides: Example 1)

Consider a drv X with pmf

$$p(x) = P(X = x) = \begin{cases} 0.5 & x = 0 \\ 0.4 & x = 1 \\ 0.1 & x = 2 \end{cases}$$

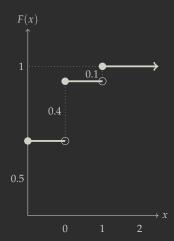
Its cdf is

$$F(x) = P(X \le x) \begin{cases} 0 & x < 0 \\ 0.5 & 0 \le x < 1 \\ 0.9 & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}$$

and its pgf is

2.1.2

$$G(z) = E[z^X] = 0.5 + 0.4z + 0.1z^2.$$



Continuous Random Variables

Definition 10 (Continuous Random Variable)

A continuous random variable (crv) takes on a continuum of values.

66 Note

Let X be a crv.

• $\exists f: X \to \mathbb{R}$ called a probability density function (pdf) such that its cdf is

$$F(x) = \int_{-\infty}^{x} f(y) \, dy,$$

and consequently by the Fundamental Theorem of Calculus, we have

$$f(x) = F'(x).$$

• *The kth moment of X is*

$$E[X^k] = \int_{\mathcal{X}} x^k f(x) \, dx$$

so long that $E[X^k]$ is defined.

• Some commonly introduced distributions are: Uniform, Exponential, Gamma, Weibull, and Normal.

Definition 11 (Moment Generating Function)

Let X be an rv. The **moment generating function** (mgf) of X is, for $t \in \mathbb{R}$ (appropriately so),

$$M_X(t) = E\left[e^{tX}\right] = \int_X e^{tx} f(x) dx$$

provided that the integral is well-defined.

The mgf is also defined for drvs.

66 Note

- The mgf uniquely determines the distribution of its rv⁴
- With the mgf, we can obtain the kth moment of an rv X by

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

⁴ This shall, also, not be proven in this

Example 2.1.3 (Lecture Notes: Example 2)

Consider an exponential rv X with pdf⁵

⁵ When not explicitly stated, it shall be assumed that domains at which we did not specify *x* shall have probability 0.

$$f(x) = 0.1e^{-0.1x}, \ x > 0.$$

Its cdf is

$$F(x) = \int_{-\infty}^{x} f(y) \, dy = \begin{cases} 1 - e^{-0.1x} & x \ge 0\\ 0 & \text{otherwise} \end{cases}$$

and its mgf is

$$M_X(t) = E\left[e^{tX}\right] = \int_0^\infty e^{tx} 0.1 e^{-0.1x} dx$$
$$= 0.1 \int_0^\infty e^{(t-0.1)x} dx$$
$$= \frac{0.1}{0.1 - t}, \ t < 0.1,$$

where we note that we must have t < 0.1, for otherwise the value of the exponent would render the integral undefined.

Definition 12 (Hazard Rate Function)

For a cro X, the hazard rate function (aka failure rate) of X is defined

$$h(x) = \frac{f(x)}{\overline{F}(x)} = -\frac{d}{dx} \ln \overline{F}(x),$$

where $\overline{F}(x) = 1 - F(x)$ is the survival function⁶

⁶ You should be familiar with this if you have studied for Exam P.

66 Note

• We may also express the survival function in terms of the hazard rate by

$$\bar{F}(x) = e^{-\int_{-\infty}^{x} h(y) \, dy}.$$

• In terms of limits, we can express the hazard rate function, for small enough $\delta > 0$, as

$$h(x) = \frac{f(x)}{\overline{F}(x)} = \frac{F'(x)}{\overline{F}(x)}$$

$$\approx \frac{F(x+\delta) - F(x)}{\delta \overline{F}(x)}$$

$$= \frac{P(x < X \le x + \delta)}{\delta F(X > x)}$$

$$= \frac{1}{\delta} P(x < X \le x + \delta \mid X > x).$$

We can make sense of this expression by recalling the notion of the probability of survival from Exam MLC7, where if a life has survived over x, the hazard rate is the probability that the life does not survive beyond another δ 8 .

- ⁷ This also tells us that the hazard rate gets its name from life insurance.
- ⁸ From the perspective of life insurance, the greater the probability, the more likely the claim is going to happen.

3 Lecture 3 Sep 13th

3.1 Review of Probability Theory (Continued)

3.1.1 Continuous Random Variables (Continued)

Example 3.1.1 (Lecture Notes: Example 3)

Suppose $X \sim Wei(\theta, \tau)$ *with pdf*

$$f(x) = \frac{\tau \left(\frac{x}{\theta}\right)^{\tau} e^{-\left(\frac{x}{\theta}\right)^{\tau}}}{x}, \quad x > 0,$$

where θ , $\tau > 0$. Find its hazard rate function.

Solution

We first require the survival function:

$$\begin{split} \bar{F}(x) &= \int_{x}^{\infty} \frac{1}{y} \tau \left(\frac{y}{\theta}\right)^{\tau} e^{-\left(\frac{y}{\theta}\right)^{\tau}} dy \\ &= \int_{\frac{x}{\theta}}^{\infty} \frac{1}{u} \tau u^{\tau} e^{-u^{\tau}} du \qquad \text{where } u = \frac{y}{\theta} \\ &= \int_{\frac{x}{\theta}}^{\infty} \tau u^{\tau - 1} e^{-u^{\tau}} du \\ &= -e^{-u^{\tau}} \Big|_{\frac{x}{\theta}}^{\infty} = e^{-\left(\frac{x}{\theta}\right)^{\tau}} \end{split}$$

The hazard rate is therefore

$$h(x) = \frac{f(x)}{\overline{F}(x)} = \frac{\tau}{x} \left(\frac{x}{\theta}\right)^{\tau}$$

3.1.2 Mixed Random Variable

We call X a mixed random variable (mixed rv) if it has both discrete and continuous components.

66 Note

 Mixed rvs are important in modeling insurance claims, e.g., the loss amount is usually a continuous random variable with a probability mass at 0.

The following is a type of mixed random variable:

Definition 14 (Deductibles)

Let X be an rv and d be a fixed value.

$$[X-d]_+ = egin{cases} X-d & x \geq d \ 0 & otherwise \end{cases}$$

66 Note

If X be an rv and d a fixed value, the deductible $[X-d]_+$ has a mass point at 0 since

$$P([X-d]_+ = 0) = P(X < d) > 0$$

66 Note

Let $\{x_1, x_2, ...\}$ be a sequence of real numbers in an increasing order. Suppose X is a rv that takes on values on the real, and has a density function f on each interval (x_i, x_{i+1}) , and has discrete mass points at the boundaries of these intervals, i.e.

$$P(X = x_i) = p(x_i) > 0 \quad i \in \mathbb{N}.$$

Since X is an rv, it must be the case that

$$\sum_{i\in\mathbb{N}} p(x_i) + \sum_{i\in\mathbb{N}} \int_{x_i}^{x_{i+1}} f(x) dx = 1.$$

In other words, we treat the discrete and continuous part of a mixed rv separately.

The cdf of a mixed rv X is

$$F(x) = P(X \le x) = \sum_{i \in \mathbb{N}} p(x_i) \mathbb{1}_{\{x_i \le x\}} + \sum_{i \in \mathbb{N}} \int_{x_i}^{x_{i+1}} f(y) \mathbb{1}_{\{y \le x\}} dy.$$

The kth moment of X is

$$E\left[X^k\right] = \sum_{i \in \mathbb{N}} (x_i)^k p(x_i) + \sum_{i \in \mathbb{N}} \int_{x_i}^{x_{i+1}} x^k f(x) \, dx.$$

The mgf of X is

$$M_X(t) = E\left[e^{tX}\right] = \sum_{i \in \mathbb{N}} e^{tx_i} p(x_i) + \sum_{i \in \mathbb{N}} \int_{x_i}^{x_{i+1}} e^{tx} f(x) dx.$$

Example 3.1.2 (Lecture Notes: Example 4)

Assume a claim amount of an insurance policy is modeled by a non-negative rv X which has probability mass of p and 0, and otherwise continuous with a pdf f over $(0, \infty)$. Find its cdf, kth moment, and mgf.

Solution

The cdf of X is

$$F(x) = \begin{cases} p + \int_0^x f(y) \, dy & x \ge 0\\ 0 & \text{otherwise} \end{cases}$$

The kth moment of X is

$$E\left[X^k\right] = \int_0^\infty x^k f(x) \, dx.$$

The mgf of X is

$$M_X(t) = p + \int_0^\infty e^{tx} f(x) \, dx.$$

3.2 Distributional Quantities and Risk Measures

This chapter introduces us to some distributional quantities for a given rv X. These distributional quantities are informative values to describe the characteristics of a risk.

3.2.1 Distributional Quantities

Definition 15 (Central Moment)

The kth central moment of an rv X is defined as

$$E\left[(X-E(X))^k\right].$$

66 Note

The second central moment is the variance. The square root of the variance is the standard deviation.

Example 3.2.1 (Lecture Notes: Example 5)

Consider an rv $Y = \begin{cases} Y_1 & U = 1 \\ Y_2 & U = 2 \end{cases}$, where $Y_1 = 0$, $Y_2 \sim \text{Exp}(10)$, and P(U = 1) = P(U = 2) = 0.5.

¹ This notation is just syntatic sugar for saying $Y_1 = Y \mid (U = 1)$ and $Y_2 = Y \mid (U = 2)$.

- 1. Find the cdf of Y.
- 2. Find the mean and variance of Y.
- 3. Let $Z = \frac{1}{2}Y_1 + \frac{1}{2}Y_2$. Does Z have the same distribution as Y? Answer this by solving the mean and variance of Z.

Solution

1. Note that

$$F(y) = P(Y_1 \le y \mid U = 1)P(U = 1) + P(Y_2 \le y \mid U = 2)P(U = 2).$$

Observe that

$$P(Y_1 \le y \mid U = 1) = \begin{cases} 1 & y \ge 0 \\ 0 & y < 0 \end{cases}$$

and

$$P(Y_2 \le y \mid U = 2) = \begin{cases} 1 - e^{-10y} & y \ge 0 \\ 0 & y < 0 \end{cases}$$

Therefore

$$F(y) = \begin{cases} 1 - \frac{1}{2}e^{-10y} & y \ge 0\\ 0 & y < 0 \end{cases}$$

2. The mean of Y is

$$E(Y) = E(Y \mid U = 1)P(U = 1) + E(Y \mid U = 2)P(U = 2) = 10 \cdot \frac{1}{2} = 5.$$

To calculate the variance of Y, we require

$$E[Y^{2}] = E[Y^{2} \mid U = 1]P(U = 1) + E[Y^{2} \mid U = 2]P(U = 2)$$
$$= (Var(Y_{2}) + E(Y_{2})^{2}) \cdot \frac{1}{2} = 100.$$

Therefore

$$Var(Y) = 100 - 5^2 = 75.$$

3. The mean of Z is

$$E[Z] = E[\frac{1}{2}Y_1 + \frac{1}{2}Y_2] = 5.$$

The variance of Z is

$$Var(Z) = \frac{1}{4} Var(Y_1) + \frac{1}{4} Var(Y_2) = 25.$$

Therefore, Z does not have the same distribution as Y.

Definition 16 (Quantiles)

The 100p% quantile (or percentile) of an rv X is a set π_v such that

$$\pi_p = \{ x \in X \mid P(X < x) \le p \le P(X \le x) \}.$$

This definition may also be presented as: any number π_p such that

$$P(X < \pi_p) \le p \le P(X \le \pi_p).$$

66 Note

• If X is a continuous random variable, we have that $P(X < \pi_p) =$ $P(X \leq \pi_p)$ and so we have to define the quantile as

$$\pi_p = F^{-1}(p)$$

where F^{-1} is the inverse function of F, the cdf of X.

- A quantile can be a set of numbers.
- $\pi_{0.5}$ is called the **median** of X.

Graphical method to interpret this notion will be included.

Find the 100p% quantile of the loss distribution $F(x) = 1 - e^{-\frac{x}{\theta}}$, x > 0.

Solution

Note that F is the cdf of an exponential distribution, which is a continuous distribution. Therefore,

$$F(\pi_p) = 1 - e^{-\frac{\pi p}{\theta}} = p \implies \pi_p = -\theta \ln(1-p).$$

Example 3.2.3 (Lecture Notes: Example 2)

Find the median $\pi_{0.5}$ for the following cdf

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.6 + 0.4(1 - e^{-\frac{x}{3}}) & x \ge 0 \end{cases}$$

Solution

Since F(0) = 0.6 and F is an increasing function, we have that F(x) = 0 for all x < 0. Therefore

$$\pi_{0.5} = 0.$$

Example 3.2.4 (Lecture Notes: Example 3)

Find the median $\pi_{0.5}$ for a loss X with pmf

$$p(0) = 0.25, p(1) = 0.25, p(2) = 0.5.$$

Solution

The cdf of X is

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.25 & 0 \le x < 1 \\ 0.5 & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}$$

since F(x) = 0.5 when $1 \le x < 2$, we have that

$$\pi_{0.5} = [1, 2].$$

4 Lecture 4 Sep 18th

4.1 Distributional Quantities and Risk Measures (Continued)

4.1.1 Risk Measures

Definition 17 (Risk Measure)

A **risk measure** is a mapping from the loss rv to the real line \mathbb{R} .

Klugman, Panjer & Wilmot (2012) ¹ on risk measure:

The level of exposure to risk is often described by one number, or at least a small set of numbers. These numbers are necessarily functions of the model and are often called 'key risk indicators'. Such key risk indicators indicate to risk managers the degree to which the company is subject to particular aspects of risk.

To ensure its solvency, insurers will have to charge on these risks, i.e. we have to **price these exposures to risks**.

Definition 18 (Premium Principle)

A premium principle (or insurance pricing) is a rule for assigning a premium to an insurance risk.

66 Note

The following are some of the common principles used by insurers:

¹ Klugman, S. A., Panjer, H. H., and Willmot, G. E. (2012). *Loss Models: From Data to Decisions*. John Wiley & Sons, Inc., 4th edition • Expectation Principle

$$\Pi(X) = (1 + \theta)E(X), \quad \theta > 0$$

• Standard Deviation Principle

$$\Pi(X) = E(X) + \theta \sqrt{\operatorname{Var}(X)}, \quad \theta > 0$$

• Dutch Principle

$$\Pi(X) = E(X) + \theta E([X - E(X)]_+), \quad \theta > 0$$

One particular measure is known as the Value-at-Risk (VaR).

4.1.1 Value-At-Risk

Definition 19 (Value-at-Risk (VaR))

The Value-at-Risk (VaR) is a quantile of the distribution of aggregate losses, i.e. the VaR of a risk X at the 100%p level is defined as²

$$\pi_p = \operatorname{VaR}_p(X) = \inf\{x \in \mathbb{R} : P(X > x) \le 1 - p\}$$
$$= \inf\{x \in \mathbb{R} : P(X \le x) \ge p\}.$$

² I must find out why we define using inf instead of min (see following remark), and I will not take "safe definition" as an answer without full justification.

66 Note

- VaR is often called a quantile risk measure.
- VaR is the standard risk measure used to evaluate exposure to risks.
- VaR measures the amount of capital required by the insurer to remain solvent, with high certainty, in the face of large claims.
- *In practice, p is generally high:* 99.95% *or as low as* 95%.

Remark

Observe that

$$B = \{x \in \mathbb{R} \mid F_X(x) \ge p\} = (A, \infty) \text{ or } [A, \infty)$$

This remark basically points out that the left endpoint of the interval *B* is always included, which should be quite clear by right-continuity of *F*.

for some $A \in \mathbb{R}$, since F is an increasing function. Now let $x_0 \in B$ such that

$$F(x_0) = P(X \le x_0) \ge p \quad \land \quad F(x_0 -) = P(X < x_0) \le p,$$

i.e. it is not necessary that $P(X = x_0) = p$ (see the two example graphs on the margin).

Let $\{x_n\}_{n\in\mathbb{N}}$ be a decreasing sequence of points on \mathbb{R} such that $x_n\to x_0$ as $n \to \infty$. Since F is right-continuous, we have that $F(x_n) \to F(x_0)$ as $n \to \infty$. Therefore,

$$B = [x_0, \infty)$$

This justifies the definition of π_n .

66 Note

• *Note that by definition, we have*

$$P(X < \pi_p) \le p \le P(X \le \pi_p)$$

• If X is a crv whose cdf is strictly increasing, i.e. no constant points, then

$$\pi_p = F^{-1}(p)$$

since $P(X < \pi_v) = P(X \le \pi_v)$.

* Warning (Shortcomings of VaR)

- VaR cannot tell us the size of the potential loss in the 100(1-p)%cases, making it difficult for us to prepare the right amount in order to safeguard against insolvency.
- VaR actually fails to satisfy properties to be a coherent risk measure³, for example, subadditivity.
- VaR is extensively used in financial risk management of trading risk over a fixed (usually short) time period, which are usually normally distributed, and VaR satisfies all coherency requirements.
- In insurance losses, instead of normal distributions, in general, skewed distributions are used, and in this cases, VaR is flawed as it lacks subadditivity.

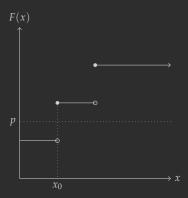


Figure 4.1: Discrete cdf

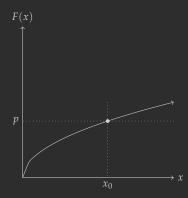


Figure 4.2: Continuous cdf The lecturer asserts that we can really define VaR using min instead of inf, but even with this, I am not completely satisfied or convinced.

³ See Appendix A.2.

Example 4.1.1

Suppose that X has a Pareto distribution with cdf

$$F(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^{\alpha}, \quad x > 0$$

where $\alpha, \theta > 0$. Find $VaR_p(X)$.

Solution

Since F is continuous and strictly increasing, we have that

$$\pi_p = F^{-1}(p) = \theta \left[(1-p)^{-\frac{1}{\alpha}} - 1 \right]$$

Example 4.1.2

Find $VaR_{0.95}(X)$, $VaR_{0.5}(X)$, and $VaR_{0.3}(X)$ for a random loss with pmf

$$p(0) = 0.25$$
, $p(1) = 0.25$, and $p(2) = 0.5$.

Solution

Note that the cdf of X is

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.25 & 0 \le x < 1 \\ 0.5 & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}.$$

Therefore,

$$VaR_{0.95}(X) = 2$$
, $VaR_{0.5}(X) = 1$, and $VaR_{0.3}(X) = 1$.

4.1.1 Tail-Value-at-Risk

To compensate for the weakness of VaR at giving us the size of the the loss *X* of which we cannot measure, we use the Tail-Value-at-Risk.

Definition 20 (Tail-Value-at-Risk (TVaR))

Let X be an rv. The **Tail-Value-at-Risk (TVaR)** of X at the 100p% level, denoted as $TVaR_p(X)$, is defined as the average of all VaR values above the level p, and expressed as

TVaR also has the following names, used by different regions:

- Conditional Tail Expectation (CTE) NA
- Tail Conditional Expectation (TCE)
- Expected Shortfall (ES) EU

$$\text{TVaR}_p(X) = \frac{1}{1-p} \int_p^1 \text{VaR}_\alpha(X) \, d\alpha = \frac{1}{1-p} \int_p^1 \pi_\alpha \, d\alpha$$

Remark

By considering the average of VaR from p's going up to 1, we take into account even the extreme cases of which VaR fails to account for.

Perhaps a clearer definition would be the following, although the expression is only sensible if *X* is a crv:

Definition 21 (Tail-Value-at-Risk (TVaR))

Let X be an rv. The **Tail-Value-at-Risk** (TVAR) of X at the 100p%*level, denoted* $TVaR_{\nu}(X)$ *, is the expected loss given that the loss exceeds* the 100p percentile (or quantile) of the distribution of X, expressible as

$$\mathrm{TVaR}_p(X) = E[X \mid X > \pi_p] = \frac{1}{\overline{F}(\pi_p)} \int_{\pi_p}^{\infty} x f(x) \, dx.$$

Note that the two definitions agree with one another:

$$\frac{1}{1-p} \int_p^1 \pi_\alpha \, d\alpha = \frac{1}{1-F(\pi_p)} \int_p^1 F^{-1}(\alpha) \, d\alpha$$
$$= \frac{1}{\overline{F}(\pi_p)} \int_{\pi_p}^1 x f(x) \, dx$$

where we let $\alpha = F(x)$ as substitution.

66 Note

While it is not difficult to notice that

$$TVaR_{\nu}(X) \geq VaR_{\nu}(X)$$
,

the proof is also simple:

$$ext{TVaR}_p(X) = rac{1}{1-p} \int_p^1 \pi_{lpha} \, dlpha \ \geq rac{1}{1-p} \pi_p \int_p^1 dlpha = \pi_p = ext{VaR}_p(X).$$

Example 4.1.3

Find TVaR_p(X) for $X \sim \text{Exp}(\theta)$.

Solution

Since X is a crv, and $F(x) = 1 - e^{-\frac{x}{\theta}}$, we have that

$$\pi_p = F^{-1}(p) = -\theta \ln(1-p).$$

Therefore,

$$TVaR_{p}(X) = \frac{1}{1-p} \int_{p}^{1} \pi_{\alpha} d\alpha = \frac{-\theta}{1-p} \int_{p}^{1} \ln(1-\alpha) d\alpha$$

$$= \frac{-\theta}{1-p} \int_{-\infty}^{\ln(1-p)} ue^{u} du \quad let \ u = \ln(1-\alpha)$$

$$= \frac{-\theta}{1-p} \left[ue^{u} \Big|_{-\infty}^{\ln(1-p)} - \int_{-\infty}^{\ln(1-p)} e^{u} du \right] \ by \ IBP$$

$$= \frac{-\theta}{1-p} \left[(1-p) \ln(1-p) - (1-p) \right]$$

$$= \theta [1 - \ln(1-p)]$$

66 Note

From the last example, by the memoryless property of $Exp(\theta)$, notice that we may also do

$$\begin{aligned} \text{TVaR}_p(X) &= E[X \mid X > \pi_p] = E[X - \pi_p + \pi_p \mid X > \pi_p] \\ &= E[X - \pi_p \mid X > \pi_p] + E[\pi_p \mid X > \pi_p] \\ &= E[X] + \pi_p \end{aligned} \tag{4.1}$$

5 Lecture 5 Sep 20th

5.1 Distrbutional Quantities and Risk Measures (Continued 2)

5.1.1 Risk Measures (Continued)

Before ending this section, we introduce a notion that is related to TVaR.

Definition 22 (Mean Excess Loss)

Let X be an rv, and $d \in \mathbb{R}$. The **mean excess loss**, denoted $e_X(d)$, is defined as

$$e_X(d) = E[X - d \mid X > d]$$

and $e_X(d) = 0$ for those d such that P(X > d) = 0.

• Proposition 1 (Relation of TVa $R_p(X)$ and $e_X(d)$)

For a crv X, we have

$$TVaR_{p}(X) = e_{X}(\pi_{p}) + VaR_{p}(X)$$

Proof

By Equation (4.1), we have that

$$\text{TVaR}_p(X) = E[X - \pi_p \mid X > \pi_p] + \pi_p = e_X(\pi_p) + \pi_p.$$

• Proposition 2 (Expection from Survival Function)

Let X be a non-negative rv such that $E[X^k] < \infty$, for any $k \in \mathbb{N} \setminus \{0\}$. Then¹

$$E\left[X^{k}\right] = k \int_{0}^{\infty} x^{k-1} \overline{F}(x) \, dx$$

¹ Note that this works for the discrete case as well, by replacing \int with Σ .

Proof

Firstly, note that since $E[X^k] < \infty$ for all $k \in \mathbb{N} \setminus \{0\}$, we have that $\overline{F}(x)$ decays faster than x^k as $x \to \infty$. Now

$$E\left[X^{k}\right] = \int_{0}^{\infty} x^{k} f(x) dx \quad \therefore \text{ Law of the Unconscious Statistician}$$

$$= \int_{0}^{\infty} x^{k} dF(x) \quad \therefore dF(x) = f(x) dx$$

$$= -\int_{0}^{\infty} x^{k} d\overline{F}(x)$$

$$= -\left[x^{k} \overline{F}(x)\right]_{0}^{\infty} - \int_{0}^{\infty} k x^{k-1} \overline{F}(x) dx\right] \quad \therefore \text{ IBP}$$

$$= k \int_{0}^{\infty} x^{k-1} \overline{F}(x) dx$$

Example 5.1.1

Calculate $e_X(d)$ and $TVaR_p(X)$ for a Pareto distribution X with cdf

$$F(x) = 1 - \left(\frac{\theta}{x + \theta}\right)^{\alpha}, \quad x > 0,$$

where $\alpha > 1$ and $\theta > 0$.

Solution

Using • Proposition 2,

$$e_X(d) = \int_0^\infty P(X - d > x \mid X > d) \, dx = \int_0^\infty \frac{P(X - d > x, X > d)}{P(X > d)} \, dx$$

$$= \int_0^\infty \frac{P(X > x + d)}{P(X > d)} \, dx = \int_0^\infty \frac{\overline{F}(x + d)}{\overline{F}(d)} \, dx$$

$$= \int_0^\infty \left(\frac{d + \theta}{x + d + \theta} \right)^\alpha dx = \frac{(d + \theta)^\alpha}{1 - \alpha} \left(\frac{1}{x + d + \theta} \right)^{\alpha - 1} \Big|_0^\infty$$

$$= \frac{d + \theta}{\alpha - 1}$$

By Example 4.1.1, we have

$$\pi_p = \theta \left[(1-p)^{-rac{1}{lpha}} - 1
ight]$$

and so

$$\begin{split} \text{TVaR}_p(X) &= e_X(\pi_p) + \pi_p \\ &= \frac{\theta \left[(1-p)^{-\frac{1}{\alpha}} - 1 \right] + \theta}{\alpha - 1} + \theta \left[(1-p)^{-\frac{1}{\alpha}} - 1 \right] \\ &= \frac{\theta (1-p)^{-\frac{1}{\alpha}}}{\alpha - 1} + \frac{\theta (\alpha - 1)(1-p)^{-\frac{1}{\alpha}}}{\alpha - 1} - \theta \\ &= \frac{\theta \alpha (1-p)^{-\frac{1}{\alpha}}}{\alpha - 1} - \theta \end{split}$$

• Proposition 3 (Expected Deductible)

We have

$$E([X-d]_+) = \int_d^\infty \bar{F}(x) \, dx$$

Proof

By the Law of the Unconscious Statistician and IBP on the last step,

$$E([X-d]_{+}) = \int_{d}^{\infty} (x-d) \, dF(x) = -\int_{d}^{\infty} (x-d) \, d\bar{F}(x) = \int_{d}^{\infty} \bar{F}(x) \, dx$$

• Proposition 4 (An Expression for Mean Excess Value)

If
$$ar{F}(d)>0$$
, we have
$$e_X(d)=rac{\int_d^\infty ar{F}(x)\,dx}{ar{F}(d)}$$

Proof

Observe that by **b** Proposition 3, we have

$$e_X(d) = E[X - d \mid X > d] = \frac{E[(X - d)\mathbb{1}_{X > d}]}{P(X > d)}$$
$$= \frac{E([X - d]_+)}{\bar{F}(d)} = \frac{\int_d^\infty \bar{F}(x) \, dx}{\bar{F}(d)}$$

5.2 Severity Distributions - Creating Severity Distributions

Definition 23 (Severity Distribution)

A severity distribution is a distribution used to describe single random losses in an insurance portfolio.

When a loss occurs, the full amount of the loss is not necessarily the amount paid by the insurer, since an insurance policy typically involves some form of adjustment (e.g. deductible, limit, coinsurance). A distinction needs to be made between the actual loss prior to any of the adjustments (aka ground-up loss) and the amount ultimately paid by the insurer.

Our goal is to find a reasonable model for the **ground-up loss** rv *X*. The following are two desirable properties for *X*:

- $Im(X) = \mathbb{R}_{>0}$, since losses are positive;
- pf of *X* is right-skewed, since we want the "tail" of the distribution to be not heavy.
 - The motivation for this property is due to the 20-80 rule: 20% of the largest claims account for 80% of the total claim amount.

THERE ARE two approaches to constructing a severity distribution:

- Parametric approach²: specify a "form" for the distribution with a finite number of parameters.
- Nonparametric approach: no form is specified; the distribution is constructed directly from the empirical data.

A weakness of the Nonparametric approach is, if there is not enough data, such as in catasthropic risks, is becomes difficult to obtain reliable information. We shall look at one such example in this approach.

² This approach shall be the focus of this course.

Definition 24 (Empirical Distribution Function)

Let $\{X_1,...,X_n\}$ be an iid sample of a risk X. Then its empricial distribution function (edf) is defined as

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{X_i \le x\}}, \quad x \in \mathbb{R}.$$

Remark

Simply put, the edf assigns a probability of $\frac{1}{n}$ to each sample point X_i .

Example 5.2.1

Consider a random sample of a risk with size 5: {30, 80, 150, 150, 200}. Find the edf of the risk.

Solution

The edf is given by

$$F_n(x) = \frac{1}{5} \sum_{i=1}^{5} \mathbb{1}_{\{X_i \le x_i\}} = \begin{cases} 0 & x < 30 \\ \frac{1}{5} & 30 \le x < 80 \\ \frac{2}{5} & 80 \le x < 150 \\ \frac{4}{5} & 150 \le x < 200 \\ 1 & x \ge 200 \end{cases}$$

A.1 Individual Risk Model: An Alternate View

This appendix serves to explain why our note of $Z_i = I_i X_i$ is wrong with as mush rigour as we can go for now. There may be hand-wavy parts, but those will be indicated.

We mentioned, as shown by Klugman, Panjer and Willmot (2012)¹, that for the Individual Risk Model, the aggregate claim is modeled by

$$S = \sum_{i=1}^{n} Z_i$$

where Z_i is a random variable for the potential loss of the i^{th} insurance policy, while n is fixed. It is claimed that we can also express each Z_i as

$$Z_i = I_i X_i$$

where I_i is an indicator function given by

$$I_i(x) = egin{cases} 1 & ext{if a claim occurs} \ 0 & ext{if there are no claims} \end{cases}$$
 ,

while X_i is the size of the claim(s) for the i^{th} policy provided that there is a claim.

ONE PROBLEM that arises is: are X_i and I_i independent? They should be if we wish to define Z_i in such a way. In fact, according to Klugman et. al. in page 177,

Let
$$X_j = I_j B_j$$
, where $I_1, ..., I_n, B_1, ..., B_n$ are independent.

where X_j is our Z_i , I_j is our I_i , and B_j is our X_i .

¹ Klugman, S. A., Panjer, H. H., and Willmot, G. E. (2012). Loss Models: From Data to Decisions. John Wiley & Sons, Inc., 4th edition

§ Z_i is not well-defined Let us be explicit about the definitions of I_i and X_i ; we have

$$I_i = \mathbb{1}_{\{Z_i > 0\}}$$
$$X_i = Z_i \mid Z_i > 0$$

However, we observe that such a defintion of X_i is undefined on $Z_i = 0$. So the equation

$$Z_i = I_i X_i$$

is note well-defined.

§ *Independence of I_i and X_i* We cannot actually tell if I_i and X_i are independent from each other, as it is equivalent to comparing apples with oranges². Recall from our earlier courses, in particular STAT₃₃₀, of the following notion:

² In fact, I think this analogy fits our case perfectly so.

Definition (Probability Space)

Let Ω be a sample space, and \mathcal{F} a σ -algebra defined on Ω^3 . A **probability space** is the measurable space (Ω, \mathcal{F}) with a probability measure, $f: \mathcal{F} \to [0,1]$, defined on the space. We denote a probability space as (Ω, \mathcal{F}, f) .

³ Note that (Ω, \mathcal{F}) is called a **measurable space**.

As mentioned in an earlier \S , X_i is not defined on $Z_i = 0$, while I_i is defined on $Z_i = 0$. So the sample space for X_i and I_i are not the same, and so their probability measures are not the same as well. Therefore, it is meaningless to ask if X_i and I_i are independent.

⁴ This statement is hand-wavy.

Our best attempt at fixing this is probably the following: let

$$Z_i = \sum_{i=1}^{I_i} X_i,$$

which we can then have X_i to be independent from I_i . However, interestingly so, this is a very similar approach to a Collective Risk Model.

A.2 Coherent Risk Measure

An excerpt from Klugman et. al. (2012)⁵:

⁵ Klugman, S. A., Panjer, H. H., and Willmot, G. E. (2012). *Loss Models: From Data to Decisions*. John Wiley & Sons, Inc., 4th edition The study of risk measures and their properties has been carried out by authors such as Wang. Specific desirable properties of risk measures were proposed as axioms in connection with risk pricing by Wang, Young, and Panjer and more generally in risk measurement by Artzer et. al. The Artzner paper introduced the concept of coherence and is considered to be the groundbreaking paper in risk measurement.

Often, we use the function $\rho(X)$ to denote risk measures. One may think of $\rho(X)$ as the amount of assets required to protect against adverse outcomes of the risk X.

Definition 25 (Coherent Risk Measure)

A coherent risk measure is a risk measure $\rho(X)$ that has the following *four properties for any two loss rvs X and Y:*

- 1. (Subadditivity) $\rho(X+Y) \leq \rho(X) + \rho(Y)$.
- 2. (Monotonicity) If $X \leq Y$ for all possible outcomes, then $\rho(X) \leq$ $\rho(Y)$.
- 3. (Positive homogeneity) $\forall c \in \mathbb{R}_{>0}$, $\rho(cX) = c\rho(X)$.
- 4. (Translation invariance) $\forall c \in \mathbb{R}_{>0}$, $\rho(X+c) = \rho(X) + c$

Interpretation of the conditions

Subadditivity

- the risk measure (and in return, the capital required to cover for it) for two risks combined will not be greater than for the risks to be treated separately;
- reflects the fact that there shuld be some diversification benefit from combining risks;
- this requirement is disputed: e.g. the merger of several small companies into a larger one exposes each of the small companies to the reputational risks of the others.

Monotonicity

 if one risk always has greater losses than the other under all circumstances⁶, then the risk measure of the greater risk should always be greater than the other.

 6 Probabilistically, this means P(X > Y) = 0

• Positive homogeneity

- the risk measure is independent of the currency used to measure it;
- doubling the exposure to a particular risk requires double the capital, which is sensible as doubling provides no diversification.

• Translation invariance

- there is no additional risk for an additional risk which has no additional uncertainty.

Bibliography

Dhaene, J. and Vyncke, D. (2010). The individual risk model. https://www.researchgate.net/publication/228232062_The_Individual_Risk_Model.

Klugman, S. A., Panjer, H. H., and Willmot, G. E. (2012). *Loss Models: From Data to Decisions*. John Wiley & Sons, Inc., 4th edition.

List of Symbols and Abbreviations

rv random variable

drv discrete random variable crv continuous random variable

pf probability function

pmf probability mass function

pdf probability density functionmgf moment generating function

pgf probability generating function

Index

aggregate claim amount, 9 Aggregate Loss, 12 Aggrement Payment, 12

Central Moment, 22 Coherent Risk Measure, 39 Collective Risk Model, 11 Continuous Random Variable, 15

Deductibles, 20 Discrete Random Variable, 13 Dutch Principle, 26

Empirical Distribution Function, 35 Expectation Principle, 26

failure rate, 17 Frequency Distribution, 12 ground-up loss, 34

Hazard Rate Function, 17

Individual Risk Model, 10 insurance pricing, 25

Liability Risk, 9

Mean Excess Loss, 31 median, 23

Mixed Random Variable, 19

moment, 14

Moment Generating Function, 16

percentile, 23 Premium Principle, 25 Probability Generating Function, 14 quantile risk measure, 26 Quantiles, 23

Random Variable, 13 Risk Measure, 25

Severity Distribution, 11, 34 standard deviation, 22 Standard Deviation Principle, 26

Tail-Value-at-Risk, 28, 29

Value-at-Risk, 26 VaR, 26 variance, 22