ADVANCED METHODS OF NON-LIFE INSURANCE

Practical Information and Introduction

Prof. dr. Tim Verdonck

Tim.Verdonck@kuleuven.be

2018 - 2019

ALMA MATER STUDIORUM A.D. 1088
UNIVERSITÀ DI BOLOGNA

Outline

Content

Periodo di svolgimento delle lezioni: 19/09/2018 - 25/10/2018		
GIORNO	ORA	
19 mercoledì settembre, 2018	08:00 - 13:00	ADVANCED METHODS OF INSURANCE 2 TIM VERDONCK AULA 4 Piano Terra Piazza Antonino Scaravilli, 1-2 - Bologna
20 giovedì settembre, 2018	08:00 - 13:00	ADVANCED METHODS OF INSURANCE 2 TIM VERDONCK AULA SEMINARI DSE 2 Piano Primo Piazza Antonino Scaravilli, 1-2 - Bologna
21 venerdì settembre, 2018	08:00 - 12:00	ADVANCED METHODS OF INSURANCE 2 TIM VERDONCK AULA II Piano Terra Via delle Belle Arti, 41 - Bologna
22 lunedì ottobre, 2018	08:00 - 12:00	ADVANCED METHODS OF INSURANCE 2 TIM VERDONCK AULA 2 Piano Terra Piazza Antonino Scaravilli, 1-2 - Bologna
23 martedì ottobre, 2018	08:00 - 12:00	ADVANCED METHODS OF INSURANCE 2 TIM VERDONCK AULA 5 Piano Terra Piazza Antonino Scaravilli, 1-2 - Bologna
24 mercoledì ottobre, 2018	08:00 - 12:00	ADVANCED METHODS OF INSURANCE 2 TIM VERDONCK AULA 2 Piano Terra Piazza Antonino Scaravilli, 1-2 - Bologna
25 giovedì ottobre, 2018	08:00 - 12:00	ADVANCED METHODS OF INSURANCE 2 TIM VERDONCK AULA 2 Piano Terra Piazza Antonino Scaravilli, 1-2 - Bologna

What is an actuary?

- From Wikipedia:

  "An actuary is a business professional who deals with the measurement and management of risk and uncertainty. Actuaries provide assessments of financial security systems, with a focus on their complexity, their mathematics, and their mechanisms.

  Actuaries of the 21st century require analytical skills, business knowledge, and an understanding of human behavior and information systems to design and manage programs that control risk.

  Actuaries use skills primarily in mathematics, particularly calculus-based probability and mathematical statistics, but also economics, computer science, finance, and business.

  Actuaries assemble and analyze data to estimate the probability and likely cost of the occurrence of an event such as death, sickness, injury, disability, or loss of property."

- Check The Actuary Song:

  `https://www.youtube.com/watch?v=PZ8fqn7M1oU`

Often considered best job

Actuary consistently ranks among top jobs in the United States. It is even regularly rated as the best job in the world.

Society of Actuaries
September 11 at 3:00 PM ·
Like Page

Actuarial science is one of the most valuable college majors in the U.S. according to Bankrate. Read the recent article with the Chicago Tribune. https://news.google.com/…/CAIiEEOwYFpxjmJ78wxtlvB0r4EqFwgEK…

CHICAGOTRIBUNE.COM
The most valuable college major? Actuarial science
A Bankrate.com report released on Monday ranked actuarial science —…

Society of Actuaries
41 mins ·

Actuary ranks as a top job in business, based on the U.S. News report.
http://ow.ly/AokK50ilhZP
See the list http://ow.ly/1ZZt50ilhZQ
Learn more about the actuarial profession
https://www.soa.org/future-actuaries/what-is-an-actuary/

BEST JOBS U.S.News RANKINGS

MONEY.USNEWS.COM
Best Business Jobs of 2018
These are the best jobs in America.

Some famous actuaries :-)

Content of this course

- Practical Information
- Introduction to Non-Life Insurance (and to statistical software R)
- Copulas
- Aggregate Loss Modeling
  We build further on the material seen in *Actuarial Mathematics* of Prof. Mulinacci
- Copulas
- Generalized Linear Models (GLMs)
- Premium Calculation Principles
- Extreme Value Theory (if we have enough time)

Study material

What are the preliminaries?

- Linear regression
- *Actuarial Mathematics* of Prof. Mulinacci
  - probability theory: random variables, cdf, pdf, mgf, (conditional) expectation and variance, etc;
  - statistics: descriptive statistics, maximum likelihood, method of moments;
  - counting and claim size distributions;
  - compound distributions.

What is the study material?

- **Slides contain all necessary material!**
- **Main references** (click on the links):
  (1) Mario Wüthrich (2013-2017), *Lecture notes on non–life insurance mathematics*;
  (2) Rob Kaas et al. (2008), *Modern Actuarial Risk Theory (using R)*;
  (3) Stuart Klugman et al. (2012), *Loss models: from data to decisions*;
  (4) Boland (2007). *Statistical and probabilistic methods in actuarial science*.
- Interesting background reading material (articles, websites and other books) can be found in slides and/or lecture notes (Wüthrich).

Evaluation

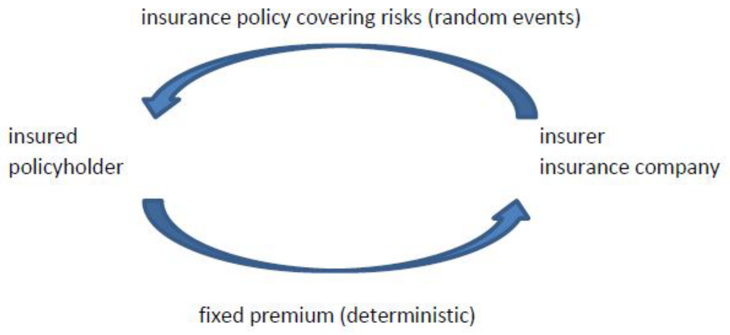
- Small assignment in R.
- Exam, which may contain
  - Exercises!
  - Theoretical questions (proof seen in class, maybe slightly adapted).
  - Questions testing insight in material
  - R code and R output that needs to be analyzed (used for answer).
- Extensive summary (formulary) may be used on exam
  - This version will be distributed to students (online). Make sure you bring it to the exam!
  - Only this summary is allowed (NO books). Solutions to exercises or other slides can NOT be used on the exam.
  - Do not write notes on this summary.

Outline

1. Practical Information
2. Introduction non-life insurance

Introduction

Insurance industry exists because people are willing to pay a price for being insured against unforeseeable random effects.



Introduction

Basic features

- every member faces similar risks
- individual members profit from **diversification benefits** in the form of a law of large numbers that applies to community.

Non-life insurance
(in US: Property and casualty insurance and in UK: General insurance)

- fire insurance, flood insurance, earthquake insurance
- property insurance. homeowner's insurance
- car insurance, motor insurance
- liability (or casualty) insurance
- accident insurance, health insurance
- credit insurance
- legal protection insurance
- traveling insurance
- ...

(⇒everything that is NOT life insurance)

Introduction

Example: Motor or car insurance
Large body of non-life actuarial literature is devoted to this line of business (LoB).
Large data bases with policyholders' characteristics and claim histories are maintained by insurance companies.
It can cover some or all of the following items

- insured party (bodile injury, fire, theft)
- insured vehicle (property damage)
- third parties (car and people, property damage and bodile injury)

Examples of types of coverage

- comprehensive (all risks)
- third party, fire and theft
- third party only: protection if vehicle owner causes harm to another party but no cover whatsoever for personal injury or damage to your vehicle.
- first party coverage: protection if vehicle owner is responsible; protects himself and his property

Third party coverage is required in most countries

Introduction

Non-life insurance policy is **contract** among two parties, insurer and insured

- typically an insurance period is specified
- provides the insurer with a fixed/deterministic amount of money (premium)
- provides the insured with a financial coverage against the random occurrence of **well-specified events**
- the right of the insured to these amounts constitutes a claim by the insured on the insurer
- the amount which the insurer is obligated to pay (in respect of a claim) is the claim amount or loss amount (the payments which make up this claim are claims payments).

Non-life insurance

- Idea of insurance is part of our civilised world: it is based on **mutual trust** of insurer and the insured. This mutual trust must be based on science and in 20th century **necessary tools (based on probability and statistics)** for dealing with matters of insurance were developed.
- In 1903 Swedish actuary Lundberg laid foundations of modern **risk theory**, which is synonym for non-life insurance mathematics. It deals with the **modeling of claims** that arrive in an insurance business and it gives advice on **how much premium** has to be charged in order to avoid bankruptcy (ruin) of the insurance company.
- In 1930 Cramer wrote that the object of the theory or risk is to give a **mathematical analysis of the random fluctuations** in an insurance business and to discuss the various means of protection against their inconvenient effects.
- In our modern world, individuals and companies continually encounter situations of risk where decisions must be made in the face of **uncertainty**.
- Risk theory can be useful in analysing possible scenarios as well as options open to the analyst, and therefore assist in the ultimate decision-making process.

Modeling approach

Many problems in (actuarial) science involve the building of a **mathematical model** (e.g. to forecast or predict future costs).

- Model is simplified mathematical description that is constructed based on knowledge and **experience** of actuary, combined with **data from the past**.
- Data guide actuary in **selecting the form** of the model.
- Data guide actuary in calibrating unknown quantities, called **parameters**.
- Data helps actuary in **testing the performance** of the model.
- Model provides balance between **simplicity** (e.g. number of parameter) and **conformity** (discrepancy between data and model) to data.

Modeling approach

We distinguish following stages in modeling process [Klugman et al.]

1. **Model choice**: select one or more models based on analyst's prior knowledge and experience and on nature and form of available data.
2. **Model calibration**: calibrate model on available data.
3. **Model validation**: fitted model is validated to determine if it adequately conforms the data.
4. Should **other models** be considered? (if former stage fails)
5. **Model selection**: compare all valid models using some criteria to select between them. Model averaging is sometimes done.
6. Modify for **future**, e.g. adjust parameters.

Repeat these 6 stages when new data become available.

Risk components and premium elements

Insurance contracts involve many different risky components

- **Pure randomness**: outcomes of claims Y_i are uncertain.
- **Model risk** *All models are wrong, but some are useful - Box*
  - model world does not provide appropriate description of real world behaviour
  - parameters in chosen model are misspecified
  - risk factors change over time so that past observations do not appropriately describe what may happen in the future.

These uncertainties ask for risk loading/margin beyond pure risk premium $\mu = E(Y_i)$. Insureds are willing to pay more, i.e.

insurance premium is μ	$+$	**risk margin to protect against risks above**
	$+$	profit margin and sales commissions to agents
	$+$	administrative expenses
	$+$	state taxes
	$-$	some financial gains on investments

Risk components and premium elements

- **Setting the price** of an insurance good can be a *perplexing problem*
  - In manufacturing, the cost of a good is (relatively) known.
  - In other areas of financial services, market prices are available.
  - In many lines of insurance, start with an **expected cost**, **add margins** to account for the product's riskiness, expenses incurred in servicing the product, and a profit/surplus allowance for the insurance company.
- For some lines of business (e.g. automobile and homeowners insurance) **analytics** sharpens the market calculating good's expectation **more precise**.
- Traditionally, insurers use **information** reported by policyholders on application forms, combined with selected external sources.
- There is interest in collecting more information about policyholders.
  - **Ethically** permissible? These debates are important.
  - From a **statistician's viewpoint** these additional sources have proven to be significant from hypothesis testing, predictive and economic viewpoints.
- Policyholders are also agreeing to let insurers gather **more data** about them for risk classification purposes.
  - Best example is the GPS and cameras that are mounting in cars for monitoring how much and how a policyholder drives (telematics).

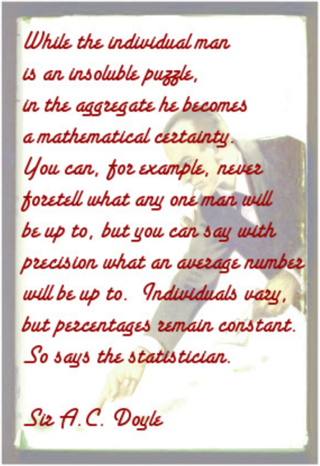
Probability theory and statistics

Consider portfolio of insurance risks covered by insurance policies issued by an insurance company. To determine this premium, the insurance company **pools similar risks** whose individual insurance claims can be described by a sequence **random variables** $Y_1, \ldots, Y_n$ where $n \in \mathbb{N}$ (most useful means of representing uncertainty is through probability).

- Probability space $(\Omega, \mathcal{F}, \mathbb{P})$
- $Y_1, \ldots, Y_n$ uncorrelated and identically distributed
- Finite mean $\mu = E(Y_1) < \infty$

Total claims paid by insurance company on all policies is sum of all payments made by insurer. Therefore it is useful to determine properties of $Y_1 + \ldots + Y_n$. First results are obtained with **Law of Large Numbers** (LLN) and **Central Limit Theorem** (CLT), but let us first briefly revise some facts/definitions/notations from probability theory and the course *Actuarial Mathematics* of Prof. Mulinacci.

Probability theory and statistics



Expectation

The distribution of a random variable X contains all of the probabilistic information about X. The entire distribution of X, however, is often too cumbersome for presenting his information. Summaries of the distribution, such as the average value, or expected value, can be useful for giving people an idea of where we expect X to be without trying to describe the entire distribution. Although the mean of a distribution is useful summary, it does not convey very much information about the distribution. It might be useful to give some measure of spread out of the distribution of X.

Originated from problem of points and problem of dice and attributed to Dutch mathematician Christiaan Huygens, but he says in foreword of his treatise (*De ratiociniis in ludo aleae*, 1957) that also for some time some of the best mathematicians of France (e.g. series of letters Pascal - Fermat - Chevalier de Méré) have occupied themselves with this kind of calculus so that no one should attribute to him the honour of the first invention.

Expectation

Definition

Assume rv $X \sim F$ with probability mass function or probability density function (pdf) f and $h : \mathbb{R} \to \mathbb{R}$ a sufficiently nice measurable function. The expected value of $h(X)$ is

$$E[h(X)] = \int_{\mathbb{R}} h(x) dF(x) = \begin{cases} \sum_{x \in \mathcal{A}} h(x) f(x) & X \text{ discrete} \\ \int_{-\infty}^{\infty} h(x) f(x) dx & X \text{ continuous} \end{cases}$$

provided the right side converges absolutely.

The middle term uses the general framework of the Riemann-Stieltjes integral $\int_{\mathbb{R}} h dF$. Important choices for function h:

- Mean, expectation or first moment of $X \sim F$

$$\mu_X = E[X] = \int_{\mathbb{R}} x dF(x)$$

- Moment of order k : measure of variation around 0

$$\alpha_k = E(X^k) = \begin{cases} \sum_x x^k f(x) & X \text{ discrete} \\ \int_{-\infty}^{\infty} x^k f(x) dx & X \text{ continuous} \end{cases}$$

If $E[X^k]$ exists ($E[|X|^k] < \infty$), all lower order moments $E[X^l]$, $l \leq k$ exist.

- Central moment order k: measure of variation around μ

$$\mu_k = E\left[(X-\mu)^k\right] = \begin{cases} \sum_x (x-\mu)^k f(x) & X \text{ discrete} \\ \int_{-\infty}^{\infty} (x-\mu)^k f(x) dx & X \text{ continuous} \end{cases}$$

- Variance: second central moment

$$\sigma_X^2 = \text{Var}(X) = E\left[(X-\mu)^2\right]$$

and its square root σ is called standard deviation.

- Coefficient of variation is ratio of σ to μ

$$\text{Vco}(X) = \frac{\sigma_X}{\mu_X}$$

- Skewness: if positive (negative), then long right (left) tail

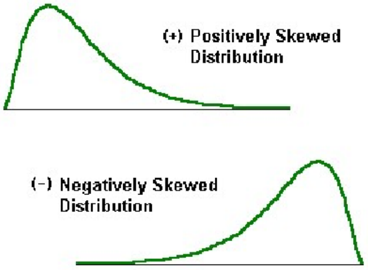
$$\zeta_X = \frac{\mu_3}{\sigma^3} = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = E[Z^3].$$

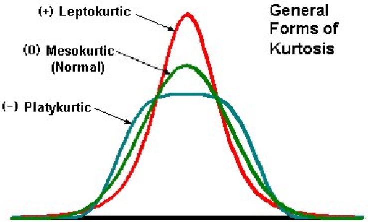
  $Z = \frac{X-\mu}{\sigma}$ is the so-called standardized rv.

- (Excess) Kurtosis: measure of peakedness (or thickness in the tails)

$$\gamma_X = \frac{\mu_4}{\sigma_4} = E[Z^4] - 3.$$

  -3 sets $\gamma_X = 0$ for normal distribution





Insurance premium

To determine this premium, the insurance company pools similar risks whose individual insurance claims can be described by a sequence rv's $Y_1, \ldots, Y_n$

- Probability space $(\Omega, \mathcal{F}, \mathbb{P})$
- $Y_1, \ldots, Y_n$ uncorrelated and identically distributed
- Finite mean $\mu = E(Y_1)$

We can apply weak law of large numbers (LLN):
[Jakob Bernouilli (*Ars Conjectandi*, 1713)]

$$\forall \varepsilon > 0 \lim_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Y_i - \mu\right| \geq \varepsilon\right) = 0 \quad \text{or} \quad \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow{P} \mu.$$

$\Rightarrow$ Total claim amount becomes **more predictable** with increasing portfolio size n
An insurance enterprise exists because of its ability to **pool risk**: by insuring many people, the individual risks are combined into an **aggregate risk** that is manageable and can be priced at a level that will attract clients.

Insurance premium

LLN can be further refined under following assumptions

- Probability space $(\Omega, \mathcal{F}, \mathbb{P})$
- $Y_1, \ldots, Y_n$ **independent** and identically distributed (iid)
- Finite mean $\mu = E(Y_1)$ and **Finite variance** $\sigma^2 = \mathrm{Var}(Y_1)$

Central limit theorem (CLT) provides asymptotic limit distribution (convergence in distribution) as $n \to \infty$:

[De Moivre (1733)-Laplace (1812); Markov-Chebyshev-Lyapunov; Lindeberg-Feller-Lévy-Polya]

$$\frac{\sum_{i=1}^n Y_i - n\mu}{\sqrt{n}\sigma} \xrightarrow{D} Z \sim \mathcal{N}(0,1) \quad \text{or} \quad P\left[\left(\frac{\sum_{i=1}^n Y_i - n\mu}{\sqrt{n}\sigma}\right) \le x\right] \xrightarrow{D} \Phi(x).$$

$\Rightarrow$ total claim amount of the portfolio becomes **predictable** in the limit because the relative confidence bounds get narrower the bigger the portfolio is.

Practically, sums of rv's can often be **approximated** by those from normal distribution

$$\sum_{i=1}^n Y_i \approx N\left(n\mu, n\sigma^2\right) \quad \text{or} \quad \overline{Y} \approx N\left(\mu, \frac{1}{n}\sigma^2\right)$$

Moment generating function (mgf)

Moment generating function (mgf) of X

$$M(r) = M_X(r) = E(e^{rX}) = \int_{\mathbb{R}} e^{rx} dF(x)$$

- An alternative for the mgf is the characteristic function

  $$\varphi_X(r) = E[e^{irX}].$$

  Always exists, uniquely determines distribution and provides moment formula. There is also connection (Lévy - Cramér Theorem) between convergence in distribution of a sequence of rv's and convergence of the corresponding characteristic functions.

- If mgf $M_X(r)$ exists, it is crucial to identify properties of X.
  - provides a moment formula
  - it **uniquely determines the distribution**!
  - connection with convergence in distribution.

Moment generating function (mgf)

Lemma 1

Choose $X \sim F$ and assume that there exists $r_0 > 0$ such that $\forall r \in (-r_0, r_0) : M_X(r) < \infty$, then $M_X(r)$ has power series expansion for $r \in (-r_0, r_0)$ with

$$M_X(r) = \sum_{k \geq 0} \frac{r^k}{k!} E[X^k]$$

Inversely,

Property of mgf

The derivatives at the origin are given by

$$\frac{d^k}{dr^k} M_X(r)|_{r=0} = E(X^r) < \infty$$

This explains the name:
given $M_X(r)$, we can **generate the moments** of X and vice versa.

Properties mgf

Properties Moment Generating Function

Let X, Y be **independent** rv's with mgf's $M_X(t)$ and $M_Y(t)$ and a real number.

1. $M_{a+bX}(r) = e^{ar} M_X(br)$
2. $M_{X+Y}(r) = M_X(r) M_Y(r)$

Proof

1. $M_{a+bX}(r) = E[e^{r(a+bX)}] = E[e^{ar+brX}] = E[e^{ar}e^{brX}] = e^{ar}E[e^{brX}]$
2. $M_{X+Y}(r) = E[e^{r(X+Y)}] = E[e^{rX}e^{rY}] \overset{\text{ind.}}{=} E[e^{rX}]E[e^{rY}]$

Therefore, if we let $S = Y_1 + \ldots + Y_n$ with Y_i's **independent** rv's, then

$$M_S(r) = \prod_{i=1}^{n} M_{Y_i}(r)$$

provided all component mgf's exist.

Exercise 1

For given $X \sim \mathcal{N}(\mu, \sigma^2)$ with $\mu \in \mathbb{R}, \sigma^2 > 0$, we have following pdf

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad x \in \mathbb{R}$$

1. Given $Z \sim \mathcal{N}(0,1)$, prove that

$$M_Z(r) = e^{r^2/2} < \infty \qquad r \in \mathbb{R}$$

2. Given $X \sim \mathcal{N}(\mu, \sigma^2)$, hence $X = \mu + \sigma Z$, prove that

$$M_X(r) = e^{r\mu + r^2\sigma^2/2} < \infty \qquad r \in \mathbb{R}$$

3. Calculate mean and variance for normal distribution $X \sim \mathcal{N}(\mu, \sigma^2)$ using mgf
4. $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ independent $\Rightarrow \sum_i X_i \sim \mathcal{N}\left(\sum_i \mu_i, \sum_i \sigma_i^2\right)$

Answer of exercise 1

1. For $Z \sim N(0,1)$ we have

$$\begin{aligned} M_Z(r) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{rx} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x^2-2rx+r^2)}{2}+\frac{r^2}{2}} dx \\ &= e^{\frac{r^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-r)^2} d(x-r) = e^{\frac{r^2}{2}} \end{aligned}$$

2. Since $X = \sigma Z + \mu$ it follows $M_X(r) = e^{\mu r} M_Z(\sigma r) = e^{\mu r + \frac{\sigma^2 r^2}{2}}$
3.

$$\begin{aligned} \mu_X &= E[X] = \frac{d}{dr} M_X(r)|_{r=0} = e^{r\mu+\frac{1}{2}r^2\sigma^2}(\mu + r\sigma^2)|_{r=0} = \mu \\ E[X^2] &= \frac{d^2}{dr^2} M_X(r)|_{r=0} = e^{r\mu+\frac{1}{2}r^2\sigma^2}\left((\mu + r\sigma^2)^2 + \sigma^2\right)|_{r=0} = \mu^2 + \sigma^2 \\ \sigma_X^2 &= \mathrm{Var}(X) = E[X^2] - E[X]^2 = \sigma^2 \end{aligned}$$

4. Since X_i are independent

$$M_{\sum_i X_i}(r) = \prod_i M_{X_i}(r) = e^{r\mu_1 + r^2\sigma_1^2/2} e^{r\mu_2 + r^2\sigma_2^2/2} \ldots = e^{r\sum_i \mu_i + r^2 \sum_i \sigma_i^2/2}$$

Exercise 2

A company insures homes in three different cities. For a particular period, denote the amounts of losses incurred by each city by Y_1, Y_2 and Y_3 respectively. Since there is sufficient distance that separates the cities, it is reasonable to assume that these losses occurring in these cities are independent. The moment generating functions for the distributions of the losses of the cities are

$$\begin{aligned} M_{Y_1}(r) &= (1-3r)^{-3} \\ M_{Y_2}(r) &= (1-3r)^{-4} \\ M_{Y_3}(r) &= (1-3r)^{-5} \end{aligned}$$

Suppose $S = Y_1 + Y_2 + Y_3$ is the combined losses from the cities.
Calculate $E(S^3)$.

Answer of exercise 2

By independence

$$\begin{aligned}
M_S(r) &= M_{Y_1}(r)M_{Y_2}(r)M_{Y_3}(r) \\
&= (1-3r)^{-(3+4+5)} = (1-3r)^{-12} \\
M_S'(r) &= -12(1-3r)^{-13}(-3) = 36(1-3r)^{-13} \\
M_S''(r) &= 36(39)(1-3r)^{-14} \\
M_S'''(r) &= 36(39)(42)(1-3r)^{-15} \\
E(S^3) &= 36 * 39 * 42 = 58968
\end{aligned}$$

Cumulant generating function (cgf)

Cumulant generating function (cgf)

Cumulant generating function (cgf) is given by

$$\log M_X(r) = \log E[e^{rX}]$$

Assume that M_X is finite on $(-r_0, r_0)$ with $r_0 > 0$. We have

$$\begin{aligned}
\frac{d}{dr}\log M_X(r)|_{r=0} &= \frac{M_X'(r)}{M_X(r)}|_{r=0} = E[X] = \mu_X \\
\frac{d^2}{dr^2}\log M_X(r)|_{r=0} &= \frac{M_X''(r)M_X(r) - (M_X'(r))^2}{(M_X(r))^2}|_{r=0} = \mathrm{Var}(X) = \sigma_X^2 \\
\frac{d^3}{dr^3}\log M_X(r)|_{r=0} &= E\left[(X - E[X])^3\right] = \zeta_X \sigma_X^3
\end{aligned}$$

Lemma 2 (cumulant generating function)

Assume that M_X is finite on $(-r_0, r_0)$ with $r_0 > 0$. Then $\log M_X(.)$ is a convex function on $(-r_0, r_0)$.

Proof

$$\begin{aligned} \frac{d^2}{dr^2}\log M_X(r) &= \frac{M_X''(r)M_X(r) - (M_X'(r))^2}{(M_X(r))^2} = \frac{M_X''(r)}{M_X(r)} - \left(\frac{M_X'(r)}{M_X(r)}\right)^2 \\ &= \frac{E[X^2 e^{rX}]}{E[e^{rX}]} - \left(\frac{E[Xe^{rX}]}{E[e^{rX}]}\right)^2 \end{aligned}$$

Define new function F_r by $F_r(x) = \frac{1}{M_X(r)} \int_{-\infty}^{x} e^{ry}\, dF(y)$.
Observe that F_r is a distribution function. Thus we can choose rv $X_r \sim F_r$ with

$$0 \leq \mathrm{Var}(X_r) = E[X_r^2] - E[X_r]^2 = \frac{E[X^2 e^{rX}]}{E[e^{rX}]} - \left(\frac{E[Xe^{rX}]}{E[e^{rX}]}\right)^2 = \frac{d^2}{dr^2}\log M_X(r)$$

$\square$

Note: F_r gives Esscher measure of F.
Introduced by Bühlmann for new premium calculation principle.

Important formula and properties

Assume that $X \sim F$ is non-negative, $\mathbb{P}$-a.s. and has finite first moment. Then

$$E(X) = \int_0^\infty x dF(x) = \int_0^\infty [1 - F(x)] dx = \int_0^\infty \mathbb{P}[X > x] dx = \int_0^\infty \overline{F}(x) dx$$

Moreover, the second moment $E[X^2]$ (when it exists) is

$$E[X^2] = \int_0^{+\infty} x^2 f_X(x) dx = \int_0^{+\infty} 2x\overline{F}(x) dx$$

$\Rightarrow$ we can write expectations using **survival function**

$$\overline{F}(x) = 1 - F(X) = \mathbb{P}[X > x].$$

Survival functions are important for study of fatness of tails of distribution functions which plays crucial role for modeling of large claims.

Important formula and properties

Tower property or double expectation theorem [Williams, 1991]

For any sub-σ-algebra $\mathcal{G} \subset \mathcal{F}$ on our probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we have for any integrable rv $X \sim F$

$$E[X] = E[E[X|\mathcal{G}]]$$

In particular, if X and Y are rv's on $(\Omega, \mathcal{F}, \mathbb{P})$ we have

$$E[X] = E[E[X|Y]]$$

where $E[E[X|Y]]$ is an abbreviation for $E[X|\sigma(Y)]$ with $\sigma(Y) \subset \mathcal{F}$ denoting the σ-algebra generated by the random variable Y.
Assume that X is square integrable then the tower proper implies

$$\text{Var}(X) = E[\text{Var}(X|\mathcal{G})] + \text{Var}(E[X|\mathcal{G}]).$$

In some cases it is easier to find $E[X]$ and $\text{Var}(X)$ by **conditioning** on some other rv Y than by calculating it directly.

Exercise 3

A component in a system and a backup unit both have mean lifetime μ. If component fails then it is immediately replaced by backup, and there is probability p that backup works. Find expected total lifetime of the system.

Answer of exercise 3

A component in a system and a backup unit both have mean lifetime μ. If component fails then it is immediately replaced by backup, and there is probability p that backup works. Find expected total lifetime of the system.
Let T be the total lifetime of the system, and X the indicator for backup success ($X = 1$ if backup works, $X = 0$ if not). Then ...

$$
\begin{aligned}
E(T) &= E[E(T|X)] \\
&= E(T|X=0)\,P(X=0) + E(T|X=1)\,P(X=1) \\
&= \mu\,(1-p) + 2\mu\,p \\
&= (1+p)\mu.
\end{aligned}
$$

Notice the intuitive result: the mean total lifetime is $E(T) = \mu + p\mu$, that is the mean lifetime of the first component, plus the proportion of the backups that works multiplied by the mean lifetime of the backup.

Inverse of F or p-quantile of $X \sim F$

- F is right-continuous and non-decreasing.
- Generalized inverse of F for $p \in (0,1)$

$$F^{\leftarrow}(p) = inf\{x; F(x) \geq p\}$$

where $inf(\emptyset) = \infty$.

- $F^{\leftarrow}(p)$ is only tricky at places where F has discontinuity or where F is not strictly increasing.
- **Properties**
  1. $F^{\leftarrow}(p)$ is non-decreasing and left-continuous.
  2. F is continuous iff $F^{\leftarrow}(p)$ is strictly increasing.
  3. F is strictly increasing iff $F^{\leftarrow}(p)$ is continuous.
  4. (If F is right-continuous then) $F(x) \geq p$ iff $F^{\leftarrow}(p) \leq x$.
  5. $F^{\leftarrow}(F(x)) \leq x$.
  6. $F(F^{\leftarrow}(p)) \geq p$.
  7. If F is strictly increasing, then $F^{\leftarrow}(F(x)) = x$.
  8. If F is continuous, then $F(F^{\leftarrow}(p)) = p$.

Terminology in Statistics

- We need to predict outcome of rv $X \sim F$.
- This problem is solved by specifying appropriate **predictor** $\hat{X}$.
- Distribution function F often involves **unknown parameters**.
- These parameters are estimated using past experience and/or expert opinion.
- If we choose predictor $\hat{X} = \hat{\mu}_X$ for predicting X, then $\hat{\mu}_X$ serves at the same time as **estimator** for μ_X and as **predictor** for X.
- We obtain estimation error

$$\mu_X - \hat{\mu}_X$$

and prediction error

$$X - \hat{X} = X - \hat{\mu}_X = \underbrace{(X - \mu_X)}_{\text{process error}} + \underbrace{(\mu_X - \hat{\mu}_X)}_{\text{estimation error}}$$

Statistical tests

Statistical tests deal with problem of making decisions.

- Assume we have observation $\boldsymbol{x}$ of random vector $\boldsymbol{X} \sim F_\theta$ with given but unknown parameter θ which lies on given set Θ of possible parameters.
- **Aim**: test whether true unknown parameter θ that has generated $\boldsymbol{x}$ may belong to some subset $\Theta_0 \subset \Theta$.
- Simplest case: a singleton $\Theta_0 = \{\theta_0\}$.
- Check whether $\boldsymbol{x}$ may have been generated by given parameter θ_0.
  $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$ (two-sided).
- We build test statistic $T(\boldsymbol{X})$ whose distribution function is known under H_0 and we check whether $T(\boldsymbol{x})$ takes unlikely value under H_0.
  $\Rightarrow$ significance level $q \in (0,1)$ (typically 5% or 1%) and critical region C_q with $\mathbb{P}[T(\boldsymbol{X}) \in C_q] \leq q$.
  $\Rightarrow$ Reject H_0 if $T(\boldsymbol{x})$ falls into C_q.
- Calculate p-value, i.e. probability at which H_0 is just rejected (for one-sided unbounded intervals). For instance, if we choose a significance level of 5% and the resulting p-value of $T(\boldsymbol{x})$ is less or equal to 5% then the test rejects the null hypothesis on the 5% significance level.

Exercise 4

Employees in a large company have been offered group rates on both life and health insurance by an insurance company. Let X be the proportion of employees who will opt for the life scheme and let Y be the proportion of employees who go for the health scheme. Assume that experience suggests that an appropriate joint distribution for X and Y is given by

$$f_{X,Y}(x,y) = \begin{cases} \frac{2}{5}(x+4y) & 0 \leqslant x \leqslant 1; \quad 0 \leqslant y \leqslant 1 \\ 0 & \text{otherwise} \end{cases}$$

1. Determine the covariance of X and Y.
2. Determine the probability of at least 30% acceptance to the life scheme.
3. Determine $f_{Y|X}(y|x)$.
4. Determine the probability that, at most, 50% will opt for the health insurance given there is a 10% acceptance rate for the life scheme.
5. Determine $E[Y|X=0.2]$
6. Are X and Y independent ? (Explain very briefly.)

Answer of exercise 4

1. We first determine

$$f_X(x) = \int_0^1 \frac{2}{5}(x+4y)dy = \frac{2}{5}(x+2) \qquad \text{for } 0 \leqslant x \leqslant 1.$$

and

$$f_Y(y) = \int_0^1 \frac{2}{5}(x+4y)dx = \frac{2}{5}(1/2+4y) \qquad \text{for } 0 \leqslant y \leqslant 1.$$

This leads to

$$E[XY] = \int_0^1 \int_0^1 \frac{2}{5}xy(x+4y)\,dy\,dx = \frac{1}{3}.$$

$$E[X] = \int_0^1 \frac{2}{5}x(x+2)dx = \frac{8}{15}.$$

$$E[Y] = \int_0^1 \frac{2}{5}(1/2+4y)ydy = \frac{19}{30}.$$

Finally

$$\text{Cov}[X,Y] = E[XY] - E[X]E[Y] = \frac{1}{3} - \frac{16}{30}\frac{19}{30} = \frac{76}{675} \approx 0.1126.$$

2.
$$P(X > 0.3) = \int_{0.3}^{1} \frac{2}{5}(x+2)dx = 0.742$$

3.
$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{2/5(x+4y)}{2/5(x+2)} = \frac{x+4y}{x+2} \text{with} \quad 0 < x < 1, 0 < y < 1.$$

4. It follows that
$$\begin{aligned} P(Y \leqslant 0.5 \,|\, X = 0.1) &= \int_0^{0.5} f_{Y|0.1}(y|0.1)\, dy \\ &= \int_0^{0.5} \frac{0.1+4y}{2.1}\, dy \\ &= 0.261905 \end{aligned}$$

5.
$$E[Y|X = 0.2] = \int_0^1 y \frac{0.2+4y}{2.2} dy = 0.65$$

6. No, since $E[XY] \neq E[X]E[Y]$.