

Actuarial Formula Cheat Sheet

My Actuarial Revision Sheet

Master in Insurance Economics/Econometrics
Master in Actuarial Science

19 juillet 2025

Insurance Keywords

1 – Life and Non-Life Insurance

The distinction between life and non-life insurance is fundamental. An insurer cannot offer both types of insurance without holding two separate companies :

- **Life insurance**, i.e., personal insurance excluding coverage for bodily injuries,
- **Non-life insurance**, which includes property and liability insurance as well as insurance for bodily injuries.

2 – The Principles of Insurance

Insurance is assumed to :

- be based on utmost good faith,
- apply only if the insured has an insurable interest in preserving the item (property insurance),
- operate under the indemnity principle :
 - not allow enrichment from a claim settlement,
 - not even through insurance accumulation,
 - include subrogation (in Liability Insurance, if the insurer compensates the insured victim, the insured cannot then claim from the party responsible for the loss.)
- not reduce the insured's efforts in prevention and protection, as a reasonable person, even if financially protected.

establish causality in Liability Insurance — I am not liable if I did not contribute to the cause of the loss.

3 – The Insurance Policy

The **insurance policy** (or contract) is the contractual document that governs the relationship between the insurance company (or mutual insurance company) and the insured (policyholder). This contract defines in particular :

1. the list of covered events, including any exclusions,
2. the coverage, i.e., the assistance provided to the insured in case of a loss,
3. the obligations of the insured :
 - any preventive measures required to reduce risk,
 - time limits for reporting a claim to the insurer,
 - the amount and payment conditions of the premium (deductible, limit),
 - the conditions for cancellation of the policy (automatic renewal),
4. the obligations of the insurance company : time limits for compensation payments.

4 – The Premium and Claims

Classically, the role of the insurer is to substitute a constant C , the **contribution** or the **premium**, for a random claim S . La **pure premium** or **technical premium** aims to compensate claims without surplus or profit, overall $C_t = \mathbb{E}[S]$

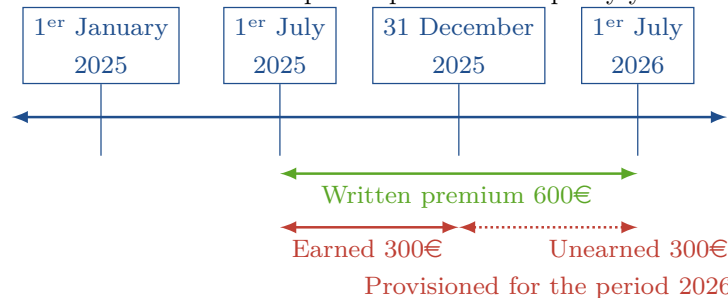
The **net premium** is higher than the pure premium. It aims to cover the cost of claims and provide a safety margin.

The **gross premium** is the net premium + overhead expenses + commissions + expected profit + taxes.

For commercial reasons, the gross premium actually charged may differ from the technical premium.

Written premium : premium charged to the insured to cover claims that may occur during the coverage period defined by the contract (generally 1 year in Property and Casualty insurance).

Earned premium : proportion of the written premium used to cover the risk over the exposure period of one policy year.



The S/P is the key indicator. For the insurer to make a profit the $S/P < 1$.

5 – Loss / Payment Triangle

Insurance accounting is broken down by the **accident year** of the claim. If a premium covers multiple calendar years, a proportional part will be allocated to each. Each payment and each claim provisioning is assigned to the accident year. The monitoring of payments or expenses is expressed through a triangle (triangular matrix) :

$$\begin{pmatrix} C_{1,1} & C_{1,2} & \dots & C_{1,n} \\ C_{2,1} & C_{2,2} & \dots & C_{2,n-1} \\ \vdots & \vdots & & \\ C_{n-1,1} & C_{n-1,2} & & \\ C_{n,1} & & & \end{pmatrix}$$

where $C_{i,j} = \sum_{k=1}^j X_{i,k}$ represents the cumulative amount of claims paid for origin year i and development year j .

6 – Solvency II and Risk Management

Solvency II is the European regulatory framework applicable to insurers and reinsurers since 2016. It is based on three interdependent pillars :

— Pillar 1 : Quantitative Requirements

Determines the capital requirements :

- **SCR** (Solvency Capital Requirement) : capital to absorb an extreme shock (99.5% over 1 year),
- **MCR** (Minimum Capital Requirement) : absolute minimum threshold,
- admissible assets to cover technical provisions and capital requirements.

— Pillar 2 : Governance, Internal Control, and Risk Management

The core link with **ERM** (*Enterprise Risk Management*). The requirements cover :

- governance : boards of directors responsible for the risk management framework ;
- an effective **internal control** system ;
- independent key functions : **actuarial, risk management, compliance, internal audit** ;
- **ORSA** (*Own Risk and Solvency Assessment*) : internal assessment of risks and solvency, a central tool aligning strategy, risk appetite, and economic capital.

— Pillar 3 : Market Discipline

Based on **transparency** and communication :

- **SFCR** (*Solvency and Financial Condition Report*) : public, summarizes solvency and financial position,

- **RSR** (*Regular Supervisory Report*) : intended for the supervisor,
- quantitative reporting : regulatory statements (**QRTs**), regular submission of financial and prudential data.

7 – Main Branches of Life and Non-Life Insurance

Life insurance covers long-term commitments, with or without a savings component :

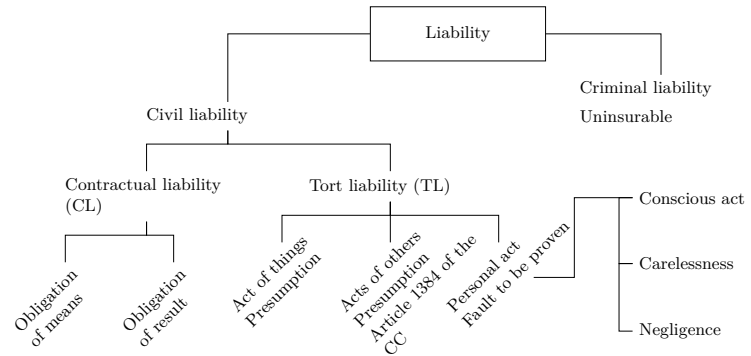
- **Life insurance** : lump sum or annuity paid if the insured is alive at a given date.
- **Death insurance** : payment if the insured dies during the covered period.
- **Endowment insurance** : combination of life and death coverage.
- **Life annuity** : periodic payments until death.
- **Savings/retirement** : products with deferred capital or deferred annuity.
- **Unit-linked policies** : benefits dependent on the value of financial assets.
- **Group contracts** : occupational pensions, group welfare insurance.

Non-life insurance covers risks occurring in the short or medium term :

- **Automobile** : third-party liability, vehicle damage.
- **Home** : fire, theft, water damage, liability.
- **General liability** : personal liability, business liability.
- **Health and welfare** : medical reimbursements, disability, incapacity.
- **Personal accident** : capital in case of accident, disability,

or death.

- **Business interruption** : financial losses related to a claim.
- **Technical risks** : construction, machinery breakdown.
- **Transport, aviation, maritime insurance** : goods in transit, specific liabilities.



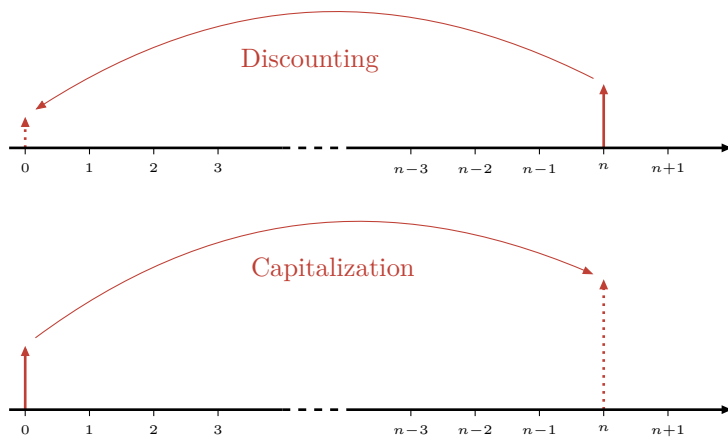
8 – Actuary

In practice, the actuary :

- prices insurance and welfare products,
- estimates technical provisions,
- projects cash flows and values long-term liabilities,
- measures economic capital (SCR, ORSA) and contributes to ERM,
- advises management on strategy, solvency, and regulatory compliance.

Financial Mathematics

9 – Capitalization Discounting



10 – Interests

Discount rate d

$$d = i / (1 + i)$$

Simple interest i

$$I_t = Pit = Pi \frac{k}{365}$$

Compound interest i

$$V_n = P(1 + i)^n = P \left(1 + \frac{p}{100} \right)^n$$

Continuous interest r

$$V_t = V_0 e^{rt}$$

Effective rate i_e

$$i_e = \left(1 + \frac{i}{m} \right)^m - 1$$

where i is the nominal rate and m the number of periods in a year.

Equivalent rate $i^{(m)}$

$$i^{(m)} = m(1 + i)^{1/m} - 1$$

Nominal rate i and periodic rate

The **nominal** or **face** rate allows calculating the interest due over one year. The **periodic** rate corresponds to the nominal rate divided by the number of periods in a year i/m . If the periodic rate is weekly, the nominal rate will be divided by 52.

11 – Present Value and Future Value

The present value (PV) represents the capital that must be invested today at an annual compound interest rate i to obtain future cash flows (F_k) at times t_k :

$$PV = \sum_{k=1}^n F_k \times \frac{1}{(1 + i)^{t_k}} \quad (1)$$

When the F_k are constant

$$PV = K \frac{1 - (1 + i)^{-n}}{i} \quad (2)$$

The future value (FV) represents the value of the capital at T which, with an annual compound interest rate i , capitalizes the

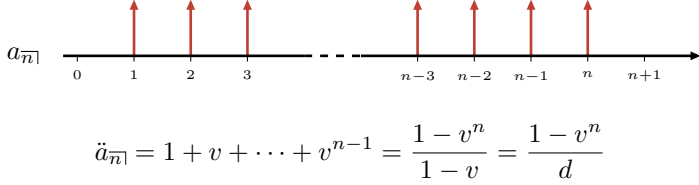
future cash flows (F_k) at times t_k .

$$FV = V_n = \sum_{k=1}^n F_k \times (1+i)^{n-t_k} \quad (3)$$

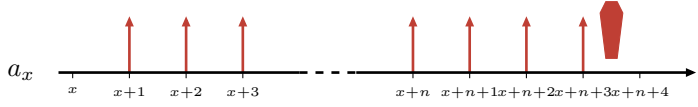
More generally $FV = (1+i)^n PV$.

12 – Annuities

Certain annuity $a_{\overline{n}|}$ (or $a_{\overline{n}|i}$ if the interest rate i needs to be specified) : this is the default case in financial mathematics. Its payments are, for example, guaranteed by a contract.



Contingent annuity \ddot{a}_x : its payments are conditional on a random event, such as a life annuity of an individual aged x . In this example, payments continue until death occurs :

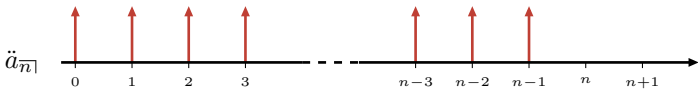


The date of death is represented here by a small coffin. This type of annuity will be extensively studied in the life actuarial section. Annuity in arrears (immediate) $a_{\overline{n}|}$: its periodic payments are made at the end of each payment period, as with a salary paid at the end of the month. This is the default case, previously illustrated for the certain annuity.

$$\ddot{a}_{\overline{n}|} = 1 + v + \dots + v^{n-1} = \frac{1-v^n}{1-v} = \frac{1-v^n}{d}$$

$$PV_{\overline{n}|}^{\text{due}} = K \ddot{a}_{\overline{n}|} = K \frac{1-v^n}{d}$$

Annuity in advance (due) $\ddot{a}_{\overline{n}|}$: its periodic payments are made at the beginning of each payment period, as with rent payments, for example.



Also denoted PV^{im} :

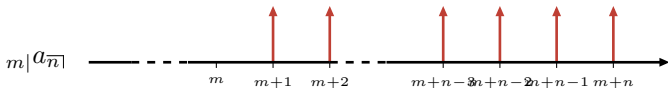
$$a_{\overline{n}|} = v + v^2 + \dots + v^n = \frac{1-v^n}{i} = v \frac{1-v^n}{1-v}$$

$$PV_{\overline{n}|}^{\text{im}} = K a_{\overline{n}|} = K \frac{1-v^n}{i}$$

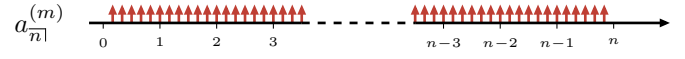
Perpetuity a or $a_{\overline{n\infty}|}$:

$$a = 1/i$$

Deferred annuity $m|a_{\overline{n}|}$: its payments do not start in the first period but after m periods, with m fixed in advance.



Periodic / monthly annuity $a^{(m)}$: the default periodicity is one year, but the unit payment can also be spread over m periods within the year.



If $i^{(m)}$ represents the equivalent nominal (annual) interest rate with m periods per year, then $i^{(m)} = m((1+i)^{1/m} - 1)$.

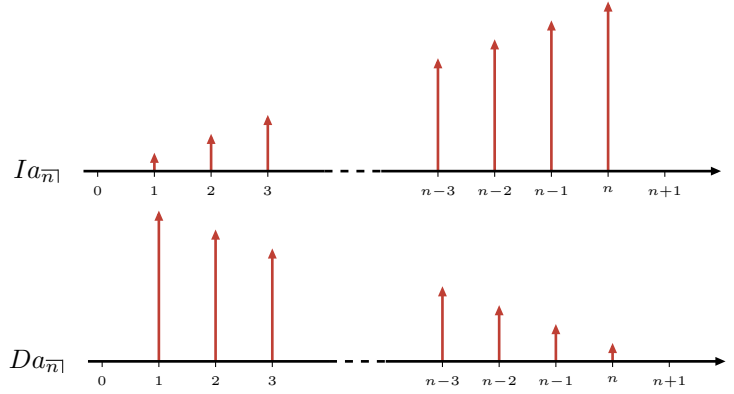
Similarly, $d^{(m)}$ is the nominal discount rate consistent with i and m : $d^{(m)} = m(1 - (1-d)^{1/m})$.

$$\ddot{a}_{\overline{n}|}^{(m)} = \frac{1}{m} \sum_{k=0}^{mn-1} v^{\frac{k}{m}} = \frac{d}{d^{(m)}} \ddot{a}_{\overline{n}|} = \frac{1-v^n}{d^{(m)}} \approx \ddot{a}_{\overline{n}|} + \frac{m-1}{2m} (1-v^n)$$

$$a_{\overline{n}|}^{(m)} = \frac{1}{m} \sum_{k=1}^{mn} v^{\frac{k}{m}} = \frac{i}{i^{(m)}} a_{\overline{n}|} = \frac{1-v^n}{i^{(m)}} \approx a_{\overline{n}|} - \frac{m-1}{2m} (1-v^n)$$

Unit annuity a : it is used when constructing annuity formulas. For a constant annuity, the total amount paid each year is 1, regardless of m .

Dynamic annuity, increasing/decreasing Ia/Da : in its simplest form, it pays an amount that starts at 1 (n) and increases (decreases) each period arithmetically or geometrically. In the following example, the progression is arithmetic. The prefix I (increasing) is used to indicate increasing annuities and D (decreasing) for decreasing annuities.



$$(I\ddot{a})_{\overline{n}|} = 1 + 2v + \dots + nv^{n-1} = \frac{1}{d} (\ddot{a}_{\overline{n}|} - nv^n) \quad (4)$$

with, we recall, $d = i/(1+i)$ and in arrears (immediate)

$$(Ia)_{\overline{n}|} = v + 2v^2 + \dots + nv^n = \frac{1}{i} (\ddot{a}_{\overline{n}|} - nv^n)$$

$$(D\ddot{a})_{\overline{n}|} = n + (n-1)v + \dots + v^{n-1} = \frac{1}{d} (n - a_{\overline{n}|})$$

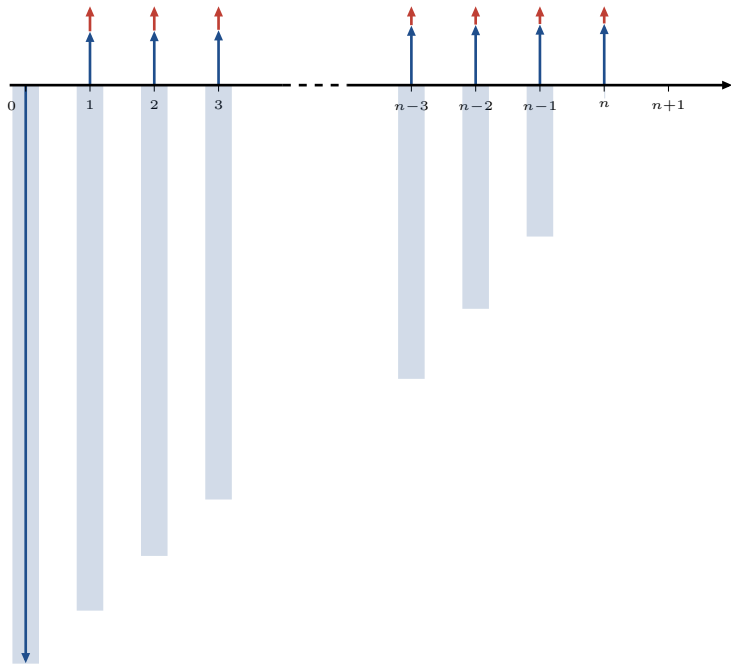
and in arrears :

$$(Da)_{\overline{n}|} = nv + (n-1)v^2 + \dots + v^n = \frac{1}{i} (n - a_{\overline{n}|})$$

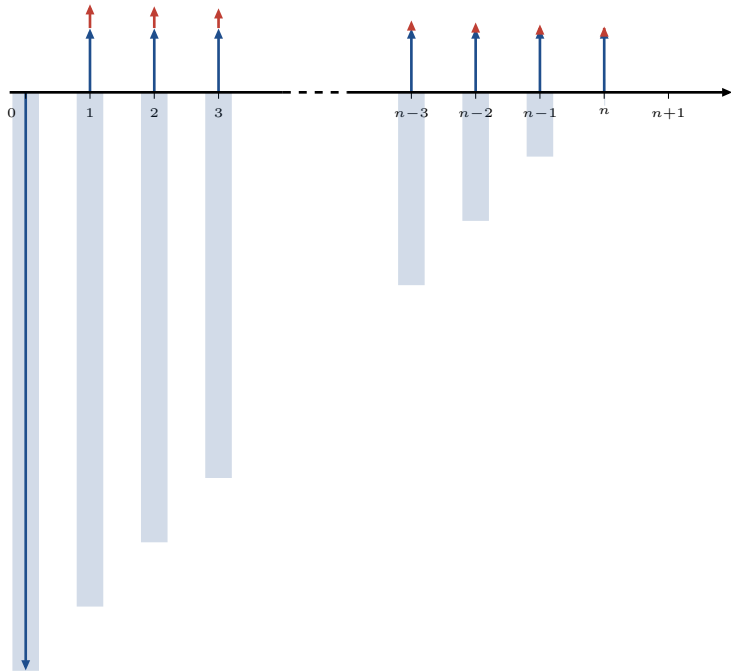
13 – The Loan (Indivisible)

The main property of the loan is to consider separately the interest from the repayment (or amortization).

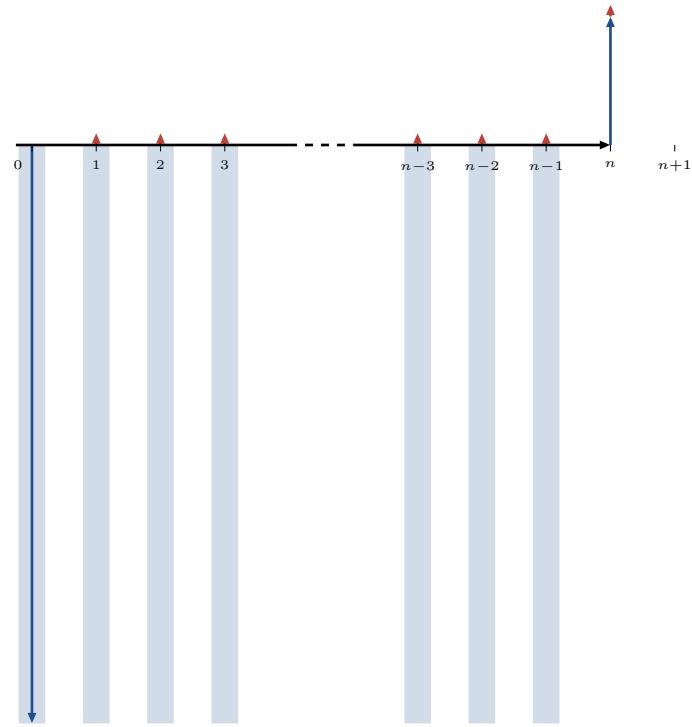
By constant repayment or constant annuity : the sum of the amortization and the interest at each period is constant.



By constant amortization.



By a bullet repayment, where the interest is constant. Only the interests are paid periodically until maturity, when the total repayment is made.



14 – Loan Amortization Schedule

	In fine	Constant amortizations	Constant annuities
Outstanding principal S_k	$T_k = S_0, T_n = 0$	$S_0 \left(1 - \frac{k}{n}\right)$	$S_0 \frac{1-v^{n-k}}{1-v^n}$
Interest U_k	$i \times S_0$	$S_0 \left(1 - \frac{k-1}{n}\right) i$	$K \left(1 - v^{n-k+1}\right)$
Amortizations T_k	$T_k = 0, T_n = S_0$	$\frac{S_0}{n}$	$K v^{n-k+1}$
Annuity K_k	$K_k = iS_0, K_n = (1+i)S_0$	$\frac{S_0}{n} (1-(n-k+1)i)$	$K = S_0 \frac{i}{1-v^n}$

15 – Market Functioning

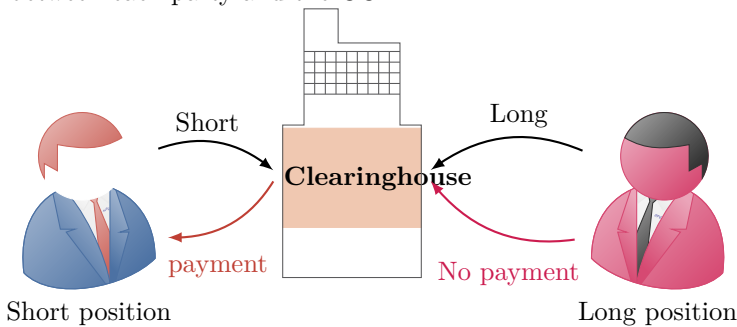
The **Exchange** – a place of exchange – enables, in fact, the physical meeting between capital demanders and suppliers. The main listings concern **equities**, **bonds** (Fixed Income), and **commodities**. Listed are **securities** such as stocks or bonds, **funds** (Exchange Traded Funds that replicate equity indices, ETC or ETN that replicate more specific indices or commodities, SICAV or FCP, subscription warrants), **futures contracts**, **options**, **swaps**, and **structured products**.

The **Financial Markets Authority (AMF)** oversees :

- the protection of invested savings ;
- the information of investors ;
- the proper functioning of the markets.

Euronext (including **Amsterdam**, **Brussels**, **Lisbon**, and **Paris**) is the main stock exchange in France. Its competitors include **Deutsche Börse** (which includes **Eurex**, **EEX**) in Europe, or **ICE** (which includes **NYSE (2012)**, **NYBOT (2005)**, **IPE (2001)**, **LIFFE**) and **CME Group** (including **CBOT**, **NYMEX**, **COMEX**) in the United States.

The **over-the-counter market (OTC, Over-The-Counter)** represents a major share of volumes traded outside organized markets. Since the Pittsburgh G20 (2009), certain standardized OTC derivatives must be cleared through a central entity. These **CCPs (Central Counterparties)** thus play the role of **clearinghouses** : they replace the bilateral contract with two contracts between each party and the CCP.



16 – The Money Market

Short-term interest-bearing securities, traded on money markets, are generally at **discounted interest**. Nominal rates are then annual and calculations use **proportional rates** to adjust for durations less than one year. These securities are quoted or valued according to the discount principle and with a Euro-30/360 calendar convention.

In the American market, public debt securities are called : Treasury Bills (T-bills) : $ZC < 1$ year, Treasury Notes (T-notes) : $ZC < 10$ years, Treasury Bonds (T-bonds) : coupon bonds with maturity > 10 years.

They are mainly :

- **BTF (fixed-rate Treasury bills, France)** : issued at 13, 26, 52 weeks, discounted rate, weekly auction, nominal 1 €, settlement at $T+2$.
- **Treasury bills > 1 year** : same rules as bonds (see next section).
- **Certificates of deposit (CDN)** : securities issued by banks at fixed/discounted rate (short term) or variable/post-discounted rate (long term), also called BMTN.

- **Eurodollars** : USD deposits outside the USA, formerly indexed on LIBOR, now declining.
- **Commercial paper** : unsecured short-term securities issued by large companies to finance their cash flow.

Price calculations of a fixed-rate Treasury bill with discounted interest

In the case of a discounted interest security according to the Euro-30/360 convention, the discount D is expressed as :

$$D = F \cdot d \cdot \frac{k}{360}$$

where F denotes the nominal value, d the annual discount rate used to value the discounted security, and k the maturity in days. If the discount rate d is known, then the price P is expressed as :

$$P = F - D = F \left(1 - d \cdot \frac{k}{360} \right)$$

Similarly, if the price P is known, then the discount rate d is derived as :

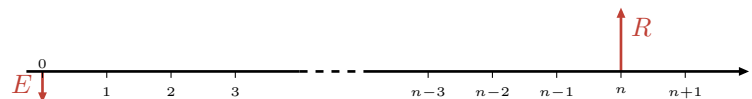
$$d = \frac{F - P}{F} \cdot \frac{360}{k}$$

The main Futures Contracts : Federal Funds Futures (US), Three-Month SOFR Futures (US), ESTR Futures (EU), SONIA Futures (UK), Euribor Futures (EU).

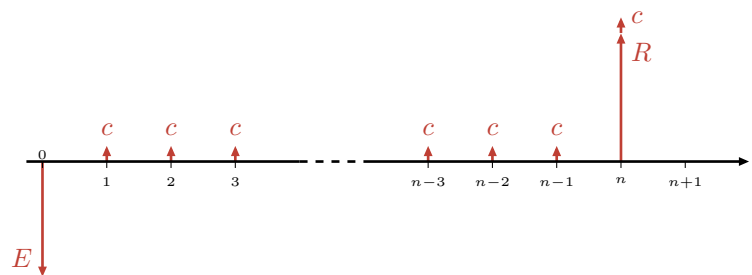
17 – Bond Market

Bonds are long-term debt securities in which the issuer (central or local government, bank, borrowing company) promises the bondholder (the lender) to pay interest (**coupons**) periodically and to repay the **nominal value** (or face value, or principal) at maturity. As mentioned in the previous section, **Treasury bills with a maturity greater than one year** will be treated as bonds with maturities under 5 years because their functioning is similar.

Zero-coupon bonds : pay only the nominal value at maturity. With E the issue price and R its redemption value :



Coupon bonds : Fixed-rate bonds have a coupon rate that remains constant until maturity. Assuming a *bullet* repayment, with E the issue price, c the coupons, and R the redemption value, it can be illustrated as follows :



Indexed bonds (inflation-linked bonds) have coupons and sometimes also the nominal value indexed to inflation or another economic indicator, such as Treasury Assimilable Bonds indexed to inflation (OATi). The values of c vary.

Bonds with **floating rate**, **variable rate**, or **resettable rate** : have a coupon rate linked to a reference interest rate (for example, the euro short- term rate (**€STR**)).

Perpetual bonds have no maturity date ; the principal is never repaid.

A distinction is often made between government bonds (Treasury bonds) and corporate bonds issued by private companies.

A bond is mainly defined by a **nominal value** F (Face Value), the **nominal rate** i , its duration or **maturity** n . In the default case, the bondholder lends the amount $E = F$ at issuance at time 0, receives each year a coupon $c = i \times F$, and at n , the principal or capital $R = F$ is returned. When $E = F$, the issue is said to be at par, and when $R = F$, the redemption is said to be at par.

The price of a bond is determined by the present value of the expected future cash flows (coupons and principal repayment) discounted at the market yield rate r .

The price calculation of bonds simply relies on the present value formula :

$$VP = \sum_{k=1}^n \frac{c}{(1+r)^k} + \frac{F}{(1+i)^n}$$

where :

- PV : price or present value of the bond,
- r : market interest rate for the relevant maturity.

For bonds with periodic coupons, the coupon is divided by the number of periods (m) per year and the formula becomes :

$$PV = \sum_{k=1}^{mn} \frac{c/m}{(1+r^{(m)})^k} + \frac{R}{(1+r^{(m)})^{mn}}$$

where c/m represents the periodic coupon payment and $r^{(m)}$ the periodic interest rate.

The bond yield is the value $r^{(m)}$, the equivalent rate of r over m periods in the year, which equates the present value VP_r with the current or market price of this bond.

The quotation of a bond is given as a percentage. Thus, a quotation of 97.9 on Euronext indicates a quoted value of $97.9/100 \times F$. It is quoted excluding **accrued coupons**, the portion of the next coupon to which the seller is entitled if the bond is sold before the payment of that coupon.

18 – Duration & Convexity

The Macaulay duration :

$$D = \sum_{t=1}^n t \cdot w_t, \quad \text{où } w_t = \frac{PV(C_t)}{P}.$$

If the payment frequency is k per year, the duration expressed in years is obtained by dividing by k . The modified duration D^* :

$$D^* = \frac{D}{1+i}.$$

Which allows approximating the portfolio change ΔP in case of interest rate changes Δ_i

$$\Delta P \approx -P D^* \Delta_i$$

Similarly, the convexity

$$C = \frac{1}{P(i)} \times \frac{d^2 P(i)}{di^2},$$

which allows refining the approximation of ΔP

$$P(i + \Delta_i) \approx P(i) \left(1 - D^* \Delta_i + \frac{1}{2} C (\Delta_i)^2 \right).$$

19 – CAPM

Capital Asset Pricing Model :

$$E(r_i) = r_f + \beta_i (E(r_m) - r_f)$$

- $E(r_i)$ is the expected return of asset i ,
- r_f is the risk-free rate,
- $E(r_m)$ is the expected market return,
- β_i is the sensitivity coefficient of asset i with respect to market variations.

The coefficient β_i measures the volatility of asset i relative to the overall market.

20 – Derivatives Market

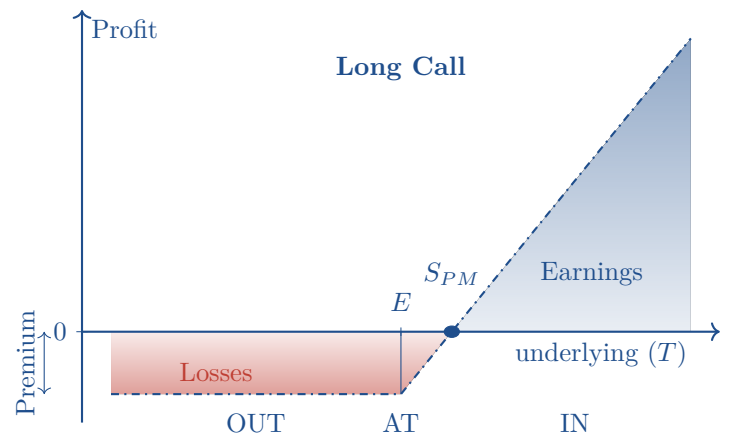
A **derivative contract** (or contingent asset) is a financial instrument whose value depends on an underlying asset or variable. Options are part of derivative contracts.

An **option** is a contract that gives the right (without obligation) to buy (call) or sell (put) an underlying asset at a fixed price (strike price) at a future date, in exchange for the payment of a premium. The buyer (long position) pays the premium ; the seller (short position) receives it. **European option** (exercise possible only at maturity) and **American option** (exercise possible at any time until maturity).

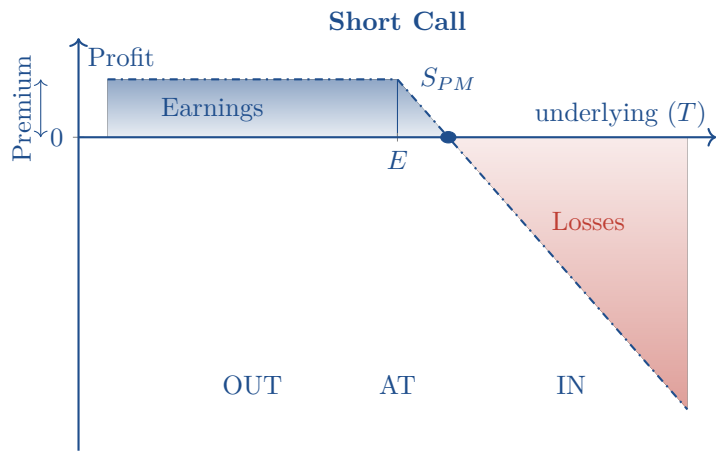
Options listed on stocks are called *stock options*.

21 – Simple Strategies

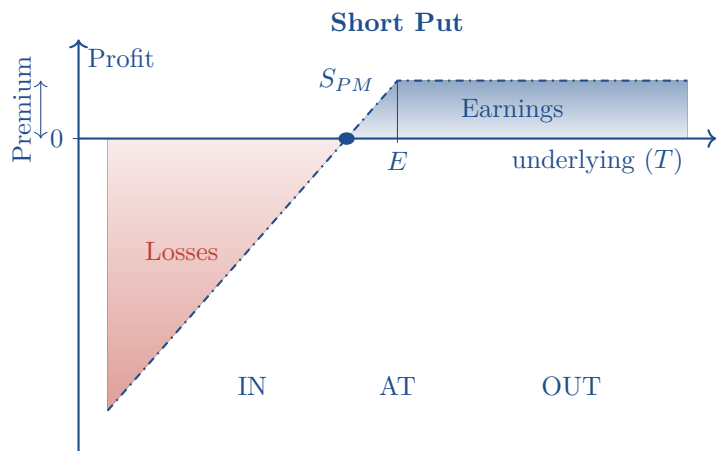
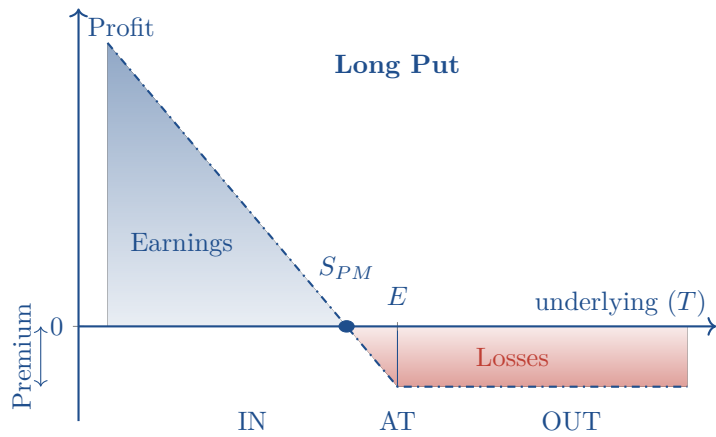
With T the maturity, K the strike price, S or S_T the underlying at maturity, the payoff is $\max(0, S_T - K) = (S_T - K)^+$. Letting C be the premium, the profit realized is $\max(0, S_T - K) - C$, with a profit if $(S_T > V_{PM} = K + C)$ (PM stands for **break-even point**).



At maturity, the payoff is $\min(0, K - S_T) = -\max(0, S_T - K) = -(S_T - K)^+$ and the profit realized is $C - \max(0, S_T - K)$.



At maturity, the payoff is $\max(0, K - S_T) = (K - S_T)^+$. Letting P be the put premium, the profit realized is $\max(0, K - S_T) - P$, positive if $V_{PM} = K - P < S_T$.



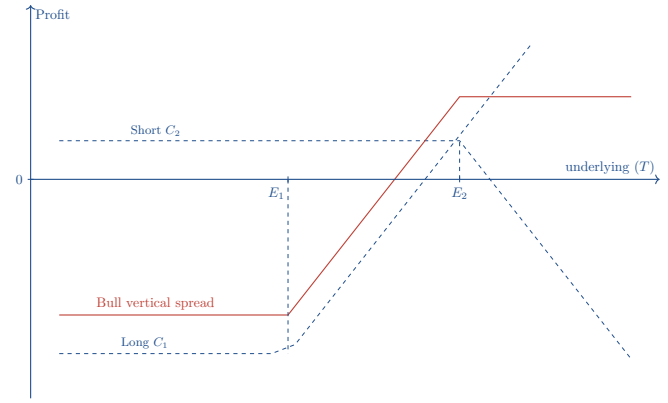
At maturity, the payoff is $\min(0, S_T - K) = -(K - S_T)^+$.

22 – Spread Strategies

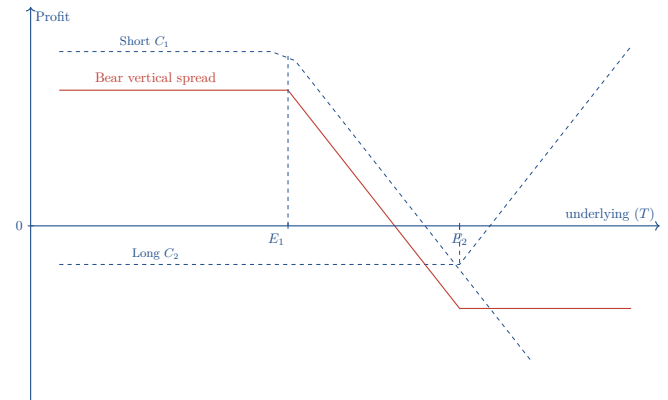
Spread strategy : uses two or more options of the same type (two call options or two put options). If the strike prices vary, it is a **vertical spread**. If the maturities change, it is a **horizontal spread**.

A vertical spread strategy (*spread trading strategy*) involves a long position and a short position on call options on the same underlying asset, with the same maturity but different strike prices. We distinguish : **bull vertical spread** (*Bull spread*) and **bear vertical spread** (*Bear spread*).

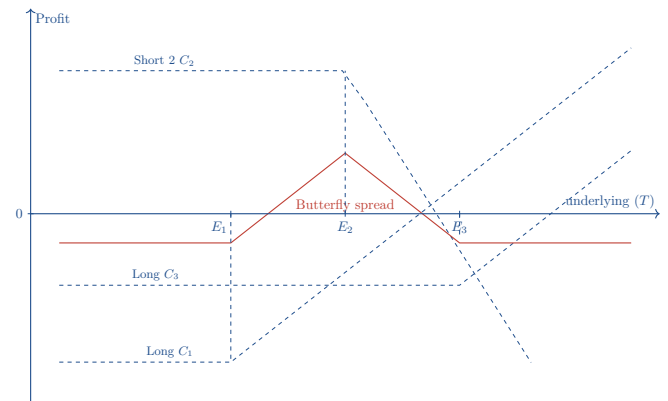
Bull vertical spread : anticipating a moderate rise in the underlying asset, the investor takes a long position on C_1 and a short position on C_2 under the condition $E_1 < E_2$. Net result at maturity :



Bear vertical spread : anticipating a moderate decline in the underlying asset, the investor sells the more expensive option and buys the cheaper one.



Butterfly spread : anticipates a small movement in the underlying asset. It is a combination of a bull vertical spread and a bear vertical spread. This strategy is suitable when large movements are considered unlikely. Requires a low initial investment.

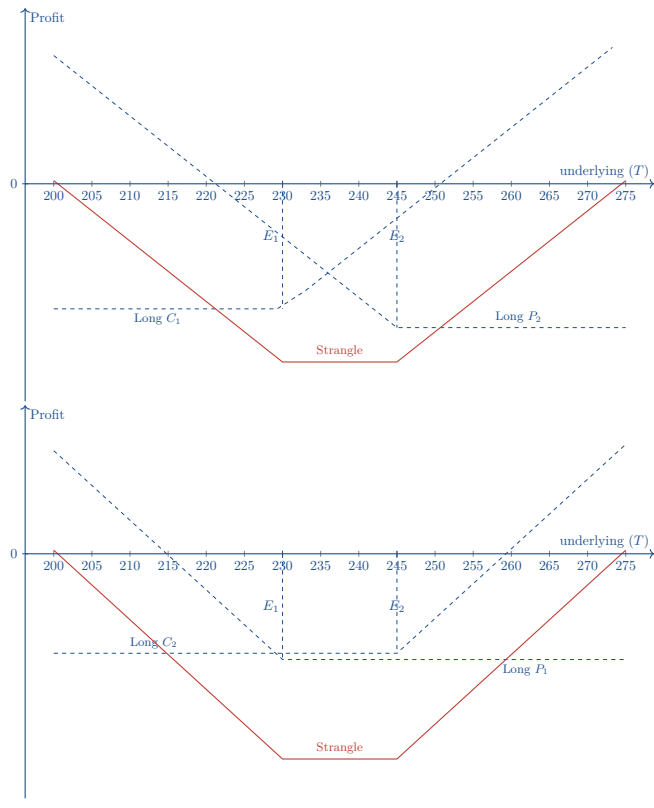


23 – Combined strategies

A **combined strategy** uses both call and put options. Notably, we distinguish between **straddles** and **strangles**.

A **straddle** combines the purchase of a call option and a put option with the same expiration date and strike price. This strategy bets on a large price movement, either upward or downward. The maximum loss occurs if the price at expiration is equal to the strike price.

A **strangle** is the purchase of a call and a put with the same expiration date but different strike prices. It assumes a very large movement in the value of the underlying asset.



24 – Absence of arbitrage opportunity

No risk-free profit is possible by exploiting price differences.

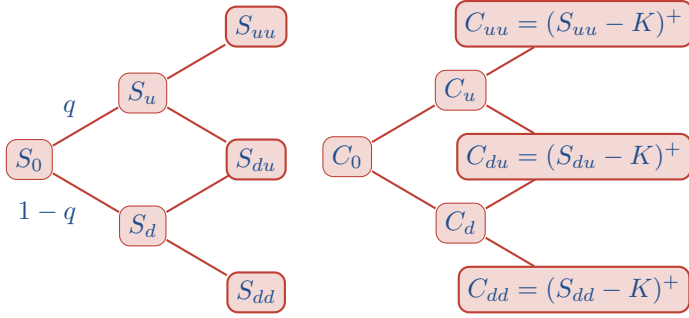
25 – Parity relation

AAO implies the following relationship between the Call and the Put :

$$S_t - C_t + P_t = Ke^{-i_f \cdot \tau}$$

26 – The Cox-Ross-Rubinstein model

It is based on a discrete-time process with two possible price movements at each period : an increase (factor u) or a decrease (factor d), with $u > 1 + i_f$ and $d < 1 + i_f$. The price at $t = 1$ is then $S_1^u = S_0 u$ or $S_1^d = S_0 d$, according to a probability q or $1 - q$.



This model extends to n periods with $n + 1$ possible prices for S_T . At expiration, the value of a call option is given by $C_1^u = (S_1^u - K)^+$ and $C_1^d = (S_1^d - K)^+$.

Absence of arbitrage opportunity implies

$$d < 1 + i_f < u$$

and a risk-neutral probability

$$q = \frac{(1 + i_f) - d}{u - d}$$

Call price (with $S_1^d < K < S_1^u$) :

$$C_0 = \frac{1}{1 + i_f} [qC_1^u + (1 - q)C_1^d]$$

We can also construct a replication portfolio composed of Δ shares and B bonds, such that :

$$\begin{cases} \Delta = \frac{S_1^u - K}{S_1^u - S_1^d}, \\ B = \frac{-S_1^d}{1 + i_f} \cdot \Delta \end{cases} \Rightarrow \Pi_0 = \Delta S_0 + B$$

Put price :

$$P_0 = \frac{1}{1 + i_f} [qP_1^u + (1 - q)P_1^d]$$

Determination of q , u , d : By calibrating the model to match the first moments of the return under the risk-neutral probability (expected value i_f , variance $\sigma^2 \delta t$), we obtain :

$$e^{i_f \delta t} = qu + (1 - q)d, \quad qu^2 + (1 - q)d^2 - [qu + (1 - q)d]^2 = \sigma^2 \delta t$$

With the constraint $u = \frac{1}{d}$, we obtain :

$$\begin{aligned} q &= \frac{e^{-i_f \delta t} - d}{u - d} \\ u &= e^{\sigma \sqrt{\delta t}} \\ d &= e^{-\sigma \sqrt{\delta t}} \end{aligned}$$

27 – The Black & Scholes Model

Assumptions of the model

- The risk-free rate R is constant. We define $i_f = \ln(1 + R)$, which implies $(1 + R)^t = e^{i_f t}$.
- The stock price S_t follows a geometric Brownian motion :

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

$$S_t = S_0 \exp \left(\sigma W_t + \left(\mu - \frac{1}{2} \sigma^2 \right) t \right)$$

- No dividend during the option's lifetime.
- The option is "European" (exercised only at maturity).
- Frictionless market : no taxes or transaction costs.
- Short selling is allowed.

The Black-Scholes-Merton equation for valuing a derivative contract f is :

$$\frac{\partial f}{\partial t} + i_f S \frac{\partial f}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = i_f f$$

At maturity, the price of a call option is $C(S, T) = \max(0, S_T - K)$, and that of a put option is $P(S, T) = \max(0, K - S_T)$.

Determinants	call	put
Underlying price	+	-
Strike price	-	+
Maturity (or time)	+	+
Volatility	+	+
Short-term interest rates	+	-
Dividend payment	-	+

The analytical solutions are :

$$\begin{aligned} C_t &= S_t \Phi(d_1) - Ke^{-i_f \tau} \Phi(d_2) \\ P_t &= Ke^{-i_f \tau} \Phi(-d_2) - S_t \Phi(-d_1) \end{aligned}$$

or :

$$d_1 = \frac{\ln(S_t/K) + (i_f + \frac{1}{2} \sigma^2) \tau}{\sigma \sqrt{\tau}}, \quad d_2 = d_1 - \sigma \sqrt{\tau}$$

- **Delta** Δ : variation in the option price depending on the underlying.
- **Gamma** Γ : delta sensitivity.
- **Thêta** Θ : sensitivity to time.
- **Véga** \mathcal{V} : sensitivity to volatility.
- **Rho** ρ : interest rate sensitivity.

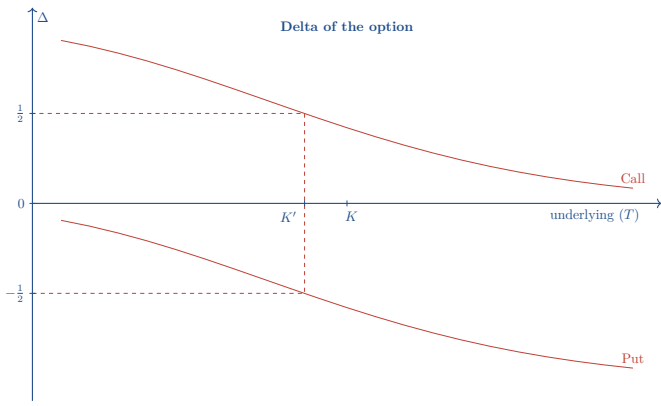
The **Delta** measures the impact of a change in the underlying asset :

$$\Delta_C = \frac{\partial C}{\partial S} = \Phi(d_1), \quad \Delta \in (0, 1)$$

$$\Delta_P = \frac{\partial P}{\partial S} = \Phi(d_1) - 1, \quad \Delta \in (-1, 0)$$

The global Delta of a portfolio Π with weights ω_i is :

$$\frac{\partial \Pi}{\partial S_t} = \sum_{i=1}^n \omega_i \Delta_i$$



28 – The Yield Curve

The **yield curve**, or the curve of returns, or $r_f(\tau)$, provides a graphical representation of risk-free interest rates as a function of maturity (or term). It is also called the **zero-coupon** yield curve, referring to a type of risk-free bond with no coupons (a debt composed only of two opposite cash flows, one at t_0 and the other at T). This curve also provides insight into market expectations regarding future interest rates (**forward rates**).

29 – The Nelson-Siegel and Svensson models

The **Nelson-Siegel** functions take the form

$$y(m) = \beta_0 + \beta_1 \frac{[1 - \exp(-m/\tau)]}{m/\tau} + \beta_2 \left(\frac{[1 - \exp(-m/\tau)]}{m/\tau} - \exp(-m/\tau) \right)$$

where $y(m)$ and m are as above, and β_0 , β_1 , β_2 , and τ are parameters :

- β_0 is interpreted as the long-term level of interest rates (the coefficient is 1, it is a constant that does not decrease),
- β_1 is the short-term component, noting that :

$$\lim_{m \rightarrow 0} \frac{[1 - \exp(-m/\tau)]}{m/\tau} = 1$$

It follows that the overnight rate such as €str will equal $\beta_0 + \beta_1$ in this model.

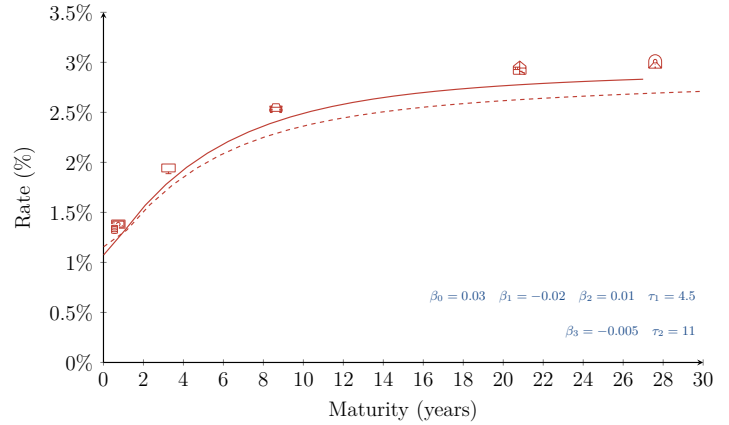
- β_2 is the medium-term component (it starts at 0, increases, then decreases back toward zero — i.e., bell-shaped),
- τ is the scale factor on maturity ; it determines where the term weighted by β_2 reaches its maximum.

Svensson (1995) adds a second bell-shaped term ; this is the Nelson-Siegel-Svensson model. The additional term is :

$$+ \beta_3 \left(\frac{[1 - \exp(-m/\tau_2)]}{m/\tau_2} - \exp(-m/\tau_2) \right)$$

and the interpretation is the same as for β_2 and τ above ; it allows for two inflection points on the yield curve.

These Nelson-Siegel and Svensson functions have the advantage of behaving well in the long term and being easy to parameterize. They are illustrated in the figure where the pictograms 🏠 🚗 🏡 represent the different usual maturities for this type of property or investment. They allow for the modeling of a broad yield curve. Once adjusted, the user can then evaluate assets or define various sensitivity measures.



30 – Vasicek model

Under a risk-neutral probability \mathbb{Q} , the short rate (r_t) follows an Ornstein-Uhlenbeck process with constant coefficients :

$$dr_t = \kappa(\theta - r_t) dt + \sigma dW_t, \quad r_0 \in \mathbb{R}$$

où :

- $\kappa > 0$ is the speed of mean reversion,
- θ is the long-term mean level,
- $\sigma > 0$ is the volatility,
- W_t is a standard Brownian motion under \mathbb{Q} .

The EDS solution (application of Itô's lemma to $Y_t = r(t)e^{\kappa t}$) :

$$r_t = r_s e^{-\kappa(t-s)} + \theta(1 - e^{-\kappa(t-s)}) + \sigma \int_s^t e^{-\kappa(t-u)} dW_u$$

So :

$$\mathbb{E}_{\mathbb{Q}}[r_t | \mathcal{F}_s] = r_s e^{-\kappa(t-s)} + \theta(1 - e^{-\kappa(t-s)})$$

$$\text{Var}_{\mathbb{Q}}[r_t | \mathcal{F}_s] = \frac{\sigma^2}{2\kappa} (1 - e^{-2\kappa(t-s)})$$

The process (r_t) is Gaussian ; negative rates are possible.

31 – Price of a zero-coupon bond (Vasicek)

The price at time t of a zero-coupon bond maturing at T is given by :

$$ZC(t, T) = A(t, T) e^{-B(t, T) r_t}$$

où :

$$B(t, T) = \frac{1 - e^{-\kappa(T-t)}}{\kappa}$$

$$A(t, T) = \exp \left[\left(\theta - \frac{\sigma^2}{2\kappa^2} \right) (B(t, T) - (T - t)) - \frac{\sigma^2}{4\kappa} B(t, T)^2 \right]$$

This formulation is possible due to the fact that $\int_t^T r_s ds$ is a Gaussian random variable conditional on \mathcal{F}_t .

$$ZC(t, T) = \mathbb{E}_{\mathbb{Q}} \left[\exp \left(- \int_t^T r_s ds \right) \middle| \mathcal{F}_t \right]$$

32 – Cox–Ingersoll–Ross (CIR) model

Under the risk-neutral measure \mathbb{Q} , the short rate (r_t) follows the dynamics :

$$dr_t = \kappa(\theta - r_t) dt + \sigma \sqrt{r_t} dW_t, \quad r_0 \geq 0$$

with :

- $\kappa > 0$: mean reversion speed,
- $\theta > 0$: long-term level,
- $\sigma > 0$: volatility,
- W_t : Brownian motion under \mathbb{Q} .

So :

- The square root $\sqrt{r_t}$ guarantees $r_t \geq 0$ if $2\kappa\theta \geq \sigma^2$ (Feller condition).
- The process (r_t) is a non-Gaussian diffusion process but with continuous trajectories.
- The rate is **mean-reverting** around θ .

Thus, the process (r_t) is a diffusion with explicit conditional distributions (under \mathbb{Q}) :

For $s < t$, the variable r_t follows a non-central χ^2 distribution :

$$r_t \mid \mathcal{F}_s \sim c \cdot \chi_d^2(\lambda)$$

with :

- $c = \frac{\sigma^2(1 - e^{-\kappa(t-s)})}{4\kappa}$
- $d = \frac{4\kappa\theta}{\sigma^2}$: degrees of freedom
- $\lambda = \frac{4\kappa e^{-\kappa(t-s)} r_s}{\sigma^2(1 - e^{-\kappa(t-s)})}$

and

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}[r_t \mid \mathcal{F}_s] &= r_s e^{-\kappa(t-s)} + \theta(1 - e^{-\kappa(t-s)}) \\ \text{Var}_{\mathbb{Q}}[r_t \mid \mathcal{F}_s] &= \frac{\sigma^2 r_s e^{-\kappa(t-s)} (1 - e^{-\kappa(t-s)})}{\kappa} \\ &\quad + \frac{\theta \sigma^2}{2\kappa} (1 - e^{-\kappa(t-s)})^2 \end{aligned}$$

33 – Price of a zero-coupon bond (CIR)

In the CIR model, the price of a zero-coupon bond at time t with maturity T is given by :

$$ZC(t, T) = A(t, T) \cdot e^{-B(t, T) r_t}$$

with :

$$\begin{aligned} B(t, T) &= \frac{2(e^{\gamma(T-t)} - 1)}{(\gamma + \kappa)(e^{\gamma(T-t)} - 1) + 2\gamma} \\ A(t, T) &= \left[\frac{2\gamma e^{\frac{(\kappa+\gamma)}{2}(T-t)}}{(\gamma + \kappa)(e^{\gamma(T-t)} - 1) + 2\gamma} \right]^{\frac{2\kappa\theta}{\sigma^2}} \end{aligned}$$

or :

$$\gamma = \sqrt{\kappa^2 + 2\sigma^2}$$

34 – Swaption, Black model

A **swaption** is an option on an interest rate swap. It gives the right (but not the obligation) to enter into a swap at a future date T .

- **Payer swaption** : right to *pay the fixed rate* and *receive the floating rate*.
- **Receiver swaption** : right to *receive the fixed rate* and *pay the floating rate*.

Notation :

- T : swaption exercise date
- K : fixed rate (strike)
- $S(T)$: swap rate on the date T
- $A(T)$: present value of future fixed flows.
- σ : swap rate volatility

The Black (1976) model is an adaptation of the Black–Scholes model for interest rate products. Here, the swap rate $S(T)$ plays the role of the underlying asset, with a European option-type payoff.

Black’s formula for a payer swaption :

$$\text{SW}_{\text{payer}} = A(T) [S_0 N(d_1) - K N(d_2)]$$

or :

$$\begin{aligned} d_1 &= \frac{\ln(S_0/K) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}} \\ d_2 &= d_1 - \sigma\sqrt{T} \end{aligned}$$

and $N(\cdot)$ is the cumulative distribution function of the standard normal distribution.

Formula for a receiver swaption :

$$\text{SW}_{\text{receiver}} = A(T) [K N(-d_2) - S_0 N(-d_1)]$$

35 – Life Table Notations

Age $x, y, z \dots$

l_x is the number of people alive, relative to an initial cohort, at age x

ω is the age limit of mortality tables.

$d_x = l_x - l_{x+1}$ is the number of people who die between the age x and age $x + 1$.

q_x is the probability of death between the ages of x et age $x + 1$.

$$q_x = d_x / l_x$$

p_x is the probability that the individual aged x survives age $x + 1$.

$$p_x + q_x = 1$$

Likewise, ${}_n d_x = d_x + d_{x+1} + \dots + d_{x+n-1} = l_x - l_{x+n}$ shows the number of people who die between the age x and age $x + n$.

${}_n q_x$ is the probability of death between the ages of x and age $x + n$.

$${}_n q_x = {}_n d_x / l_x$$

${}_n p_x$ is the probability of a person of age x to survive the age $x + n$.

$${}_n p_x = l_{x+n} / l_x$$

${}_m | q_x$, the probability that the individual of age x dies in the $m + 1^e$ year.

$${}_m | q_x = \frac{d_{x+m}}{l_x} = \frac{l_{x+m} - l_{x+m+1}}{l_x}$$

e_x is the life expectancy for a person still alive at the age x . This is the number of birthdays you hope to live.

$$e_x = \sum_{t=1}^{\infty} {}_t p_x$$

36 – Coefficient or commutations

These coefficients or commutations established by actuarial functions which depend on a mortality table and a rate i ($v = 1/(1+i)$) to establish the actuarial table.

$$D_x = l_x \cdot v^x$$

can be seen "as" the actualized number of survivors. The sums

$$N_x = \sum_{k \geq 0} D_{x+k} = \sum_{k=0}^{\omega-x} D_{x+k}$$

$$S_x = \sum_{k \geq 0} N_{x+k} = \sum_{k \geq 0} (k+1) \cdot D_{x+k}$$

will be used to simplify the calculations. Likewise

$$C_x = d_x v^{x+1}$$

can be seen "as" the number of deaths discounted to age x . The sums

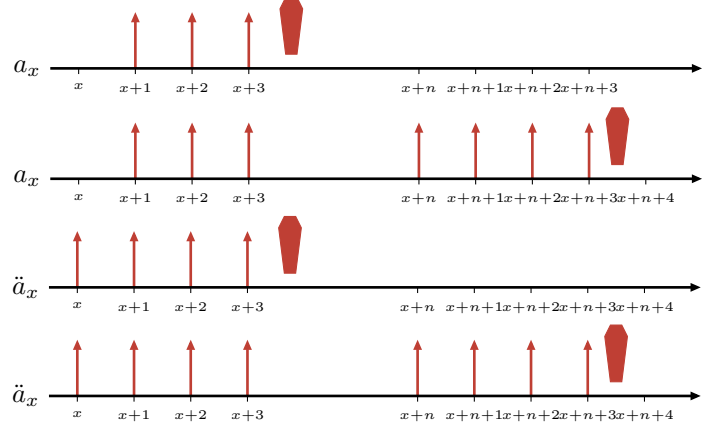
$$M_x = \sum_{k=0}^{\omega-x} C_{x+k}$$

$$R_x = \sum_{k=0}^{\omega-x} M_{x+k}$$

will be used to simplify the calculations.

The coefficients D_x N_x and S_x will be used for calculations on operations in case of life and C_x M_x and R_x for operations in case of death.

37 – Life annuities or annuities



$$a_x = \sum_{k=1}^{\infty} {}_k p_x v^k = \ddot{a}_x - 1 = \frac{N_{x+1}}{D_x}$$

$$\ddot{a}_x = \sum_{k=0}^{\infty} {}_k p_x v^k = \frac{N_x}{D_x}$$

If the periodicity corresponds to m periods per year :

$$\ddot{a}_x^{(m)} = \sum_{k=0}^{\infty} \frac{1}{m} {}_k p_x v^{\frac{k}{m}} \approx \ddot{a}_x - \frac{m-1}{2m}$$

Similarly, if he pays $1/m$ at the start of the m periods

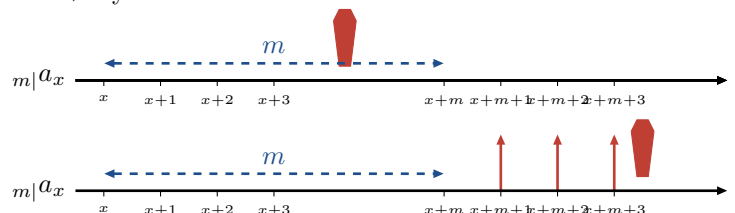
$$a_x^{(m)} \approx a_x + \frac{m-1}{2m}$$

Temporary life annuities. Whole life annuity guaranteed for n years

$$a_{x:\overline{n}|} = \sum_{k=1}^n {}_k p_x v^k = \frac{N_{x+1} - N_{x+n+1}}{D_x}$$

$$\ddot{a}_{x:\overline{n}|} = \sum_{k=0}^{n-1} {}_k p_x v^k = \frac{N_x - N_{x+n}}{D_x}$$

Deferred life annuities ${}_m | a_x$ represent the annuities on the individual of age x deferred m years. The first payment occurs in $m + 1$ years in the case of life.



38 – Death or survival benefits

Death benefits (Whole life insurance noted SP_x or A_x)

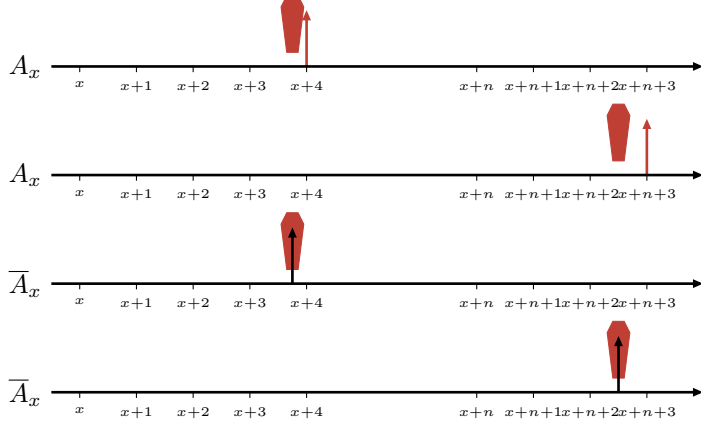
A_x indicates a death benefit at the end of the year of death (amount of 1), regardless of the date of occurrence, for an individual insured at age x at the time of subscription.

$A_{x:\overline{n}|}$ denotes a capital paid upon death if it occurs and at the latest in n years (Endowment).

$A_{x:\overline{n}|}^1$ denotes a death benefit paid if x dies within the next n years (Term insurance).

$A_x^{(12)}$ indicates a benefit payable at the end of the month of death.

\overline{A}_x indicates a benefit paid on the date of death.

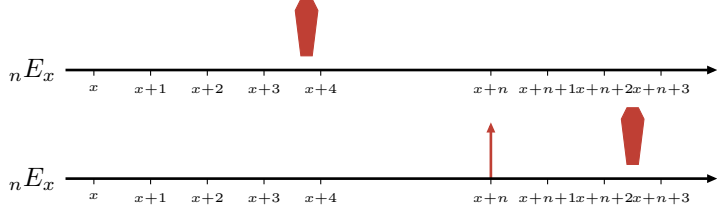


Whole life benefit

$$A_x = \sum_{k=0}^{\infty} k|q_x \nu^{k+1} = \frac{M_x}{D_x}$$

$$A_{x:\overline{n}|}^1 = \sum_{k=0}^{n-1} k|q_x \nu^{k+1} = \frac{M_x - M_{x+n}}{D_x}$$

Deferred capital (Pure Endowment, unique capital in the event of survival) noted $A_{x:\overline{n}|}^1$ or ${}_nE_x$.



$${}_nE_x = {}_np_x \cdot v^n = \frac{l_{x+n}}{l_x} \cdot v^n = \frac{D_{x+n}}{D_x}$$

Death benefit with payment of the capital in the event of survival (Endowment)

$$A_{x:\overline{n}|} = A_{x:\overline{n}|}^1 + A_{x:\overline{n}|}$$

39 – Life insurance on several individuals

a_{xyz} is an annual annuity, paid at the end of the first year and for as long as they live (x) , (y) and (z) .

$a_{\overline{xy}z}$ is an annual annuity, paid at the end of the first year and for as long as they live (x) , (y) or (z) .

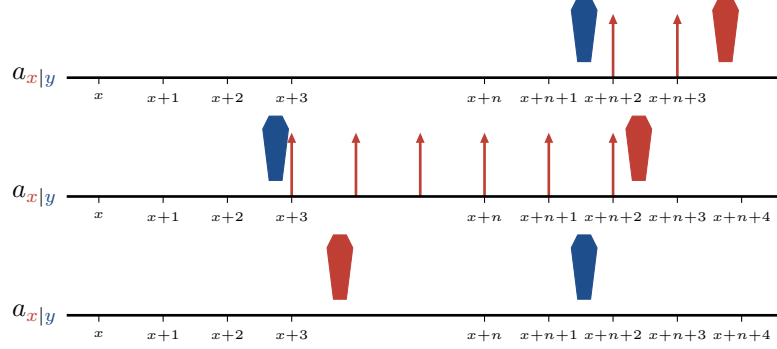
$$a_{\overline{xy}} = a_y + a_x - a_{xy}$$

A_{xyz} is an insurance that comes into effect at the end of the year of the first death of (x) , (y) and (z) .

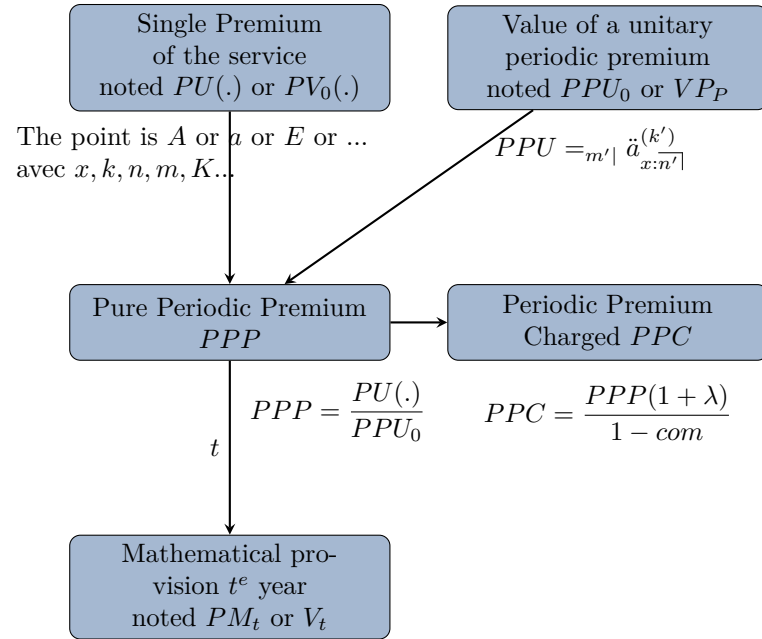
The vertical bar indicates conditionality :

$a_{x|y}$ is a survivor's annuity which benefits (x) after the death of (y) .

$A_{x|yz}$ is a first-to-die insurance (y) and (z) .



40 – Simplified pricing schemes for periodic premiums and provisions



$$V_t = PV_t(.) - PPP \times PPU_t$$

$PV_t(.)$ recalculated with $x + t, k, n, m - t, K...$ if $t \geq m$

or $x + t, k, n - (t - m), m = 0, K...$ si $t > m$

likewise $PPU_t = {}_{(m'-t)+|} \ddot{a}_{x+t:n'-(t-m)+}^{(k')}$

41 – Axiomatique

Un **univers** Ω , est l'ensemble de tous les résultats possibles qui peuvent être obtenus au cours d'une expérience aléatoire.

L'**événement aléatoire** est un événement ω_i de l'univers dont l'issue (le résultat) n'est pas certaine.

L'**événement élémentaire** :

- deux événements élémentaires distincts ω_i et ω_j sont incompatibles,
- la réunion de tous les événements élémentaires de l'univers Ω correspond à la certitude.

Les **ensembles** :

- $E = \{\omega_{i1}, \dots, \omega_{ik}\}$ un sous-ensemble de Ω (k éléments).
- \bar{E} le complémentaire de E ,
- $E \cap F$ l'intersection de E et F ,
- $E \cup F$ l'union de E et F ,
- $E \setminus F = E \cap \bar{F}$ E moins F ,
- \emptyset l'événement impossible ou vide.

Soit E un ensemble. On appelle **tribu** ou σ -**algèbre** sur E , un ensemble \mathcal{A} de parties de E qui vérifie :

- $\mathcal{A} \neq \emptyset$,
- $\forall A \in \mathcal{A}, \bar{A} \in \mathcal{A}$,
- si $\forall n \in \mathbb{N}, A_n \in \mathcal{A}$ alors $\cup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

On appelle **probabilité** \mathbb{P} toute application de l'ensemble des événements \mathcal{A} dans l'intervalle $[0, 1]$, telle que :

$$\mathbb{P} : \mathcal{A} \mapsto [0, 1]$$

satisfaisant les propriétés (ou axiomes) suivantes :

- (P1) $A \subseteq \mathcal{A}$ alors $\mathbb{P}(A) \geq 0$,
- (P2) $\mathbb{P}(\Omega) = 1$,
- (P3) $A, B \subseteq \mathcal{A}$, si $A \cap B = \emptyset$ alors $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.

L'**espace de probabilité** se définit par

$$\{\Omega, \mathcal{A}, \mathbb{P}(\cdot)\}$$

L'**égalité de poincarré** s'écrit :

$$\forall A \in \mathcal{A}, \forall B \in \mathcal{A}, \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

42 – Bayes

En théorie des probabilités, la **probabilité conditionnelle** d'un événement A , sachant qu'un autre événement B de probabilité non nulle s'est réalisé.

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Le réel $\mathbb{P}(A|B)$ se lit 'probabilité de A , sachant B . Le théorème de Bayes permet d'écrire :

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}.$$

43 – Variables aléatoires

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. On appelle **variable aléatoire** X de Ω vers \mathbb{R} , toute fonction mesurable $X : \Omega \mapsto \mathbb{R}$.

$$\{X \leq x\} \equiv \{e \in \Omega \mid X(e) \leq x\} \in \mathcal{A}$$

L'ensemble des événements de Ω n'est souvent pas explicite.

La **fonction de répartition** (F_X) d'une variable aléatoire réelle caractérise sa loi de probabilité.

$$F_X(x) = \mathbb{P}(X \leq x), x \in \mathbb{R}$$

où le membre de droite représente la probabilité que la variable aléatoire réelle X prenne une valeur inférieure ou égale à x . La probabilité que X se trouve dans l'intervalle $]a, b]$ est donc, si $a < b$, $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$

Une loi de probabilité possède une **densité de probabilité** f , si f est une fonction définie sur \mathbb{R}^+ , Lebesgue-intégrable, telle que la probabilité de l'intervalle $[a, b]$ est donnée par

$$\mathbb{P}(a < X \leq b) = \int_a^b f(x)dx \text{ pour tous nombres tq } a < x < b.$$

44 – Espérance

L'espérance mathématique dans le cas discret (variables qualitatives ou quantitatives discrètes) :

$$\mathbb{E}[X] = \sum_{j \in \mathbb{N}} x_j \mathbb{P}(x_j)$$

où $\mathbb{P}(x_j)$ est la probabilité associée à chaque événement x_i .

L'espérance mathématique dans le cas continu :

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

où f désigne la fonction de densité de la variable aléatoire x , définie dans notre cas sur \mathbb{R} . S'agissant de somme ou d'intégrale, l'espérance est linéaire, c'est-à-dire :

$$\mathbb{E}[c_0 + c_1 X_1 + c_2 X_2] = c_0 + c_1 \mathbb{E}[X_1] + c_2 \mathbb{E}[X_2]$$

$$\mathbb{E}[X] = \int x \cdot f(x) dx = \int_0^1 F^{-1}(p) dp = \int \bar{F}(x) dx$$

45 – Convolution ou loi de la somme

La convolution de deux fonctions f et g , notée $(f * g)(x)$, est définie par :

$$(f * g)(x) = \int f(t)g(x - t) dt$$

Elle est souvent utilisée pour déterminer la loi d'une somme de deux variables aléatoires indépendantes.

La convolution mesure comment $f(t)$ et $g(t)$ interagissent à différents points tout en tenant compte du décalage (ou translation) entre

46 – Loi composée ou modèle fréquence/gravité

Soit N une variable aléatoire discrète dans \mathbb{N}^+ , (X_i) une suite de variable aléatoire *iid* d'espérance et variance finies, alors pour $S = \sum_{i=1}^N X_i$:

$$\mathbb{E}(S) = \mathbb{E}(\mathbb{E}[S | N]) = \mathbb{E}(N \cdot \mathbb{E}(X_1)) = \mathbb{E}(N) \cdot \mathbb{E}(X_1)$$

$$\text{Var}(S) = \mathbb{E}(\text{Var}[S | N]) + \text{Var}(\mathbb{E}[S | N])$$

47 – Théorèmes fondamentaux

Soit X une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, et supposée presque sûrement positive ou nulle. L'**Inégalité de Markov** donne :

$$\forall \alpha > 0, \mathbb{P}(X \geq \alpha) \leq \frac{\mathbb{E}[X]}{\alpha}.$$

L'**inégalité de Bienaymé-Tchebychev** : Pour tout réel strictement positif α , avec $\mathbb{E}[X] = \mu$ et $\text{Var}[X] = \sigma^2$

$$\mathbb{P}(|X - \mu| \geq \alpha) \leq \frac{\sigma^2}{\alpha^2}.$$

La **loi faible des grands nombres** considère une suite $(X_i)_{i \geq n \in \mathbb{N}^*}$ de variables aléatoires indépendantes définies sur un même espace probabilisé, ayant mêmes espérance et variance finies notées respectivement $\mathbb{E}[X]$ et $\text{Var}(X)$.

$$\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mathbb{P}\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - \mathbb{E}[X]\right| \geq \varepsilon\right) = 0$$

Considérons une suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires indépendantes qui suivent la même loi de probabilité, intégrables, i. e. $E(|X_0|) < +\infty$.

En reprenant les notations, la **loi forte des grands nombres** précise que $(Y_n)_{n \in \mathbb{N}}$ converge vers $E(X)$ « presque sûrement ».

$$\mathbb{P}\left(\lim_{n \rightarrow +\infty} Y_n = E(X)\right) = 1$$

Considérons la somme $S_n = X_1 + X_2 + \dots + X_n$.

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}},$$

l'espérance et l'écart-type de Z_n valent respectivement 0 et 1 : la variable est ainsi dite centrée et réduite.

Le **théorème central limite** stipule alors que la loi de Z_n converge en loi vers la loi normale centrée réduite $\mathcal{N}(0, 1)$ lorsque n tend vers l'infini. Cela signifie que si Φ est la fonction de répartition de $\mathcal{N}(0, 1)$, alors pour tout réel z :

$$\lim_{n \rightarrow \infty} \mathbb{P}(Z_n \leq z) = \Phi(z),$$

ou, de façon équivalente :

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq z\right) = \Phi(z)$$

48 – Variables multidimensionnelles

Une loi de probabilité est dite **multidimensionnelle**, ou n -dimensionnelle, lorsque la loi décrit plusieurs valeurs (aléatoires) d'un phénomène aléatoire. Le caractère multidimensionnel apparaît ainsi lors du transfert, par une variable aléatoire, de l'espace probabilisé (Ω, \mathcal{A}) vers un espace numérique E^n de dimension n .

Soit une variable aléatoire X sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans \mathbb{R}^n muni de la tribu borélienne réelle produit $\mathcal{B}(\mathbb{R})^{\otimes n}$. La loi de la variable aléatoire X est la mesure de probabilité \mathbb{P}_X définie par :

$$\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}(X \in B).$$

pour tout $B \in \mathcal{B}(\mathbb{R})^{\otimes n}$.

Le théorème de **Cramer-Wold** assure que la loi $(n$ -dimensionnelle) de ce vecteur aléatoire est entièrement déterminée par les lois (unidimensionnelles) de toutes les combinaisons linéaires de ces composantes :

$$\sum_{i=1}^n a_i X_i \text{ pour tous } a_1, a_2, \dots, a_n$$

49 – Loi marginale

La loi de probabilité de la i^e coordonnée d'un vecteur aléatoire est appelée la i^e loi marginale. La **loi marginale** \mathbb{P}_i de \mathbb{P} s'obtient par la formule :

$$\mathbb{P}_i(B) = \mathbb{P}_{X_i}(B) = \iint \mathbb{1}_{\omega_i \in B} \mathbb{P}(d(\omega_1, \dots, \omega_n)), \forall B \in \mathcal{B}(\mathbb{R}).$$

Les lois marginales d'une loi absolument continue s'expriment à l'aide de leurs densités marginales.

La fonction de densité conditionnelle X_2 sachant la valeur x_1 de X_1 , peut s'écrire :

$$f_{X_2}(x_2 | X_1 = x_1) = \frac{f_{X_1, X_2}(x_1, x_2)}{f_{X_1}(x_1)},$$

$$f_{X_2}(x_2 | X_1 = x_1) f_{X_1}(x_1) = f_{X_1, X_2}(x_1, x_2) = f_{X_1}(x_1 | X_2 = x_2) f_{X_2}(x_2)$$

50 – Indépendance

(X_1, X_2, \dots, X_n) est une famille de **variables aléatoires indépendantes** si l'une des deux conditions suivantes est remplie :

$$\forall (A_1, \dots, A_n) \in \mathcal{E}_1 \times \dots \times \mathcal{E}_n$$

$$\mathbb{P}(X_1 \in A_1 \text{ et } X_2 \in A_2 \dots \text{ et } X_n \in A_n) = \prod_{i=1}^n \mathbb{P}(X_i \in A_i),$$

on a l'égalité

$$\mathbb{E}\left[\prod_{i=1}^n \varphi_i(X_i)\right] = \prod_{i=1}^n \mathbb{E}[\varphi_i(X_i)],$$

pour n'importe quelle suite de fonctions φ_i définies sur (E_i, \mathcal{E}_i) , à valeurs dans \mathbb{R} , dès que les espérances ci-dessus ont un sens.

$$f_X(x) = \prod_{i=1}^n f_{X_i}(x_i)$$

51 – Dépendance parfaite en dimension 2

Soient F_1, F_2 fonctions de répartition $\mathbb{R} \rightarrow [0, 1]$.

Les **classes de Fréchet** $\mathcal{F}_{(F_1, F_2)}$ regroupent l'ensemble des fonctions de répartition $\mathbb{R}^2 \rightarrow [0, 1]$ dont les lois marginales sont précisément F_1, F_2 .

Pour tout $F \in \mathcal{F}(F_1, F_2)$, et pour tout x in \mathbb{R}^d

$$F^-(x) \leq F(x) \leq F^+(x)$$

où $F^+(x) = \min\{F_1(x_1), F_2(x_2)\}$, et $F^-(x) = \max\{0, F_1(x_1) + F_2(x_2) - 1\}$.

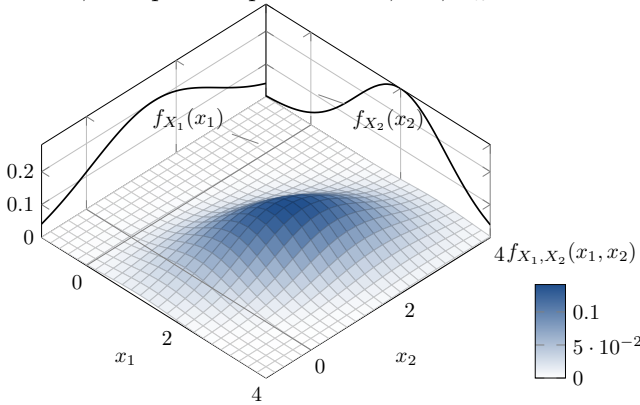
1. Le couple $\mathbf{X} = (X_1, X_2)$ est dit **comonotone** si et seulement s'il admet F^+ comme fonction de répartition.
2. Le couple $\mathbf{X} = (X_1, X_2)$ est dit **antimonotone** si et seulement s'il admet F^- comme fonction de répartition.

Le couple $\mathbf{X} = (X_1, X_2)$ est dit **comonotone (antimonotone)** s'il existe des fonctions non-décroissantes (non-croissante) g_1 et g_2 d'une variable aléatoire Z telles que

$$\mathbf{X} = (g_1(Z), g_2(Z))$$

52 – Le vecteur gaussien

Un vecteur $\mathbf{X} = (X_1, \dots, X_n)$ est dit **vecteur gaussien**, de loi $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, lorsque toute combinaison linéaire $\sum_{j=1}^n \alpha_j X_j$ de ses composantes est la loi normale univariée. En particulier, chaque composante X_1, \dots, X_n est de loi normale.



- $\boldsymbol{\mu}$ de \mathbb{R}^N sa localisation,
- $\boldsymbol{\Sigma}$ semi-définie positive de $\mathcal{M}_N(\mathbb{R})$, sa variance-covariance.

Si $\boldsymbol{\Sigma}$ est bien définie positive, donc inversible, alors

$$f_{(\boldsymbol{\mu}, \boldsymbol{\Sigma})}(\mathbf{x}) = \frac{1}{(2\pi)^{N/2} |\boldsymbol{\Sigma}|^{1/2}} e^{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})}.$$

où $|\boldsymbol{\Sigma}|$ est le déterminant de $\boldsymbol{\Sigma}$.

53 – Trois mesures de lien (corrélations)

On nomme le coefficient de **corrélation linéaire de Pearson** la valeur

$$\rho_P = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

où σ_{xy} désigne la covariance entre les variables x et y , et σ_x, σ_y leur écart type. ρ prend ses valeurs dans $[-1, 1]$ (application du théorème de Cauchy-Schwartz).

$X \perp Y \Rightarrow \rho_P = 0$, Attention, $\rho_P = 0 \not\Rightarrow X \perp Y$.

Le **tau de Kendall** se définit par

$$\tau_K = \mathbb{P}((X - X')(Y - Y') > 0) - \mathbb{P}((X - X')(Y - Y') < 0)$$

où (X, Y) (X', Y') sont deux couples indépendant de même densité jointe. Cela correspond à la probabilité des concordants diminuée de celle des discordants :

$$\begin{aligned} \tau_K &= \mathbb{P}(\text{sgn}(X - X') = \text{sgn}(Y - Y')) - \\ &\quad \mathbb{P}(\text{sgn}(X - X') \neq \text{sgn}(Y - Y')) \\ &= \mathbb{E}[\text{sgn}(X - X') \text{sgn}(Y - Y')] \\ &= \text{Cov}(\text{sgn}(X - X'), \text{sgn}(Y - Y')) \\ &= 4\mathbb{P}(X < X', Y < Y') - 1 \end{aligned}$$

Le coefficient de corrélation **rho de Spearman** de (X, Y) est défini comme le coefficient de corrélation de Pearson des rangs des variables aléatoires X et Y . Pour un échantillon n , les n valeurs X_i, Y_i sont converties par leurs rangs x_i, y_i , et ρ est calculé :

$$\rho_S = \frac{1/n \sum_i (x_i - \mathbb{E}[x])(y_i - \mathbb{E}[y])}{\sqrt{1/n \sum_i (x_i - \mathbb{E}[x])^2 \times 1/n \sum_i (y_i - \mathbb{E}[y])^2}}.$$

Si on note $x_i = R(X_i)$ de 1 à N et $d_i = x_i - y_i$:

$$\rho_S = 1 - \frac{6 \sum_i d_i^2}{n(n^2 - 1)}$$

54 – Copule

Une **copule** est une fonction de répartition, notée \mathcal{C} , définie sur $[0, 1]^d$ dont les marges sont uniformes sur $[0, 1]$. Une caractérisation est alors que $\mathcal{C}(u_1, \dots, u_d) = 0$ si une des composantes u_i est nulle, $\mathcal{C}(1, \dots, 1, u_i, 1, \dots, 1) = u_i$, et \mathcal{C} est d -croissante.

Soit $F^{(d)}$ une fonction de répartition en dimension d où les F_i sont les lois marginales de F .

Le **théorème de Sklar** indique que $F^{(d)}$ admet une représentation copule :

$$F^{(d)}(x_1, \dots, x_d) = \mathcal{C}(F_1(x_1), \dots, F_d(x_d))$$

Si ces lois marginales sont toutes continues, la copule \mathcal{C} est alors unique, et donnée par la relation

$$\mathcal{C}(u_1, \dots, u_d) = F^{(d)}(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))$$

Dans ce cas, on pourra alors parler de la copule associée à un vecteur aléatoire (X_1, \dots, X_d) . Ce théorème est très important puisque nous pouvons séparer la partie marge de distribution de la partie dépendance.

La **Copule Gaussienne** est une distribution sur le cube unitaire de dimension d , $[0, 1]^d$. Elle est construite sur la base d'une loi normale de dimension d sur \mathbb{R}^d .

Soit la matrice de corrélation $\Sigma \in \mathbb{R}^{d \times d}$, la copule Gaussienne de paramètre Σ peut s'écrire :

$$\mathcal{C}_{\Sigma}^{Gauss}(u) = \Phi_{\Sigma}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d)),$$

où Φ^{-1} est la fonction de répartition inverse de la loi normale standard et Φ_{Σ} est la distribution jointe d'une loi normale de dimension d , de moyenne nulle et de matrice de covariance égale à la matrice de corrélation Σ .

Une copule \mathcal{C} est qualifiée d'**archimédienne** si elle admet la représentation suivante :

$$\mathcal{C}(u_1, \dots, u_d) = \psi^{-1}(\psi(u_1) + \dots + \psi(u_d))$$

où ψ est alors appelé **générateur**.

Souvent, les copules admettent une formulation explicite de \mathcal{C} . Un seul paramètre permet d'accentuer la dépendance de toute la copule, quelque soit sa dimension d .

Cette formule fournit une copule si et seulement si ψ est d -monotone sur $[0, \infty)$ i.e. la dérivé k^e de ψ satisfait

$$(-1)^k \psi^{(k)}(x) \geq 0$$

pour tout $x \geq 0$ et $k = 0, 1, \dots, d-2$ et $(-1)^{d-2} \psi^{d-2}(x)$ est non-croissante et convexe.

Les générateurs suivants sont tous monotones, i.e. d -monotone pour tout $d \in \mathbb{N}$.

Nom	Générateur $\psi^{-1}(t)$,	$\psi(t)$	Paramètre
Ali-Mikhail-Haq	$\frac{1-\theta}{\exp(t)-\theta}$	$\log\left(\frac{1-\theta+\theta t}{t}\right)$	$\theta \in [0, 1)$
Clayton	$(1+\theta t)^{-1/\theta}$	$\frac{1}{\theta}(t^{-\theta}-1)$	$\theta \in (0, \infty)$
Frank	$-\frac{1}{\theta}\exp(-t)$ $\times \log(1-(1-\exp(-\theta)))$	$-\log\left(\frac{\exp(-\theta t)-1}{\exp(-\theta)-1}\right)$	$\theta \in (0, \infty)$
Gumbel	$\exp(-t^{1/\theta})$	$(-\log(t))^{\theta}$	$\theta \in [1, \infty)$
\perp	$\exp(-t)$	$-\log(t)$	
Joe	$1-(1-\exp(-t))^{1/\theta}$	$-\log(1-(1-t)^{\theta})$	$\theta \in [1, \infty)$

55 – Mouvement brownien, filtration et martingales

Une **filtration** $(\mathcal{F}_t)_{t \geq 0}$ est une famille croissante de σ -algèbres ou tribu représentant l'information disponible jusqu'au temps t . Un processus (X_t) est dit **\mathcal{F}_t -adapté** si X_t est mesurable par rapport à \mathcal{F}_t pour tout t .

Un processus $(B_t)_{t \geq 0}$ est un **mouvement brownien standard** (ou processus de Wiener) s'il vérifie :

- $B_0 = 0$;
- accroissements indépendants : $B_t - B_s$ indépendant de \mathcal{F}_s ;
- accroissements stationnaires : $B_t - B_s \sim \mathcal{N}(0, t-s)$;
- trajectoires continues presque sûrement.

Un processus (M_t) est une **martingale** (par rapport à \mathcal{F}_t) si :

$$\mathbb{E}[|M_t|] < \infty \quad \text{et} \quad \mathbb{E}[M_t | \mathcal{F}_s] = M_s \quad \forall 0 \leq s < t$$

Exemples : le mouvement brownien, les intégrales stochastiques de la forme $\int_0^t \theta_s dB_s$ (sous conditions) sont des martingales.

Variation quadratique est noté $\langle B \rangle_t = t$, $\langle cB \rangle_t = c^2 t$

Covariation : pour deux processus d'Itô X, Y ,

$$\langle X, Y \rangle_t := \lim_{\|\Pi\| \rightarrow 0} \sum_i (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i})$$

convergence en probabilité, où $\Pi = \{t_0 = 0 < t_1 < \dots < t_n = t\}$ est une partition de $[0, t]$.

56 – Processus d'Itô et calcul différentiel stochastique

Un processus (X_t) est un **processus d'Itô** s'il peut s'écrire :

$$X_t = X_0 + \int_0^t \phi_s ds + \int_0^t \theta_s dB_s$$

ou en différentielle

$$dX_t = \phi_t dt + \theta_t dB_t$$

avec ϕ_t, θ_t \mathcal{F}_t -adaptés et L^2 -intégrables.

Formule d'Itô (1D) : pour $f \in C^2(\mathbb{R})$, on a :

$$df(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)d\langle X \rangle_t$$

Exemple : si $dX_t = \mu dt + \sigma dB_t$ alors :

$$dX_t^2 = 2X_t dX_t + d\langle X \rangle_t$$

Formule d'Itô (multi-dimension) :

Si $X = (X^1, \dots, X^d)$ est un processus d'Itô, $f \in C^2(\mathbb{R}^d)$:

$$df(X_t) = \sum_i \frac{\partial f}{\partial x_i}(X_t)dX_t^i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_t)d\langle X^i, X^j \rangle_t$$

Intégration par parties (Itô) :

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + d\langle X, Y \rangle_t$$

57 – Équations différentielles stochastiques (EDS)

Une EDS (*SDE*) est une équation stochastique de la forme :

$$dX_t = b(t, X_t)dt + a(t, X_t)dB_t, \quad X_0 = x$$

où :

- $b(t, x)$ est la **dérive** (*drift*) : fonction $\mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$;
- $a(t, x)$ est la **diffusion** : fonction $\mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$;
- B_t est un mouvement brownien ;
- X_t est la solution, processus stochastique adapté.

Forme intégrale :

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t a(s, X_s)dB_s$$

Conditions d'existence et d'unicité :

- **Lipschitz** : il existe $L > 0$ tel que :

$$|b(t, x) - b(t, y)| + |a(t, x) - a(t, y)| \leq L|x - y|$$

- **Croissance linéaire** :

$$|b(t, x)|^2 + |a(t, x)|^2 \leq C(1 + |x|^2)$$

Exemples classique :

- **Brownien géométrique** : $dS_t = \mu S_t dt + \sigma S_t dB_t$
- **Ornstein-Uhlenbeck** : $dX_t = \theta(\mu - X_t)dt + \sigma dB_t$

Méthodes numériques : Euler-Maruyama, Milstein.

58 – Probabilité risque neutre

Une probabilité \mathbb{Q} est dite **risque neutre** si, sous \mathbb{Q} , tout actif S_t a un prix actualisé $\frac{S_t}{B_t}$ qui est une martingale où (B_t) est le numéraire (par exemple $B_t = e^{rt}$).

L'absence d'arbitrage $\iff \exists \mathbb{Q} \sim \mathbb{P}$ telle que les prix actualisés soient des martingales. C'est le **théorème fondamental de l'évaluation des actifs**.

Application :

Sous \mathbb{Q} , la valeur à l'instant t d'un actif donnant un retour H à la date T est :

$$S_t = B_t \mathbb{E}^{\mathbb{Q}} \left[\frac{H}{B_T} \middle| \mathcal{F}_t \right]$$

Remarque :

La mesure \mathbb{Q} est équivalente à \mathbb{P} , mais reflète un monde "sans préférence pour le risque", utile en valorisation.

Les simulations permettent en particulier d'approximer l'espérance par la moyenne empirique des réalisations x_1, \dots, x_n :

$$\frac{1}{n}(x_1 + \dots + x_n) \approx \int x dF(x) = \mathbb{E}[X]$$

Ensuite, en vertu du TLC, nous estimons l'incertitude ou intervalle de confiance sur la base de la loi normale :

$$\left[\bar{x} - 1,96 \frac{S_n}{\sqrt{n}}, \bar{x} + 1,96 \frac{S_n}{\sqrt{n}} \right]$$

où S_n estime sans biais la variance de X :

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

la convergence est dite en $\mathcal{O}(\frac{\sigma}{\sqrt{n}})$. Cet intervalle permet de décider le nombre de simulation à réaliser.

60 – Générateur pseudo aléatoire sur $[0, 1]^d$

L'ordinateur ne sait pas lancer le dé ($\Omega = \{\square, \square, \square, \square, \square, \square\}$). Il génère un pseudo aléa, c'est à dire un algorithme déterministe qui ressemble à un événement aléatoire. Les générateurs produisent généralement un aléa sur $[0, 1]^d$. Si la valeur initiale (*seed*) est définie ou identifiée, les tirages suivants sont connus et répliquables.

L'algorithme le plus simple es appelée méthode des congruences linéaires.

$$x_{n+1} = \Phi(x_n) = (a \times x_n + c) \mod m$$

chaque x_n est un entier compris entre 0 et $m-1$. a le multiplicateur [*multiplier*], c l'accroissement [*increment*], et m le modulo [*modulus*] de la forme $2^p - 1$, c'est-à-dire un nombre premier de Mersenne (p nécessairement premier) :

Générateur de Marsaglia : $a = 69069, b = 0, m = 2^{32}$

Générateur Knuth&Lewis : $a = 1664525, b = 1013904223, m = 2^{32}$

Générateur de Haynes : $a = 6364136223846793005, b = 0, m = 2^{64}$

Le générateur de Tausworthe, constitue une extension 'autorégressi-ve' :

$$x_n = (a_1 \times x_{n-1} + a_2 \times x_{n-2} + \dots + a_k \times x_{n-k}) \mod m \text{ avec } n \geq k$$

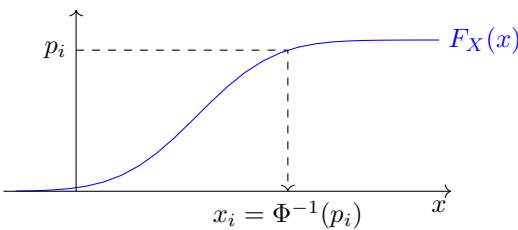
La période du générateur est $m^k - 1$, avec tous les a_i premiers entre eux. Si m est de la forme 2^p , les temps de calculs machine sont réduits. Le générateur aléatoire par défaut est généralement l'algorithme de Mersenne-Twister. Il est basé sur une récurrence linéaire sur une matrice F_2 (matrice dont les valeurs sont en base 2, i.e. 0 ou 1). Son nom provient du fait que la longueur de la période choisie pour être un nombre premier de Mersenne.

1. sa période est de $2^{19937} - 1$
2. il est uniformément distribué sur un grand nombre de dimensions (623 pour les nombres de 32 bits) ;
3. il est plus rapide que la plupart des autres générateurs ,
4. il est aléatoire quel que soit le poids du bit considéré, et passe les tests Diehard.

61 – Simuler une variable aléatoire

Simuler X d'une loi quelconque F_X revient souvent à simuler $(p_i)_{i \in [1, n]}$ de loi $\mathcal{Uni}(0, 1)$.

Si F_X est inversible, $x_i = F_X^{-1}(p_i)$ (ou fonction quantile) livre $(x_i)_{i \in [1, n]}$ un jeu de n simulations de loi F_X .



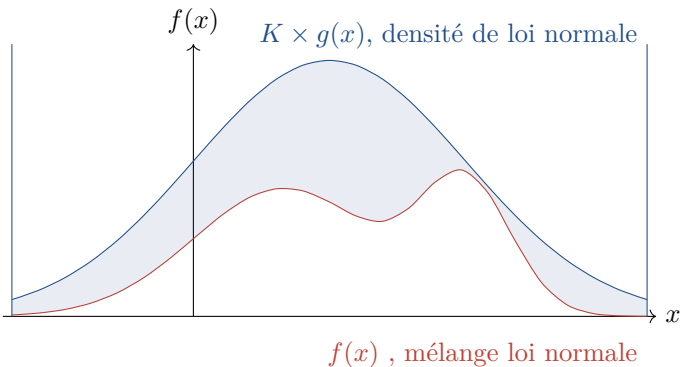
Si c'est une variable discrète (F^{-1} n'existe pas) $X_\ell = \min_{\ell} F(X_\ell) > p_i$, où $(X_\ell)_\ell$ l'ensemble dénombrable des valeurs possibles, ordonnées de manière croissante.

Dans la méthode du **changement de variable**, on suppose qu'on sait simuler une loi X , et qu'il existe ϕ tel que $Y = \phi(X)$ suit une loi F_Y . L'exemple naturelle est celui de $X \sim \mathcal{N}(0, 1)$ et de faire le changement $Y = \exp(X)$ pour obtenir Y qui suit une loi lognormale.

La méthode du rejet est utilisée dans des cas plus complexe, par exemple lorsque F^{-1} n'est pas explicite ou exige beaucoup de temps de calcul. Soit f une fonction de densité de probabilités. On suppose qu'il existe une densité de probabilités g telle que :

$$\exists K > 0, \forall x \in \mathbb{R}, f(x) \leq K g(x)$$

On simule alors Z suivant la loi de densité g , et $Y \sim \mathcal{U}([0; Kg(Z)])$. Alors la variable aléatoire $X = \{Z | Y \leq f(Z)\}$ suit la loi de densité f .



La performance de l'algorithme dépend du nombre de rejet, représenté par la surface bleu sur le graphique.

62 – Méthodes de Monte Carlo

Les **méthodes de Monte Carlo** reposent sur la simulation répétée de variables aléatoires pour approximer des quantités numériques.

Convergence :

- Par la **loi des grands nombres**, l'estimateur converge presque sûrement vers la valeur attendue.
- Par le **théorème central limite**, l'erreur standard est en $\mathcal{O}(N^{-1/2})$:

$$\sqrt{N}(\hat{\mu}_N - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

- Cette convergence lente justifie l'usage de techniques d'**amélioration de la convergence**.

Techniques de réduction de variance :

- **Variables antithétiques** : on simule X et $-X$ (ou $1 - U$ si $U \sim \mathcal{U}[0, 1]$), puis on moyenne les résultats. Réduction efficace si f est monotone.
- **Méthode de contrôle** : si $\mathbb{E}[Y]$ est connue, on simule $(f(X), Y)$ et on corrige :

$$\hat{\mu}_{\text{corr}} = \hat{\mu} - \beta(\bar{Y} - \mathbb{E}[Y])$$

où β optimal minimise la variance.

- **Stratification** : on divise l'espace des simulations en strates (sous-ensembles), et on simule proportionnellement dans chaque strate.
- **Importance sampling** : on modifie la loi de simulation pour accentuer les événements rares, puis on répond :

$$\mathbb{E}[f(X)] = \mathbb{E}^Q \left[f(X) \frac{dP}{dQ}(X) \right]$$
 utilisée notamment pour estimer les queues de distribution (VaR, TVaR).

63 – Le bootstrap

Le **bootstrap** est une méthode de *rééchantillonnage* permettant d'estimer l'incertitude d'un estimateur sans supposer de forme paramétrique pour la loi sous-jacente.

Soit un échantillon $\xi = (X_1, X_2, \dots, X_n)$ de variables iid suivant une loi inconnue F . On cherche à estimer une statistique $\theta = T(F)$ (ex. moyenne, médiane, variance), via son estimateur empirique $\hat{\theta} = T(\hat{F}_n)$.

1. On approxime F par la fonction de répartition empirique :

$$\hat{F}_n(x) = \frac{1}{n} \sum_{k=1}^n \mathbf{1}_{\{X_k \leq x\}}$$

2. On génère B échantillons bootstrap $\xi^{*(b)} = (X_1^{*(b)}, \dots, X_n^{*(b)})$ en tirant **avec remise** dans l'échantillon initial.
3. Pour chaque échantillon simulé, on calcule l'estimation $T^{*(b)} = T(\hat{F}_n^{*(b)})$.

Les réalisations $T^{*(1)}, \dots, T^{*(B)}$ forment une approximation de la distribution de l'estimateur $\hat{\theta}$.

On peut en déduire :

- un **biais estimé** : $\widehat{\text{bias}} = \overline{T^*} - \hat{\theta}$;
- un **intervalle de confiance** à $(1 - \alpha)$: $[q_{\alpha/2}, q_{1-\alpha/2}]$ des quantiles empiriques de $T^{*(b)}$;
- une estimation de la **variance** : $\widehat{\text{Var}}(T^*)$.

Remarque : Le bootstrap est particulièrement utile lorsque la distribution de T est inconnue ou difficile à estimer analytiquement.

64 – Bootstrap paramétrique

Le **bootstrap paramétrique** repose sur l'hypothèse que les données suivent une famille de lois paramétrée $\{F_\theta\}$.

Soit $\xi = (X_1, \dots, X_n)$ un échantillon iid selon une loi F_θ inconnue. On procède comme suit :

1. Estimer le paramètre $\hat{\theta}$ à partir de ξ (ex. par maximum de vraisemblance).
2. Générer B échantillons $\xi^{*(b)}$ de taille n , simulés selon la loi $F_{\hat{\theta}}$.
3. Calculer $T^{*(b)} = T(\xi^{*(b)})$ pour chaque échantillon.

Ce procédé approxime la distribution de l'estimateur $T(\xi)$ en supposant connue la forme de F . Il est plus efficace que le bootstrap non paramétrique si l'hypothèse de modèle est bien spécifiée. Le bootstrap paramétrique est plus rapide, mais hérite des biais du modèle.

65 – Validation croisée

La **validation croisée** est une méthode d'évaluation de la performance prédictive d'un modèle statistique, utilisée notamment en machine learning ou en tarification.

Principe :

- Diviser les données en K blocs (ou folds).
- Pour chaque $k = 1, \dots, K$:
 - Entraîner le modèle sur les $K - 1$ autres blocs.

- Évaluer la performance (erreur, log-vraisemblance...) sur le k -ième bloc.

- Agréger les erreurs pour obtenir une estimation globale de la performance hors-échantillon.

66 – Méthodes quasi-Monte Carlo

Les méthodes **quasi-Monte Carlo** visent à **accélérer la convergence** de l'estimateur d'espérance sans recourir à l'aléatoire. L'erreur typique est de l'ordre :

$$\mathcal{O} \left(\frac{(\ln N)^s}{N} \right)$$

où N est la taille de l'échantillon et s la dimension du problème. Ces méthodes reposent sur l'utilisation de suites à **discrédance faible** dans $[0, 1]^s$. La discrédance à l'origine (*star discrepancy*), notée $D_N^*(P)$ pour un ensemble de points $P = \{x_1, \dots, x_N\}$, mesure l'écart maximal entre la proportion de points contenus dans des rectangles *ancrés à l'origine* et leur volume. Elle est définie par :

$$D_N^*(P) = \sup_{u \in [0, 1]^s} \left| \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{[0, u)}(x_i) - \lambda_s([0, u)) \right|$$

avec :

- $[0, u) = \prod_{j=1}^s [0, u_j)$ un rectangle ancré à l'origine dans $[0, 1]^s$,
- $\mathbf{1}_{[0, u)}(x_i)$ l'indicatrice de l'appartenance de x_i à ce rectangle,
- $\lambda_s([0, u)) = \prod_{j=1}^s u_j$ le volume de ce rectangle.

Une discrédance faible signifie que les points sont bien répartis dans l'espace, ce qui améliore la convergence de l'estimation.

Suite de van der Corput (dimension 1) : Soit n un entier. On l'écrit en base b :

$$n = \sum_{k=0}^{L-1} d_k(n) b^k$$

puis on inverse les chiffres autour de la virgule pour obtenir :

$$g_b(n) = \sum_{k=0}^{L-1} d_k(n) b^{-k-1}$$

Par exemple, pour $b = 5$ et $n = 146$, on a $146 = (1\,0\,4\,1)_5$, donc :

$$g_5(146) = \frac{1}{5^4} + \frac{0}{5^3} + \frac{4}{5^2} + \frac{1}{5} = 0,3616$$

Séquence de Halton (dimension s) : On généralise la suite de van der Corput en utilisant s bases entières premières distinctes b_1, \dots, b_s :

$$x(n) = (g_{b_1}(n), \dots, g_{b_s}(n))$$

Cette construction fournit une suite de points bien répartis dans $[0, 1]^s$.

Inégalité de Koksma-Hlawka

Pour une fonction f de variation finie $V(f)$ (au sens de Hardy-Krause) sur $[0, 1]^s$:

$$\left| \int_{[0, 1]^s} f(u) du - \frac{1}{N} \sum_{i=1}^N f(x_i) \right| \leq V(f) D_N$$

où D_N est la discrédance de la suite utilisée.

Cette borne explique pourquoi les méthodes quasi-Monte Carlo sont souvent plus efficaces que les méthodes de Monte Carlo.

Distribution	Densité & support	Moments &	
		fonction de répartition	Fonction génératrice de moment
$\text{Bin}(m, q)$ ($0 < p < 1, m \in \mathbb{N}$)	$\binom{m}{x} p^x (1-p)^{m-x}$ $x = 0, 1, \dots, m$	$E = mp, \text{Var} = mp(1-p)$ $\gamma = \frac{mp(1-p)(1-2p)}{\sigma^3}$	$(1-p+pe^t)^m$
$\text{Ber}(q)$	$\equiv \text{Binomial}(1, p)$		
$\mathcal{D}\text{Uni}(n)$ ($n > 0$)	$\frac{1}{n}, x = 0, 1, \dots, n$	$\mathbb{E} = (n+1)/2$ $\text{Var} = (n^2-1)/12$	$\frac{e^t(1-e^{nt})}{n(1-e^t)}$
$\mathcal{P}\text{ois}(\lambda)$ ($\lambda > 0$)	$e^{-\lambda} \frac{\lambda^x}{x!}, x = 0, 1, \dots$	$\mathbb{E} = \text{Var} = \lambda$ $\gamma = 1/\sqrt{\lambda}$ $\kappa_j = \lambda, j = 1, 2, \dots$	$\exp[\lambda(e^t-1)]$
$\mathcal{N}\text{Bin}(m, q)$ ($m > 0, 0 < p < 1$)	$\binom{m+x-1}{x} p^m (1-p)^x$ $x = 0, 1, 2, \dots$	$\mathbb{E} = m(1-p)/p$ $\text{Var} = \mathbb{E}/p$ $\gamma = \frac{(2-p)}{p\sigma}$	$\left(\frac{p}{1-(1-p)e^t}\right)^m$
$\mathcal{G}\text{eo}(q)$	$\equiv \mathcal{N}\text{Bin}(1, q)$		
$\mathcal{C}\text{Uni}(a, b)$ ($a < b$)	$\frac{1}{b-a}; a < x < b$	$\mathbb{E} = (a+b)/2,$ $\text{Var} = (b-a)^2/12,$ $\gamma = 0$	$\frac{e^{bt}-e^{at}}{(b-a)t}$
$\mathcal{N}(\mu, \sigma^2)$ ($\sigma > 0$)	$\frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-(x-\mu)^2}{2\sigma^2}$	$\mathbb{E} = \mu, \text{Var} = \sigma^2, \gamma = 0$ ($\kappa_j = 0, j \geq 3$)	$\exp(\mu t + \frac{1}{2}\sigma^2 t^2)$
$\mathcal{G}\text{am}(k, \theta)$ ($k, \theta > 0$)	$\frac{\theta^k}{\Gamma(k)} x^{k-1} e^{-\theta x}, x > 0$	$\mathbb{E} = k/\theta, \text{Var} = k/\theta^2,$ $\gamma = 2/\sqrt{k}$	$\left(\frac{\theta}{\theta-t}\right)^k (t < \theta)$
$\mathcal{E}_{xp}(\lambda)$	$\equiv \mathcal{G}\text{am}(1, \lambda)$	$\mathbb{E} = 1/\lambda$ $\text{Var} = 1/\lambda^2$	
$\chi^2(k) (k \in \mathbb{N})$	$\equiv \mathcal{G}\text{am}(k/2, 1/2)$		
$\mathcal{I}\mathcal{N}(\alpha, \beta)$ ($\alpha > 0, \beta > 0$)	$\frac{\alpha x^{-3/2}}{\sqrt{2\pi\beta}} \exp\left(\frac{-(\alpha-\beta x)^2}{2\beta x}\right)$	$\mathbb{E} = \alpha/\beta, \text{Var} = \alpha/\beta^2,$ $\gamma = 3/\sqrt{\alpha}$	$e^{\alpha(1-\sqrt{1-2t/\beta})}$ $(t \leq \beta/2)$
	$F(x) = \Phi\left(\frac{-\alpha}{\sqrt{\beta x}} + \sqrt{\beta x}\right) + e^{2\alpha} \Phi\left(\frac{-\alpha}{\sqrt{\beta x}} - \sqrt{\beta x}\right), x > 0$		
$\text{Beta}(\alpha, \beta)$ ($\alpha > 0, \beta > 0$)	$\Gamma(\alpha+\beta) \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)}, 0 < x < 1$	$\mathbb{E} = \frac{\alpha}{\alpha+\beta}, \text{Var} = \frac{\mathbb{E}(1-\mathbb{E})}{\alpha+\beta+1}$	
$\mathcal{L}\mathcal{N}\text{orm}(\mu, \sigma^2)$ ($\sigma > 0$)	$\frac{1}{x\sigma\sqrt{2\pi}} \exp \frac{-(\log x - \mu)^2}{2\sigma^2}, x > 0$	$\mathbb{E} = e^{\mu+\sigma^2/2}, \text{Var} = e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2}$ $\gamma = c^3 + 3c$ où $c^2 = \text{Var}/\mathbb{E}^2$	
$\mathcal{P}\text{areto}(\alpha, x_m)$ ($\alpha, x_m > 0$)	$\frac{\alpha x_m^\alpha}{x^{\alpha+1}}, x > x_m$	$\mathbb{E} = \frac{\alpha x_m}{\alpha-1} \quad \alpha > 1, \text{Var} = \frac{\alpha x_m^2}{(\alpha-1)^2(\alpha-2)} \quad \alpha > 2$	
$\text{Weibull}(\alpha, \beta)$ ($\alpha, \beta > 0$)	$\alpha\beta(\beta y)^{\alpha-1} e^{-(\beta y)^\alpha}, x > 0$	$\mathbb{E} = \Gamma(1+1/\alpha)/\beta$ $\text{Var} = \Gamma(1+2/\alpha)/\beta^2 - \mathbb{E}^2$ $\mathbb{E}[Y^t] = \Gamma(1+t/\alpha)/\beta^t$	

67 – Concept d'utilité

L'utilité modélise les préférences d'un individu entre deux paniers de biens x et y dans un ensemble S , via la relation $x \succ y$ (préférée ou indifférent).

Une fonction $U : S \rightarrow \mathbb{R}$ représente les préférences si :

$$x \succ y \iff U(x) \geq U(y)$$

Axiomes nécessaires à l'existence d'une fonction d'utilité :

1. **Complétude** : Pour tout $x, y \in S$, soit $x \succ y$, soit $y \succ x$
2. **Transitivité** : Si $x \succ y$ et $y \succ z$, alors $x \succ z$
3. **Continuité** : Si $x_n \rightarrow x$ et $y_n \rightarrow y$, et $x_n \succ y_n$ pour tout n , alors $x \succ y$

68 – Fonction d'utilité

Une fonction $u : \mathbb{R}_+ \rightarrow \mathbb{R}$ représente les préférences d'un agent face à l'incertitude.

Critère d'utilité espérée : L'agent préfère X à Y si :

$$\mathbb{E}[u(X)] > \mathbb{E}[u(Y)]$$

Il choisit X tel que $\mathbb{E}[u(X)]$ soit maximale.

Propriétés de u :

- $u' > 0$: l'agent préfère plus de richesse (monotonie)
- $u'' < 0$: l'agent est averse au risque (concavité)

Exemples classiques :

- *Linéaire* (neutre au risque) : $u(x) = x$
- *Logarithmique* : $u(x) = \ln(x)$
- *CRRA* (aversion relative constante) : $u(x) = \frac{x^{1-\gamma}}{1-\gamma}$, $\gamma \neq 1$
- *CARA* (aversion absolue constante) : $u(x) = -e^{-ax}$

69 – Aversion au risque

Un agent est dit **averse au risque** (ou risquophobe) si :

$$u(\mathbb{E}[X]) > \mathbb{E}[u(X)]$$

Ce qui est équivalent à u concave, c.-à-d. $u''(x) < 0$

70 – Mesure de l'aversion au risque

Indice d'aversion absolue :

$$A_a(x) = -\frac{u''(x)}{u'(x)}$$

Indice d'aversion relative :

$$A_r(x) = -x \cdot \frac{u''(x)}{u'(x)}$$

Inégalité de Jensen (cas concave) :

$$u(\mathbb{E}[X]) \geq \mathbb{E}[u(X)]$$

Avec l'égalité si et seulement si X est constante.

71 – Primes de risque

La **prime de risque** π est le montant maximal qu'un individu est prêt à payer pour remplacer une loterie de gain aléatoire H par son espérance certaine $\mathbb{E}[H]$. Elle vérifie :

$$\mathbb{E}[u(w + H)] = u(w + \mathbb{E}[H] - \pi)$$

π est aussi appelée *mesure de Markowitz* : elle capture l'écart entre utilité espérée et utilité certaine.

Inversement, la **prime compensatoire** $\tilde{\pi}$ est le montant que l'on doit offrir à un individu pour qu'il accepte la loterie H au lieu d'un gain certain. Elle vérifie :

$$\mathbb{E}[u(w + H + \tilde{\pi})] = u(w + \mathbb{E}[H])$$

72 – Diversification et utilité

Soient deux actifs :

- A : risqué
- B : certain, avec $\mathbb{E}[A] = B$

Un agent averse au risque préfère une combinaison $Z = \alpha A + (1 - \alpha)B$, avec $0 < \alpha < 1$, à l'actif risqué seul. Si u est concave, alors

$$\mathbb{E}[u(Z)] > \mathbb{E}[u(A)]$$

Portefeuille optimal : choix des pondérations (w_i) maximisant l'utilité espérée :

$$\max \mathbb{E}[u(X)], \quad \text{où } X = \sum_i w_i X_i, \quad \text{s.c. } \sum w_i = 1$$

Principe : la diversification réduit le risque (variance) sans affecter l'espérance.

73 – Méthode de Lagrange pour l'optimisation sous contrainte

La méthode des multiplicateurs de Lagrange permet de résoudre un problème d'optimisation sous contrainte.

Objectif : maximiser/minimiser $f(\mathbf{x})$ sous la contrainte $g(\mathbf{x}) = c$, où $\mathbf{x} \in \mathbb{R}^d$ est un vecteur de variables.

Étapes de la méthode :

1. **Identification** : déterminer la fonction objectif $f(\mathbf{x})$ et la contrainte $g(\mathbf{x}) = c$
2. **Lagrangien** :

$$\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda(g(\mathbf{x}) - c)$$

3. **Système d'équations** : résoudre

$$\nabla_{\mathbf{x}} \mathcal{L} = \nabla f(\mathbf{x}) + \lambda \nabla g(\mathbf{x}) = \mathbf{0}, \quad \frac{\partial \mathcal{L}}{\partial \lambda} = g(\mathbf{x}) - c = 0$$

4. **Résolution** du système pour obtenir \mathbf{x}^*, λ^*
5. **Vérification** : s'assurer que les solutions satisfont bien la contrainte et le type d'optimum (max/min)

Exemple (dimension 2) : maximiser $f(x, y) = xy$ sous la contrainte $x + y = 10$

$$\mathcal{L}(x, y, \lambda) = xy + \lambda(x + y - 10)$$

On dérive :

$$\frac{\partial \mathcal{L}}{\partial x} = y + \lambda = 0, \quad \frac{\partial \mathcal{L}}{\partial y} = x + \lambda = 0, \quad \frac{\partial \mathcal{L}}{\partial \lambda} = x + y - 10 = 0$$

On résout le système :

$$\begin{cases} y + \lambda = 0 \\ x + \lambda = 0 \\ x + y = 10 \end{cases} \Rightarrow \begin{cases} \lambda = -y \\ x = -\lambda = y \\ x + y = 10 \Rightarrow 2x = 10 \end{cases} \Rightarrow \begin{cases} x^* = y^* = 5, \\ f(5, 5) = 25 \end{cases}$$

Exemple (Choix optimal et contrainte budgétaire)

Un agent rationnel est face à un choix de consommation (c_1, c_2) entre deux biens, sous la contrainte :

$$p_1 c_1 + p_2 c_2 = R$$

où p_1, p_2 sont les prix et R le revenu total.

Problème : $\max_{c_1, c_2} u(c_1, c_2)$ s.c. $p_1 c_1 + p_2 c_2 = R$

Méthode : introduire le **lagrangien**

$$\mathcal{L}(c_1, c_2, \lambda) = u(c_1, c_2) + \lambda(R - p_1 c_1 - p_2 c_2)$$

Conditions du premier ordre (FOC) :

$$\begin{cases} \frac{\partial u}{\partial c_1} = \lambda p_1 \\ \frac{\partial u}{\partial c_2} = \lambda p_2 \\ p_1 c_1 + p_2 c_2 = R \end{cases}$$

En divisant les deux premières équations :

$$\frac{\partial u / \partial c_1}{\partial u / \partial c_2} = \frac{p_1}{p_2}$$

Ce rapport est appelé **taux marginal de substitution (TMS)** : il mesure la quantité de bien 2 à laquelle l'agent est prêt à renoncer pour obtenir une unité supplémentaire de bien 1, tout en maintenant son niveau d'utilité constant.

74 – Demande d'assurance (Mosin)

Un agent possède un patrimoine initial w et fait face à une perte aléatoire L . Il existe une **demande d'assurance** pour l'assurance qui verse l'indemnité $0 < I(L) < L$ ssi $u(w - \pi_I) \geq \mathbb{E}(u(w - L))$ et l'**assurance optimale** maximise $u(w - \pi_I)$.

Dans Mosin (1968) ou Borch (1961) ou Smith (1968), le modèle de perte L se définit simplement par s compris entre 0 et w :

$$L = \begin{cases} 0 & \text{avec proba. } 1 - p \\ s & \text{avec proba. } p \end{cases}$$

La prime devient $\pi_I = (1 + \lambda)\mathbb{E}(I(L)) = (1 + \lambda)pI(s)$ avec λ un chargement. On note π le cas où $I(L) = L$ avec $\pi = ps$. Si $\lambda = 0$, alors on parle de prime pure ou actuariellement juste.

Coassurance (partage du risque) : $I(l) = \alpha l$ sachant $L = l$ pour $\alpha \in [0, 1]$, $\pi_I(\alpha) = \alpha \pi$ et :

$$w_f = w - L + I(L) - \pi(\alpha) = w - L + \alpha L - \alpha \pi = w - (1 - \alpha)L - \alpha \pi$$

$$U(\alpha) = (1 - p)u(w - \alpha \pi) + pu(w - (1 - \alpha)s - \alpha \pi)$$

L'assurance partielle ($\alpha^* < 1$) est optimale ssi $\lambda > 0$. L'assurance totale ($\alpha^* = 1$) est optimale si le chargement est nul.

L'assurance avec franchise (part d'autoassurance) : Avec franchise d l'assureur verse une indemnité $I(l) = (l - d)_+$ sachant $L = l$.

$$\pi(d) = (1 + \lambda)\mathbb{E}((L - d)_+) = (1 + \lambda)(s - d)p$$

$$w_f = w - X + (L - d)_+ - \pi(d) = w - \min(X, d) - (1 + \lambda)(s - d)p$$

$$U(d) = (1 - p)u(\underbrace{w + (1 + \lambda)(d - s)p}_{w_f^+}) + pu(\underbrace{w - d + (1 + \lambda)(d - s)p}_{w_f^-})$$

Dans le modèle avec franchise, l'assurance partielle ($d^* > 0$) est optimale ssi la prime n'est pas actuariellement juste. De même, l'assurance totale ($d^* = 0$) est optimale si le chargement est nul.

Modèle généralisé : Le risque de perte $L > 0$ aléatoire est défini sur \mathbb{R} , avec fonction de répartition F_L ,

$$\pi_I = (1 + \lambda)\mathbb{E}(I(L)) = (1 + \lambda) \int_0^\infty I(l) dF_L(l)$$

1. L'assurance totale ($d^* = 0$) ou ($\alpha^* = 1$) est optimale si et seulement si la prime est actuariellement juste.
2. Si $A_a(u, x)$ est décroissante, alors le niveau de franchise d^* ou le taux de couverture α^* augmente avec la richesse initiale. Pour les préférences CARA, d^* est indépendant de w ou α^* est constant.
3. Le niveau de couverture décroît avec le coefficient de chargement λ lorsque $A_a(u, x)$ est croissante ou constante.
4. Un agent plus averse au risque choisit une couverture plus élevée.

75 – Information et assurance

Mosin avec hétérogénéité : Deux types d'individus : H pour haut risque et Lo pour faible risque. $\theta \in [0, 1]$ la proportion d'individu H . Les individus de type $i \in \{Lo, H\}$ font face à un risque de même montant s survenant avec une probabilité différente p_i telle que $1 > p_H > p_{Lo} > 0$.

$$L_i = \begin{cases} 0 & \text{avec probabilité } 1 - p_i, \\ s & \text{avec probabilité } p_i. \end{cases}$$

La probabilité du marché :

$$p_m = \theta p_H + (1 - \theta)p_{Lo}.$$

Absence d'antisélection : Dans ce modèle, en présence d'information totale, l'assureur préfère une assurance individuelle $I_i = s$ et $\pi_i = sp_i$, $\forall i$.

Le problème d'anti-sélection : L'assureur propose un contrat non individualisé du marché $M = (\pi_m = p_m I, I_m(s) = I(s))$, qui ne dépend pas de i . La fortune finale d'un individu de type i est $W_i^m = w - \pi_m - X_i + I_m$.

En présence d'un contrat unique, les individus de type H préfèrent un contrat d'assurance tel que $I_H(s) = s$ et $\pi_H = sp_m$, tandis que les individus de type Lo préfèrent une couverture partielle avec $I_L^* < s$ et $\pi_{Lo} = I_L^* p_m$.

Aléa moral : L'assuré réduit ou interrompt ses efforts maintenant qu'il est assuré. Les efforts de

- prévention réduisent la probabilité de sinistre,
- protection réduisent le montant de sinistre.

En l'absence d'effort e de prévention ou de réduction du risque, la fortune finale w_f est simplement définie par

$$\begin{cases} w_f^- = w - s - \pi(I) + I & \text{avec probabilité } p \\ w_f^+ = w - \pi(I) & \text{avec probabilité } 1 - p \end{cases}$$

S'il y a prévention de risque d'effort e , on considère

$$\begin{cases} w_f^- = w - s - \pi(I) + I - e & \text{avec probabilité } p(e) \\ w_f^+ = w - \pi(I) - e & \text{avec probabilité } 1 - p(e) \end{cases}$$

S'il y a réduction de risque d'effort e , on considère

$$\begin{cases} w_f^- = w - s(e) - \pi(I) + I - e & \text{avec probabilité } p \\ w_f^+ = w - \pi(I) - e & \text{avec probabilité } 1 - p \end{cases}$$

avec

- $e \mapsto p(e)$ est strictement décroissante et strictement convexe.
- $e \mapsto s(e)$ est strictement décroissante et strictement convexe.
- $I \leq s$ implique $w_f^- \leq w_f^+$

76 – Définitions

Série temporelle - succession d'observations quantitatives d'un phénomène classées par ordre temporelle.

Il existe plusieurs variantes de séries temporelles :

- **Données de panel** - consistant en une série temporelle pour chaque observation d'une section transversale.
- **Sections transversales regroupées** - combinant des sections transversales de différentes périodes.

Processus stochastique - séquence de variables aléatoires indexées dans le temps.

77 – Composantes d'une série temporelle

- **Tendance** - mouvement général à long terme d'une série.
- **Variations saisonnières** - oscillations périodiques qui se produisent sur une période égale ou inférieure à une année et qui peuvent être facilement identifiées d'une année à l'autre (elles sont généralement dues à des raisons climatologiques).
- **Variations cycliques** - sont des oscillations périodiques qui se produisent sur une période supérieure à une année (elles sont le résultat du cycle économique).
- **Variations résiduelles** - sont des mouvements qui ne suivent pas une oscillation périodique reconnaissable (elles sont le résultat de phénomènes ponctuels).

78 – Types de modèles de séries temporelles

- **Modèles statiques** - la relation entre y et x est contemporaine. Conceptuellement :

$$y_t = \beta_0 + \beta_1 x_t + u_t$$

- **Modèles à décalage distribué** - la relation entre y et x n'est pas contemporaine. Conceptuellement :

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + \dots + \beta_s x_{t-(s-1)} + u_t$$

L'effet cumulatif à long terme dans y lorsque Δx est : $\beta_1 + \beta_2 + \dots + \beta_s$

- **Modèles dynamiques** - décalages de la variable dépendante (endogénéité). Conceptuellement :

$$y_t = \beta_0 + \beta_1 y_{t-1} + \dots + \beta_s y_{t-s} + u_t$$

- **Combinaisons** de ce qui précède, comme les modèles rationnels à décalage distribué (décalage distribué + dynamique).

79 – Hypothèses du modèle OLS dans le cadre des séries temporelles

Sous ces hypothèses, l'estimateur OLS présentera de bonnes propriétés. **Hypothèses de Gauss-Markov** étendues aux séries temporelles :

t1. Linéarité des paramètres et faible dépendance.

- a. y_t doit être une fonction linéaire des β .
- b. Le stochastique $\{(x_t, y_t) : t = 1, 2, \dots, T\}$ est stationnaire et faiblement dépendant.

t2. Pas de colinéarité parfaite.

- Il n'y a pas de variables indépendantes qui sont constantes : $\text{Var}(x_j) \neq 0, \forall j = 1, \dots, k$
- Il n'y a pas de relation linéaire exacte entre les variables indépendantes.

t3. Moyenne conditionnelle nulle et corrélation nulle.

- a. Il n'y a pas d'erreurs systématiques : $\mathbb{E}(u | x_1, \dots, x_k) = \mathbb{E}(u) = 0 \rightarrow$ **exogénéité forte** (a implique b).
- b. Il n'y a pas de variables pertinentes laissées hors du modèle : $\text{Cov}(x_j, u) = 0, \forall j = 1, \dots, k \rightarrow$ **exogénéité faible**.

t4. Homoscédasticité.

La variabilité des résidus est la même pour tout x : $\text{Var}(u | x_1, \dots, x_k) = \sigma_u^2$

t5. Pas d'autocorrélation.

Les résidus ne contiennent aucune information sur les autres résidus : $\text{Corr}(u_t, u_s | x_1, \dots, x_k) = 0, \forall t \neq s$

t6. Normalité.

Les résidus sont indépendants et identiquement distribués (**i.i.d.** etc.) : $u \sim \mathcal{N}(0, \sigma_u^2)$

t7. Taille des données.

Le nombre d'observations disponibles doit être supérieur à $(k + 1)$ paramètres à estimer. (Cette condition est déjà satisfaite dans des situations asymptotiques)

80 – Propriétés asymptotiques de l'OLS

Sous les hypothèses du modèle économétrique et le théorème de la limite centrale :

- Si t1 à t3a sont vérifiées : l'OLS est **non biaisée**. $\mathbb{E}(\hat{\beta}_j) = \beta_j$
- Si t1 à t3 sont vérifiées : l'OLS est **cohérent**. $\text{plim}(\hat{\beta}_j) = \beta_j$ (pour t3b, t3a est omis, exogénéité faible, biaisé mais cohérent)
- Maintenir t1 à t5 : **normalité asymptotique** de l'OLS (alors, t6 est nécessairement satisfait) : $u \underset{a}{\sim} \mathcal{N}(0, \sigma_u^2)$
- Maintenir t1 à t5 : **estimation non biaisée** de σ_u^2 . $\mathbb{E}(\hat{\sigma}_u^2) = \sigma_u^2$
- Maintenir t1 à t5 : OLS est **BLUE, Best Linear Unbiased Estimator** (meilleur estimateur linéaire non biaisé) ou **efficace**.
- Maintenir t1 à t6 : les tests d'hypothèse et les intervalles de confiance peuvent être effectués de manière fiable.

81 – Tendances et saisonnalité

Régression parasite - se produit lorsque la relation entre y et x est due à des facteurs qui affectent y et qui sont corrélés à x , $\text{Corr}(x_j, u) \neq 0$. Il s'agit du **non-respect de t3**.

Deux séries temporelles peuvent avoir la même tendance (ou une tendance contraire), ce qui devrait entraîner un niveau élevé de corrélation. Cela peut provoquer une fausse apparence de causalité, le problème étant la **régression fallacieuse**. Étant donné le modèle :

$$y_t = \beta_0 + \beta_1 x_t + u_t$$

où :

$$\begin{aligned} y_t &= \alpha_0 + \alpha_1 \text{Tendance} + v_t \\ x_t &= \gamma_0 + \gamma_1 \text{Tendance} + v_t \end{aligned}$$

L'ajout d'une tendance au modèle peut résoudre le problème :

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 \text{Tendance} + u_t$$

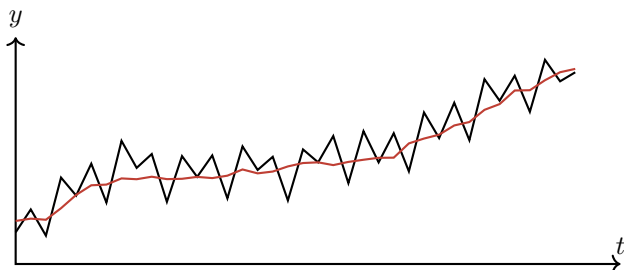
La tendance peut être linéaire ou non linéaire (quadratique, cubique, exponentielle, etc.).

Une autre méthode consiste à utiliser le **filtre de Hodrick-Prescott** pour extraire la tendance et la composante cyclique.

82 – Saisonnalité

Une série temporelle peut présenter une saisonnalité. Cela signifie que la série est soumise à des variations ou à des schémas saisonniers, généralement liés aux conditions climatiques.

Par exemple, le PIB (en noir) est généralement plus élevé en été et plus faible en hiver. Série corrigée des variations saisonnières (en orange) à titre de comparaison.



— Ce problème est une **régression parasite**. Un ajustement saisonnier peut le résoudre.

Un simple **ajustement saisonnier** pourrait consister à créer des variables binaires stationnaires et à les ajouter au modèle. Par exemple, pour les séries trimestrielles (Q_{qt} sont des variables binaires) :

$$y_t = \beta_0 + \beta_1 Q_{2t} + \beta_2 Q_{3t} + \beta_3 Q_{4t} + \beta_4 x_{1t} + \dots + \beta_k x_{kt} + u_t$$

Une autre méthode consiste à ajuster les variables en fonction des variations saisonnières (sa), puis à effectuer la régression avec les variables ajustées :

$$z_t = \beta_0 + \beta_1 Q_{2t} + \beta_2 Q_{3t} + \beta_3 Q_{4t} + v_t \rightarrow \hat{v}_t + \mathbb{E}(z_t) = \hat{z}_t^{sa} \\ \hat{y}_t^{sa} = \beta_0 + \beta_1 \hat{x}_{1t}^{sa} + \dots + \beta_k \hat{x}_{kt}^{sa} + u_t$$

Il existe des méthodes bien plus efficaces et complexes pour ajuster saisonnièrement une série temporelle, comme la méthode **X-13ARIMA-SEATS**.

83 – Autocorrélation

Le résidu de toute observation, u_t , est corrélé avec le résidu de toute autre observation. Les observations ne sont pas indépendantes. Il s'agit d'un cas de **non-respect** de **t5**.

$$\text{Corr}(u_t, u_s \mid x_1, \dots, x_k) = \text{Corr}(u_t, u_s) \neq 0, \forall t \neq s$$

84 – Conséquences

— Les estimateurs OLS restent non biaisés.

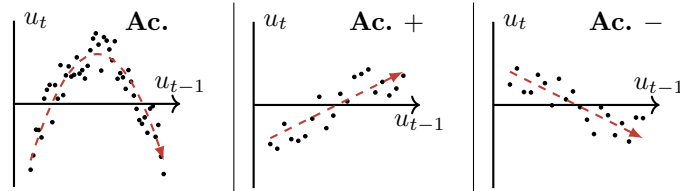
— Les estimateurs OLS restent cohérents.

— L'OLS n'est **plus efficace**, mais reste un LUE (estimateur linéaire non biaisé).

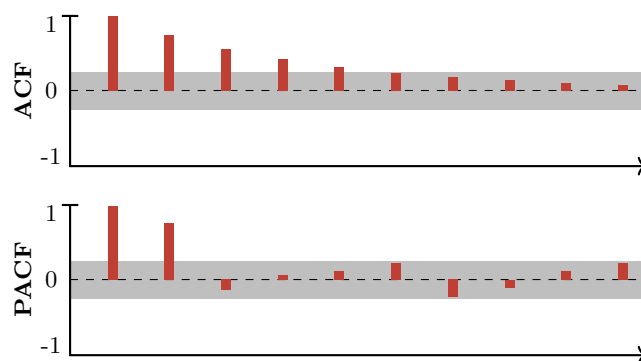
— Les **estimations de variance** des estimateurs sont **biaisées** : la construction des intervalles de confiance et les tests d'hypothèse ne sont pas fiables.

85 – Détection

— **Diagrammes de dispersion** - recherchez des modèles de dispersion sur u_{t-1} par rapport à u_t .



— **Corrélogramme** - fonction d'autocorrélation (ACF) et fonction d'autocorrélation partielle (PACF).
— Axe Y : corrélation.
— Axe X : nombre de décalages.
— Zone grise : $\pm 1,96/T^{0,5}$



Processus MA(q). ACF : seuls les premiers coefficients q sont significatifs, les autres sont brusquement annulés. PACF : décroissance exponentielle rapide atténuée ou ondes sinusoïdales.

Processus AR(p). ACF : décroissance exponentielle rapide atténuée ou ondes sinusoïdales. PACF : seuls les premiers coefficients p sont significatifs, les autres sont brusquement annulés.

Processus ARMA(p, q). ACF et PACF : les coefficients ne sont pas brusquement annulés et présentent une décroissance rapide.

Si les coefficients ACF ne décroissent pas rapidement, cela indique clairement un manque de stationnarité dans la moyenne.

— **Tests formels** - En général, H_0 : pas d'autocorrélation. En supposant que u_t suit un processus AR(1) :

$$u_t = \rho_1 u_{t-1} + \varepsilon_t$$

où ε_t est un bruit blanc.

Test t AR(1) (régresseurs exogènes) :

$$t = \frac{\hat{\rho}_1}{\text{se}(\hat{\rho}_1)} \sim t_{T-k-1, \alpha/2}$$

— H_1 : Autocorrélation d'ordre un, AR(1).

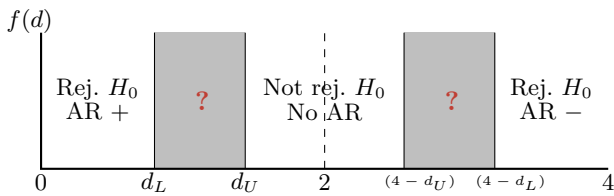
Statistique de Durbin-Watson (régresseurs exogènes et normalité des résidus) :

$$d = \frac{\sum_{t=2}^n (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^n \hat{u}_t^2} \approx 2 \cdot (1 - \hat{\rho}_1)$$

Où $0 \leq d \leq 4$

— H_1 : Autocorrélation d'ordre un, AR(1).

$$\begin{array}{c|c|c|c} d = & 0 & 2 & 4 \\ \hline \rho \approx & 1 & 0 & -1 \end{array}$$



h de Durbin (régresseurs endogènes) :

$$h = \hat{\rho} \cdot \sqrt{\frac{T}{1 - T \cdot v}}$$

où v est la variance estimée du coefficient associé à la variable endogène.

— H_1 : Autocorrélation d'ordre un, AR(1).

Test de Breusch-Godfrey (régresseurs endogènes) : il permet de détecter les processus MA(q) et AR(p) (ε_t est w. bruit) :

— MA(q) : $u_t = \varepsilon_t - m_1 u_{t-1} - \dots - m_q u_{t-q}$

— AR(p) : $u_t = \rho_1 u_{t-1} + \dots + \rho_p u_{t-p} + \varepsilon_t$

Sous H_0 : Pas d'autocorrélation :

$$T \cdot R_{\hat{u}_t}^2 \underset{a}{\sim} \chi_q^2 \quad \text{ou} \quad T \cdot R_{\hat{u}_t}^2 \underset{a}{\sim} \chi_p^2$$

— H_1 : Autocorrélation d'ordre q (ou p).

Test Q de Ljung-Box :

— H_1 : Autocorrélation jusqu'au décalage h .

86 – Correction — Utiliser la méthode des moindres carrés ordinaires (OLS) avec un estimateur de matrice de variance-covariance **robuste à l'hétéroscédasticité et à l'autocorrélation** (HAC), par exemple celui proposé par Newey-West.

— Utiliser les **moindres carrés généralisés** (GLS). Supposons que $y_t = \beta_0 + \beta_1 x_t + u_t$, avec $u_t = \rho u_{t-1} + \varepsilon_t$, où $|\rho| < 1$ et ε_t est un bruit blanc.

— Si ρ est **connu**, utilisez un **modèle quasi-différencié** :

$$y_t - \rho y_{t-1} = \beta_0(1 - \rho) + \beta_1(x_t - \rho x_{t-1}) + u_t - \rho u_{t-1}$$

$$y_t^* = \beta_0^* + \beta_1^* x_t^* + \varepsilon_t$$

où $\beta_1' = \beta_1$; et estimez-le par OLS.

— Si ρ n'est **pas connu**, l'estimer par exemple par la **méthode itérative de Cochrane-Orcutt** (la méthode de Prais-Winsten est également valable) :

1. Obtenir \hat{u}_t à partir du modèle original.
2. Estimez $\hat{u}_t = \rho \hat{u}_{t-1} + \varepsilon_t$ et obtenez $\hat{\rho}$.
3. Créez un modèle quasi-différencié :

$$y_t - \hat{\rho} y_{t-1} = \beta_0(1 - \hat{\rho}) + \beta_1(x_t - \hat{\rho} x_{t-1}) + u_t - \hat{\rho} u_{t-1}$$

$$y_t^* = \beta_0^* + \beta_1^* x_t^* + \varepsilon_t$$

où $\beta_1' = \beta_1$; et l'estimer par OLS.

4. Obtenir $\hat{u}_t^* = y_t - (\hat{\beta}_0^* + \hat{\beta}_1^* x_t) \neq y_t - (\hat{\beta}_0^* + \hat{\beta}_1^* x_t^*)$.
5. Répéter à partir de l'étape 2. L'algorithme se termine lorsque les paramètres estimés varient très peu entre les itérations.

— Si le problème n'est pas résolu, rechercher une **forte dépendance** dans la série.

87 – Lissage exponentiel $f_t = \alpha y_t + (1 - \alpha) f_{t-1}$
où $0 < \alpha < 1$ est le paramètre de lissage.

88 – Prévisions

Deux types de prévisions :

- De la valeur moyenne de y pour une valeur spécifique de x .
- D'une valeur individuelle de y pour une valeur spécifique de x .

Statistique U de Theil - compare les résultats prévus avec les résultats des prévisions réalisées à partir d'un minimum de données historiques.

$$U = \sqrt{\frac{\sum_{t=1}^{T-1} \left(\frac{y_{t+1} - y_t}{y_t} \right)^2}{\sum_{t=1}^{T-1} \left(\frac{y_{t+1} - y_t}{y_t} \right)^2}}$$

- < 1 : la prévision est meilleure qu'une simple estimation.
- $= 1$: la prévision est à peu près aussi bonne qu'une simple estimation.
- > 1 : La prévision est moins bonne qu'une simple estimation.

89 – Stationnarité

La stationnarité permet d'identifier correctement les relations entre les variables qui restent inchangées dans le temps.

- **Processus stationnaire** (stationnarité stricte) : si un ensemble de variables aléatoires est pris et décalé de h périodes (changements de temps), la distribution de probabilité conjointe doit rester inchangée.
- **Processus non stationnaire** : par exemple, une série avec une tendance, où au moins la moyenne change avec le temps.
- **Processus stationnaire de covariance** - il s'agit d'une forme plus faible de stationnarité :
 - $\mathbb{E}(x_t)$ est constant. — $\text{Var}(x_t)$ est constant.
 - Pour tout $t, h \geq 1$, $\text{Cov}(x_t, x_{t+h})$ dépend uniquement de h , et non de t .

90 – Faible dépendance

La faible dépendance remplace l'hypothèse d'échantillonnage aléatoire pour les séries temporelles.

- Un processus stationnaire $\{x_t\}$ est **faiblement dépendant** lorsque x_t et x_{t+h} sont presque indépendants lorsque h augmente sans limite.
- Un processus stationnaire de covariance est **faiblement dépendant** si la corrélation entre x_t et x_{t+h} tend vers 0 suffisamment rapidement lorsque $h \rightarrow \infty$ (ils ne sont pas corrélés de manière asymptotique).

Les processus faiblement dépendants sont appelés **intégrés d'ordre zéro**, I(0). Quelques exemples :

— **Moyenne mobile** - $\{x_t\}$ est une moyenne mobile d'ordre q , MA(q) :

$$x_t = e_t + m_1 e_{t-1} + \dots + m_q e_{t-q}$$

où $\{e_t : t = 0, 1, \dots, T\}$ est une séquence *i.i.d.* avec une moyenne nulle et une variance σ_e^2 .

— **Processus autorégressif** - $\{x_t\}$ est un processus autorégressif d'ordre p , AR(p) :

$$x_t = \rho_1 x_{t-1} + \dots + \rho_p x_{t-p} + e_t$$

où $\{e_t : t = 1, 2, \dots, T\}$ est une séquence *i.i.d.* avec une moyenne nulle et une variance σ_e^2 .

Condition de stabilité : si $1 - \rho_1 z - \dots - \rho_p z^p = 0$ pour $|z| > 1$, alors $\{x_t\}$ est un processus AR(p) stable qui est faiblement dépendant. Pour AR(1), la condition est : $|\rho_1| < 1$.

— **Processus ARMA** - est une combinaison de AR(p) et MA(q) ; $\{x_t\}$ est un ARMA(p, q) :

$$x_t = e_t + m_1 e_{t-1} + \dots + m_q e_{t-q} + \rho_1 x_{t-1} + \dots + \rho_p x_{t-p}$$

91 – Racines unitaires

Un processus est I(d), c'est-à-dire intégré d'ordre d , si l'application de différences d fois rend le processus stationnaire.

Lorsque $d \geq 1$, le processus est appelé **processus à racine unitaire** ou on dit qu'il a une racine unitaire.

Un processus a une racine unitaire lorsque la condition de stabilité n'est pas remplie (il existe des racines sur le cercle unitaire).

92 – Forte dépendance

La plupart du temps, les séries économiques sont fortement dépendantes (ou très persistantes). Quelques exemples de **racine unitaire** I(1) :

— **Marche aléatoire** - un processus AR(1) avec $\rho_1 = 1$.

$$y_t = y_{t-1} + e_t$$

où $\{e_t : t = 1, 2, \dots, T\}$ est une séquence *i.i.d.* avec une moyenne nulle et une variance σ_e^2 .

— **Marche aléatoire avec dérive** - un processus AR(1) avec $\rho_1 = 1$ et une constante.

$$y_t = \beta_0 + y_{t-1} + e_t$$

où $\{e_t : t = 1, 2, \dots, T\}$ est une séquence *i.i.d.* avec une moyenne nulle et une variance σ_e^2 .

93 – Tests de racine unitaire

Test	H_0	Rejeter H_0
ADF	I(1)	tau < Valeur critique
KPSS	Niveau I(0)	mu > Valeur critique
	Tendance I(0)	tau > Valeur critique
Phillips-Perron	I(1)	Z-tau < Valeur critique
Zivot-Andrews	I(1)	tau < Valeur critique

94 – De la racine unitaire à la faible dépendance

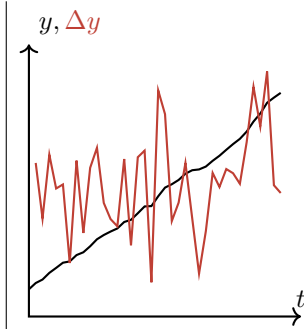
Intégré d'ordre **un**, I(1), signifie que **la première différence** du processus est **faiblement dépendante** ou I(0) (et généralement stationnaire). Par exemple, soit $\{y_t\}$ une marche aléatoire :

$$\Delta y_t = y_t - y_{t-1} = e_t$$

où $\{e_t\} = \{\Delta y_t\}$ est *i.i.d.*

Remarque :

- La première différence d'une série supprime sa tendance.
- Les logarithmes d'une série stabilisent sa variance.



De la racine unitaire au pourcentage de variation

Lorsqu'une série I(1) est strictement positive, elle est généralement convertie en logarithmes avant de prendre la première différence pour obtenir le pourcentage de variation (approximatif) de la série :

$$\Delta \log(y_t) = \log(y_t) - \log(y_{t-1}) \approx \frac{y_t - y_{t-1}}{y_{t-1}}$$

95 – Cointégration

Lorsque **deux séries sont I(1), mais qu'une combinaison linéaire de celles-ci est I(0)**. Dans ce cas, la régression d'une série sur l'autre n'est pas fallacieuse, mais exprime quelque chose sur la relation à long terme. Les variables sont dites cointégrées si elles ont une tendance stochastique commune.

Par exemple, $\{x_t\}$ et $\{y_t\}$ sont I(1), mais $y_t - \beta x_t = u_t$ où $\{u_t\}$ est I(0). (β est le paramètre de cointégration).

96 – Test de cointégration

En suivant l'exemple ci-dessus :

1. Estimer $y_t = \alpha + \beta x_t + \varepsilon_t$ et obtenir $\hat{\varepsilon}_t$.
2. Effectuer un test ADF sur $\hat{\varepsilon}_t$ avec une distribution modifiée.

Le résultat de ce test est équivalent à :

- $H_0 : \beta = 0$ (pas de cointégration)
- $H_1 : \beta \neq 0$ (cointégration)

si la statistique du test > valeur critique, rejeter H_0 .

97 – Hétéroscédasticité sur les séries temporelles

L'hypothèse affectée est **t4**, ce qui conduit à **une inefficacité de l'OLS**.

Utilisez des tests tels que Breusch-Pagan ou White, où H_0 : pas d'hétéroscédasticité. Il est **important** pour que les tests fonctionnent qu'il n'y ait **pas d'autocorrélation**.

98 – ARCH

Une hétéroscédasticité conditionnelle autorégressive (ARCH) est un modèle permettant d'analyser une forme d'hétéroscédasticité dynamique, où la variance de l'erreur suit un processus AR(p).

Étant donné le modèle : $y_t = \beta_0 + \beta_1 z_t + u_t$ où il y a AR(1) et hétéroscédasticité :

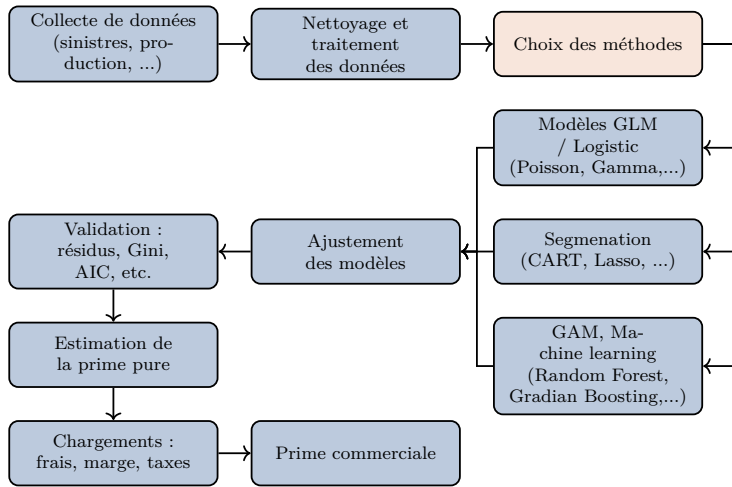
$$\mathbb{E}(u_t^2 | u_{t-1}) = \alpha_0 + \alpha_1 u_{t-1}^2$$

99 – GARCH

Un modèle général d'hétéroscédasticité conditionnelle autorégressive (GARCH) est similaire au modèle ARCH, mais dans ce cas, la variance de l'erreur suit un processus ARMA(p, q).

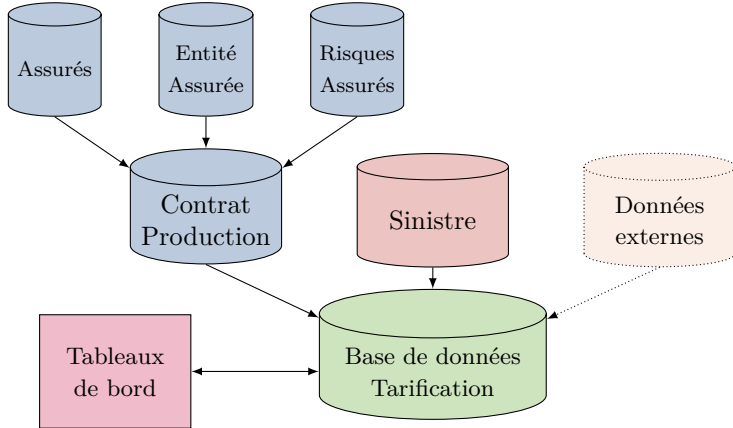
100 – La tarification en assurance non-vie

Une approche générale, mais non exhaustive, car les possibles sont nombreux :



101 – Structure générale des données en assurance

Une structure classique des données en assurance. Là encore, les possibles sont nombreux :



102 – Provision

L'actuaire non-vie évalue principalement les provisions suivantes :

- Des provisions pour sinistres à payer (PSAP, *Reserves for claims reported but not settled (RBNS)*)
- Des provisions pour sinistres non encore manifestés (PS-NEM, *Reserve for claims incurred but not reported (IBNR)*)
- Des provisions pour primes non acquises (PPNA, *Reserves for unearned premiums*)
- Des provisions pour risques en cours (PREC, *Reserves for outstanding risks (non-life)*)

103 – Chain Ladder déterministe

Soit C_{ik} le montant, cumulé jusqu'en l'année de développement k , des sinistres survenus en l'année d'accident i , pour $1 \leq i, k \leq n$. C_{ik} peut représenter soit le montant payé, soit le cout total estimé (paiement déjà effectué plus réserve) du sinistre. Les montants C_{ik} sont connus pour $i + k \leq n + 1$ et on cherche à estimer les valeurs des C_{ik} pour $i + k > n + 1$, et en particulier les valeurs ultimes C_{in} pour $2 \leq i \leq n$. Ces notations sont illustrées dans le triangle suivant : La méthode de Chain

Ladder estime les montants inconnus, C_{ik} pour $i + k > n + 1$, par

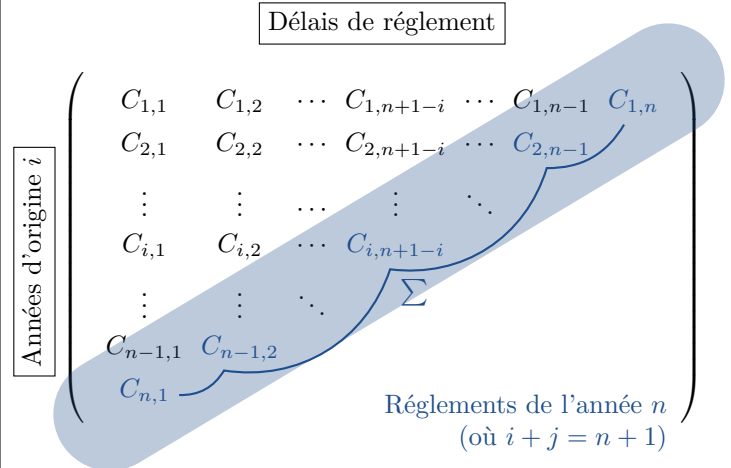
$$\hat{C}_{ik} = C_{i,n+1-i} \cdot \hat{f}_{n+1-i} \cdots \hat{f}_{k-1} \quad i + k > n + 1 \quad (5)$$

où

$$\hat{f}_k = \frac{\sum_{i=1}^{n-k} C_{i,k+1}}{\sum_{i=1}^{n-k} C_{ik}} \quad 1 \leq k \leq n - 1. \quad (6)$$

La réserve de sinistre pour l'année d'accident (R_i , $2 \leq i \leq n$), est alors estimée par

$$\begin{aligned} \hat{R}_i &= C_{in} - C_{i,n+1-i} \\ &= C_{i,n+1-i} \cdot \hat{f}_{n+1-i} \cdots \hat{f}_{n-1} - C_{i,n+1-i} \end{aligned}$$



104 – Méthode de Mack

Les deux premières hypothèses sont les suivantes :

$$E(C_{i,k+1} | C_{i1}, \dots, C_{ik}) = C_{ik} f_k \quad 1 \leq i \leq n, 1 \leq k \leq n - 1 \quad (7)$$

$$\{C_{i1}, \dots, C_{in}\}, \{C_{j1}, \dots, C_{jn}\} \quad \forall i, j \text{ sont indépendants} \quad (8)$$

Mack démontre que si on estime les paramètres du modèle (7) par (6) alors ce modèle stochastique (7), combiné avec l'hypothèse (8) fournit exactement les mêmes réserves que la méthode originale de Chain Ladder (5).

Avec la notation $f_{i,k} = \frac{C_{i,k+1}}{C_{i,k}}$, \hat{f}_k est la moyenne des $f_{i,k}$ pondérée par les $C_{i,k}$:

$$\hat{f}_k = \frac{\sum_{i=1}^{n-k} C_{i,k} \times f_{i,k}}{\sum_{i=1}^{n-k} C_{ik}}$$

La variance s'écrit :

$$\begin{aligned} \hat{\sigma}_k^2 &= \frac{1}{n-k-1} \sum_{i=1}^{n-k} C_{ik} \left(\frac{C_{i,k+1}}{C_{ik}} - \hat{f}_k \right)^2 \\ &= \frac{1}{n-k-1} \sum_{i=1}^{n-k} \left(\frac{C_{i,k+1} - C_{i,k} \hat{f}_k}{\sqrt{C_{i,k}}} \right)^2 \end{aligned}$$

Le troisième hypothèse concerne la distribution de R_i pour pouvoir construire facilement des intervalles de confiance sur les réserves estimées. Si la distribution est normale, de moyenne la valeur estimée \hat{R}_i et d'écart-type donné par l'erreur standard se (\hat{R}_i) . Un intervalle de confiance à 95% est alors donné par $[\hat{R}_i - 2 \text{ se}(\hat{R}_i), \hat{R}_i + 2 \text{ se}(\hat{R}_i)]$.

Si la distribution est supposée lognormale, les bornes d'un intervalle de confiance à 95% seront alors données par

$$\left[\hat{R}_i \exp\left(\frac{-\sigma_i^2}{2} - 2\sigma_i\right), \hat{R}_i \exp\left(\frac{-\sigma_i^2}{2} + 2\sigma_i\right) \right]$$

105 – Le modèle risque collectif

Le modèle collectif est le modèle de base en actuariat non-vie ? X_i désigne le montant du i^e , N désigne le nombre de sinistres et S le montant total au cours d'une année

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

en sachant que $S = 0$ lorsque $N = 0$ et que $\{X_i\}_{i=1}^\infty$ est une séquence iid et $N \perp \{X_i\}_{i=1}^\infty$. La difficulté est d'obtenir la distribution de S , alors même que $\mathbb{E}[N]$ n'est pas grand au sens du TCL.

106 – La distribution de S

Soit $G(x) = \mathbb{P}(S \leq x)$, $F(x) = \mathbb{P}(X_1 \leq x)$, et $p_n = \mathbb{P}(N = n)$ de sorte que $\{p_n\}_{n=0}^\infty$ soit la fonction de probabilité pour le nombre de sinistres.

$$\{S \leq x\} = \bigcup_{n=0}^\infty \{S \leq x \text{ et } N = n\}$$

$$\mathbb{P}(S \leq x \mid N = n) = \mathbb{P}\left(\sum_{i=1}^n X_i \leq x\right) = F^{n*}(x)$$

Ainsi, pour $x \geq 0$

$$G(x) = \sum_{n=0}^\infty p_n F^{n*}(x)$$

où F^{n*} désigne la convolution n^e , malheureusement elle n'existe pas sous forme fermée pour de nombreuses distributions.

Si $E[X] = m$

$$E[S] = E[Nm] = E[N]m$$

Ce résultat est très intéressant, car il indique que le montant total attendu des sinistres est le produit du nombre attendu de sinistres et du montant attendu de chaque sinistre. De même, en utilisant le fait que $\{X_i\}_{i=1}^\infty$ sont des variables aléatoires indépendantes,

$$V[S \mid N = n] = V\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n V[X_i]$$

$$\begin{aligned} V[S] &= E[V(S \mid N)] + V[E(S \mid N)] \\ &= E[N]V[X_i] + V[N]m^2 \end{aligned}$$

107 – La classe de distributions $(a, b, 0)$

Une distribution de comptage est dite $(a, b, 0)$ si sa fonction de probabilité $\{p_n\}_{n=0}^\infty$ peut être calculée de manière récursive à partir de la formule

$$p_n = \left(a + \frac{b}{n}\right) p_{n-1}$$

pour $n = 1, 2, 3, \dots$, où a et b sont des constantes.

Il existe exactement trois distributions non triviales dans la classe $(a, b, 0)$, à savoir Poisson, binomiale et binomiale négative. Voici les valeurs de a et b pour les principales distributions $(a, b, 0)$:

	a	b
$\mathcal{P}_{ois}(\lambda)$	0	λ
$\mathcal{B}_{in}(n, q)$	$-q/(1-q)$	$(n+1)q/(1-q)$
$\mathcal{NB}_{in}(k, q)$	$1-q$	$(1-q)(k-1)$
$\mathcal{G}_{eo}(q)$	$1-q$	0
Distribution de Panjer	$\frac{\lambda}{\alpha+\lambda}$	$\frac{(\alpha-1)\lambda}{\alpha+\lambda}$

La loi géométrique est un cas particulier de la binomiale négative où $k=1$.

108 – Algorithme d'agrégation de Panjer

L'algorithme de Panjer vise l'estimation de distribution d'une loi composée coût-fréquence dans des conditions particulières.

- $(X_i)_{i=1}^N$ iid discrètes définies sur $\{0, h, 2h, 3h, \dots\}$
- la loi du nombre dans la classe dite $(a, b, 0)$

Puisque nous supposons désormais que les montants individuels des demandes sont répartis sur les entiers non négatifs, il s'ensuit que S est également réparti sur les entiers non négatifs. Comme $S = \sum_{i=1}^N X_i$, il s'ensuit que $S = 0$ si $N = 0$ ou si $N = n$ et $\sum_{i=1}^n X_i = 0$. Comme $\sum_{i=1}^n X_i = 0$ uniquement si chaque $X_i = 0$, il s'ensuit par indépendance que

$$\mathbb{P}\left(\sum_{i=1}^n X_i = 0\right) = f_0^n$$

$$\begin{cases} g_0 = p_0 + \sum_{n=1}^\infty p_n f_0^n = P_N(f_0) \text{ si } a \neq 0, \\ g_0 = p_0 \cdot \exp(f_0 b) \text{ si } a = 0, \\ g_k = \frac{1}{1-af_0} \sum_{j=1}^k \left(a + \frac{bj}{k}\right) f_j g_{k-j} \end{cases}$$

g_x est exprimé en fonction de g_0, g_1, \dots, g_{x-1} , de sorte que le calcul de la fonction de probabilité est récursif. Dans toutes les applications pratiques de cette formule, un ordinateur est nécessaire pour effectuer les calculs. Cependant, l'avantage de la formule de récursivité de Panjer par rapport à la formule pour g_x est qu'il n'est pas nécessaire de calculer les convolutions, ce qui est beaucoup plus efficace d'un point de vue computationnel.

109 – Panjer et la loi de Poisson

Lorsque la fréquence suit une loi de Poisson, cela implique que $a = 0$ et $b = \lambda$.

$$\begin{cases} g_0 = e^{-\lambda(1-f_0)} \\ g_k = \frac{\lambda}{k} \sum_{j=1}^k j \cdot f_j \cdot g_{k-j} \end{cases}$$

L'algorithme de Panjer nécessite la discrétisation de la variable Y_i , l'excédent de sinistre entre la la franchise et la limite.

110 – Panjer et Pollaczek-Khinchine-Beekman

Soit τ_1 le premier instant où $R_t < \kappa (= \kappa_0)$. On pose alors $L_1 = \kappa - R_{\tau_1}$. On redémarre le processus avec $\kappa_1 = \kappa_0 - R_{\tau_1}$ pour trouver τ_2 et $L_2 = \kappa_1 - R_{\tau_2}$. En continuant de la sorte, on constate que :

$$M = \sup_{t \geq 0} \{S_t - ct\} = \sum_{k=1}^K L_k$$

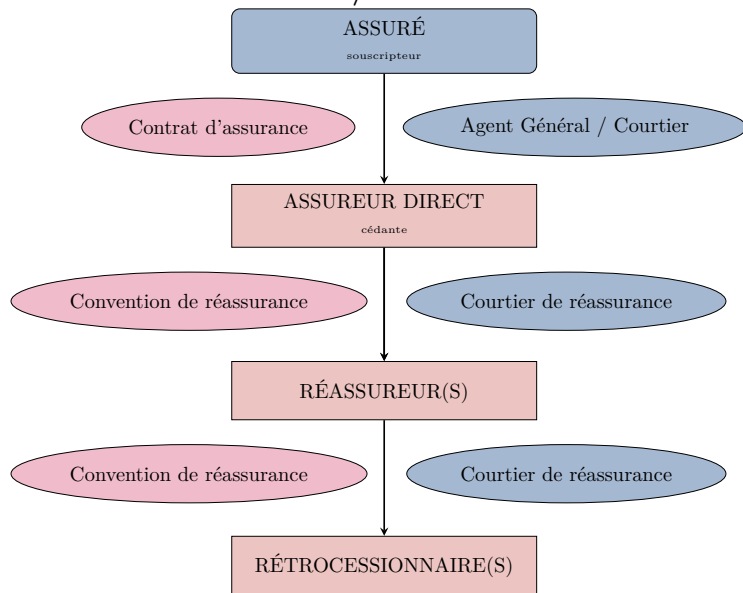
où $K \sim \mathcal{Geo}(q)$ avec $q = 1 - \psi(0)$. En remarquant que les variables $(L_k)_{1 \leq k \leq K}$ sont iid (F), on a alors $\psi(k) = \mathbb{P}[M > \kappa]$ donnée par la formule de Pollaczek-Khinchine-Beekman.

La représentation

$$\psi(\kappa) = \mathbb{P}\left[\sum_{j=1}^K L_j > \kappa\right]$$

permet d'évaluer la probabilité de ruine sur horizon infini à l'aide de l'algorithme de Panjer.

111 – Schémas de cession/rétrocession



112 – Les mots clés de la réassurance

Cédante : client du réassureur, c'est-à-dire l'assureur direct, qui transfère (cède) des risques au réassureur contre le versement d'une **prime de réassurance**.

Cession : transfert de risques par l'assureur direct au réassureur.

Capacité (Value Exposure) : limite du montant du risque couvert par un contrat de (ré)assurance.

Réassurance proportionnelle : participation proportionnelle du réassureur aux primes et aux sinistres de l'assureur direct.

Réassurance en quote-part (Quota Share) : type de réassurance proportionnelle où le réassureur participe à un pourcentage donné de tous les risques souscrits par un assureur direct dans une branche déterminée.

Réassurance en excédent de plein (Surplus Share) : type de réassurance proportionnelle où le réassureur couvre les risques au delà du plein de conservation de l'assureur direct. Ce ratio se calcule sur la capacité du risque souscrit (\approx Sinistre maximal possible).

Commission de réassurance : rémunération que le réassureur accorde à l'assureur ou aux courtiers en dédommagement des frais d'acquisition et de gestion des contrats d'assurance.

Réassurance non proportionnelle (ou réassurance en excédent de sinistre, Excess Reinsurance) : prise en charge par le réassureur des sinistres excédant un certain montant, contre le versement par l'assureur direct d'une prime de réassurance spécifique.

Rétrocession : part des risques que le réassureur cède à d'autres réassureurs.

Coassurance : participation de plusieurs assureurs directs au même risque.

On utilise alors l'expression **pool de réassurance**. Le réassureur principal est appelé **apériteur**.

Traité de réassurance : contrat conclu entre l'assureur direct et le réassureur sur un ou plusieurs portefeuilles de l'assureur.

Réassurance facultative : Elle diffère du traité de réassurance par une souscription risque par risque (ou police par police) (du cas par cas, un risque à la fois).

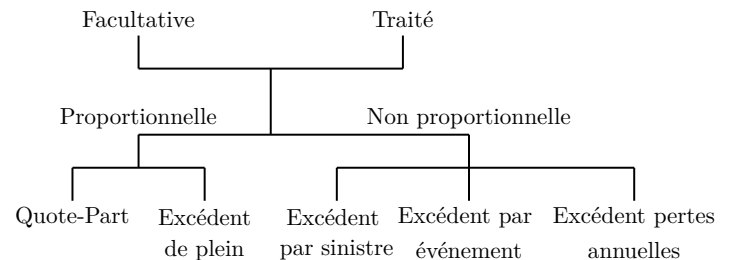
113 – Le rôle économique de la réassurance

L'assurance et la réassurance partagent la même finalité : la mutualisation des risques, en particuliers les risques :

- indépendants, mais unitairement coûteux (avion, navire, sites industriels...),
- de petits montants (bris, auto, ...) mais corrélés lors d'événements de grandes ampleurs, engendrant des cumuls onéreux,
- agrégés au sein d'un portefeuille de polices d'assurance,
- mal connus ou nouveaux.

La réassurance permet d'augmenter la capacité d'émission d'affaires, assurer la stabilité financière de l'assureur, surtout en cas de catastrophes, réduire leur besoin en capital, bénéficier de l'expertise du réassureur.

114 – Les types d'ententes en réassurance

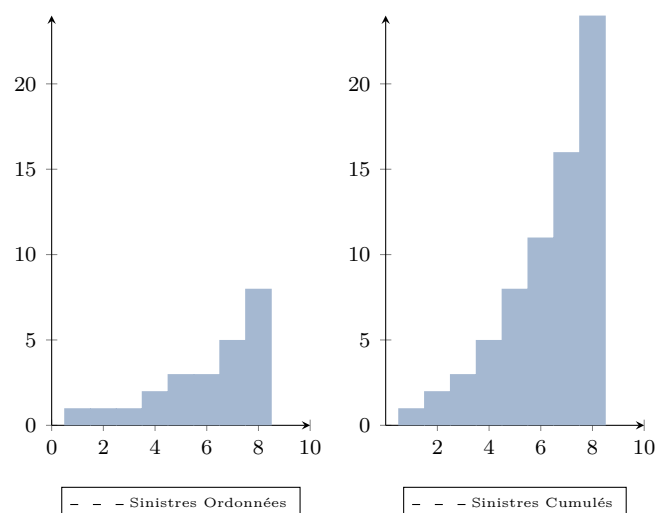


115 – Les types de réassurance à travers un exemple

Notre assureur réassure $N = 30$ polices d'assurance, d'un total de primes est de 10M€ ($P = \sum_{i=1, \dots, 30} P_i$). La capacité totale est de 180M€ ($\sum_{i=1, \dots, 30} K_i$). S_r sera la part de sinistre pris en charge par l'assureur et P_r la prime de réassurance. Voici les polices sinistrées :

Num de sinistre	1	2	3	4	5	6	7	8
Prime (k€)	500	200	100	100	50	200	500	200
Capacité (M€)	8	5	3	2	3	5	8	8
Sinistres (M€)	1	1	1	2	3	3	5	8

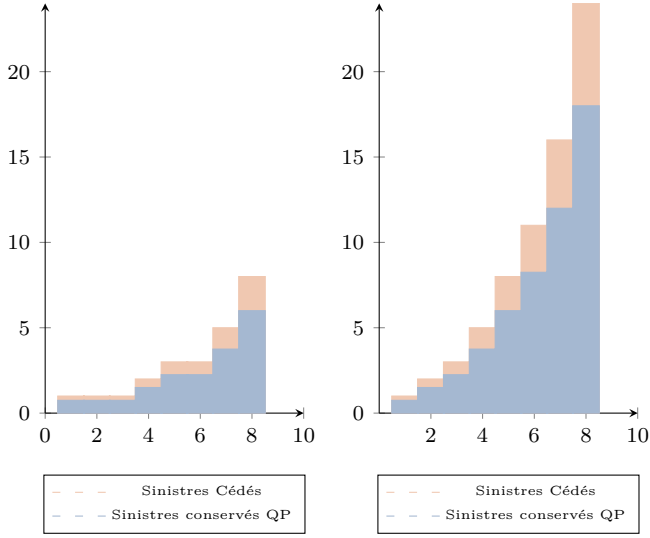
Le S/P est de 240%.



Quote-part :

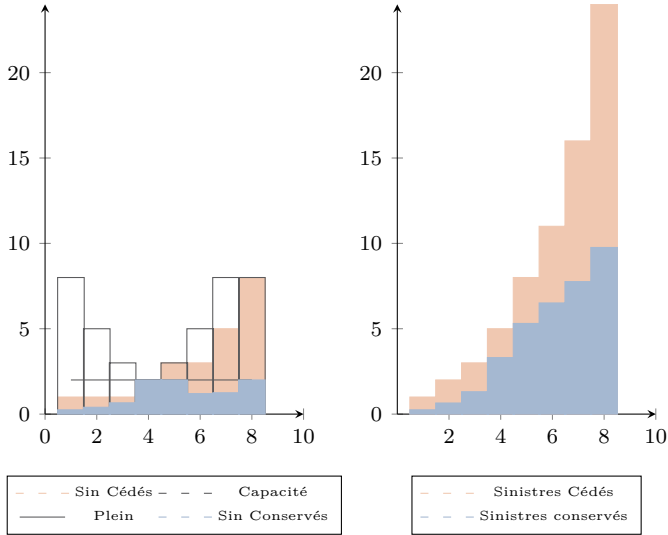
$$S_r = \alpha \sum_{i=1 \dots N} S_i \quad P_r = \alpha \sum_{i=1 \dots N} P_i$$

où $\alpha \in [0, 1]$ (25% dans la figure) est la part cédée en Quote-part.



Excédent de plein, le plein est noté K (2M€ dans l'exemple), α_i représente le taux de cession de la police i .

$$S_r = \sum_{i=1 \dots N} \underbrace{\left(\frac{(K_i - K)_+}{K_i} \right)}_{\alpha_i} S_i \quad P_r = \sum_{i=1 \dots N} \left(\frac{(K_i - K)_+}{K_i} \right) P_i$$

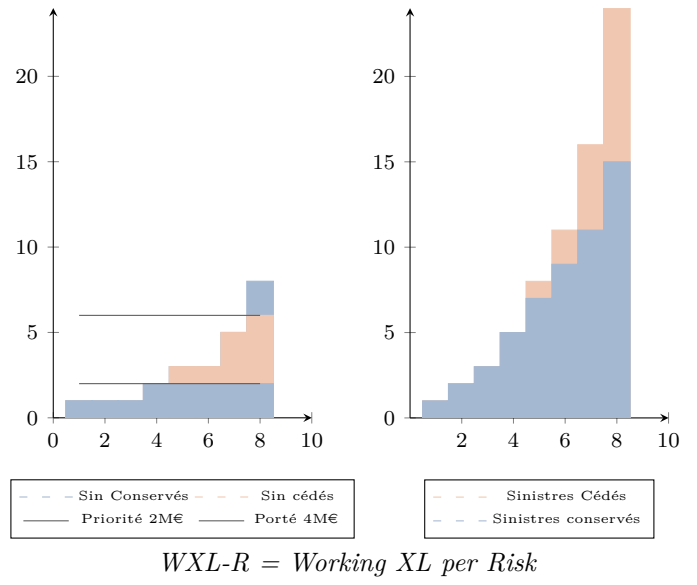


Excédent par sinistre

L'assureur fixe la priorité a et la portée b (respectivement 2M€ et 4M€ dans la figure).

$$S_r = \sum_{i=1 \dots N} \min \left((S_i - a)^+, b \right)$$

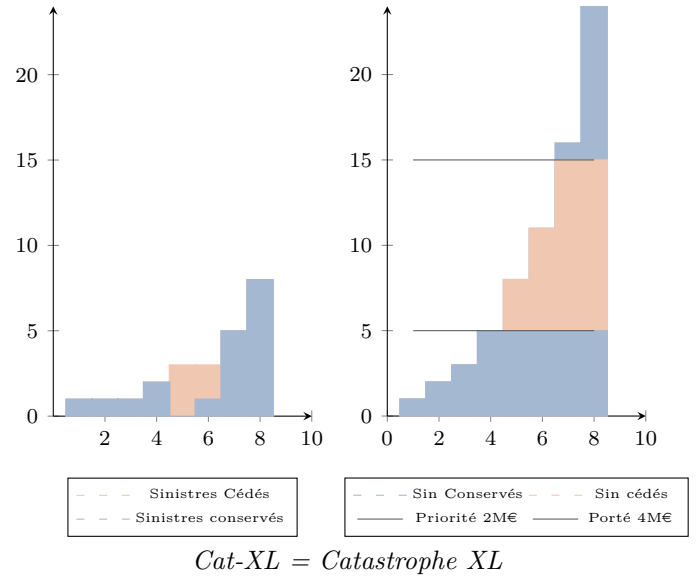
La prime est fixée par le réassurance, en fonction de son estimation de $\mathbb{E}[S_r]$.



Excédent par événement

$$S_r = \sum_{i=1 \dots N} \min \left(\mathbb{1}_{i \in \text{Cat}_j} \times (S_i - a)^+, b \right)$$

Dans l'illustration, les sinistres font référence à un seul événement, avec une priorité à 5M€ et une portée à 10M€.



Excédent de pertes annuelles

Cette réassurance (*Stop Loss*) intervient lorsque le cumul des pertes annuelles est dégradé. Il s'exprime sur la base du ratio S/P avec une priorité et une portée du XL exprimées en %.

$$S_r = \min \left(\left(\sum_{i=1 \dots N} S_i - aP \right)^+, bP \right)$$

116 – Les principales clauses en réassurance

La franchise a^{ag} et la **limite aggregate b^{ag}** s'appliquent après le calcul du S_r .

$$S_r^{ag} = \min \left((S_r - a^{ag})^+, b^{ag} \right)$$

L'objectif de la **clause d'indexation** est de conserver les modalités du traité sur plusieurs exercices successifs. Les bornes du traité s'alignent sur un indice économique (salaire, devise, indice de prix ...).

Avec **clause de stabilisation**, lorsque le sinistre souffre d'un règlement est long, voire très long (au moins ≥ 1 an), les bornes du traité sont actualisées dans le calcul du S_r afin que les parts respectives du réassureur et de la cédante prévues initialement soient globalement respectées.

Avec la **clause de partage des intérêts**, si lors d'une transaction ou d'un jugement d'un tribunal une distinction a été faite entre l'indemnité et les intérêts, les intérêts courus entre la date du sinistre et celle du paiement effectif de l'indemnité seront répartis entre la cédante et le réassureur proportionnellement à leur charge respective résultant de l'application du traité hors intérêts.

La **clause de reconstitution de garantie** concerne uniquement les *traités en excédent de sinistre par risque ou par événement* qui pourraient être déclenchée à plusieurs reprises dans l'année. Le réassureur limite sa prestation à N fois la portée de l' XS , contre le versement d'une prime complémentaire. La reconstitution peut se faire au prorata temporis (temps qui reste à courir jusqu'à la date d'échéance du traité) ou au prorata des capitaux absorbés, ou les deux (double proata).

La **clause de superposition** (*Interlocking Clause*) est utilisée dans les traités en XS par événement, qui fonctionnent par exercice de souscription et non par exercice de survenance. La clause de superposition qui aura pour effet de recalculer les bornes du traité, parce qu'un même événement peut déclencher le traité des souscriptions n et $N - 1$.

117 – La réassurance publique

La **Caisse Centrale de Réassurance (CCR)** propose, avec la garantie de l'État, des couvertures illimitées pour des branches spécifiques au marché français.

- les risques exceptionnels liés à un transport,
- la RC des exploitants de navires et installations nucléaires,
- les risques de catastrophes naturelles,
- les risques d'attentats et d'actes de terrorisme,

— le Complément d'Assurance crédit Public (CAP).

Elle gère également pour le compte de l'État certains Fonds Publics, en particulier le régime Cat Nat.

Également, le **GAREAT** est un Groupement d'Intérêt Économique (GIE) à but non lucratif, mandaté par ses adhérents, qui gère la réassurance des risques d'attentats et actes de terrorisme avec le soutien de l'État via la CCR.

118 – Titrisation / CatBonds

Pourquoi ? Les capacités financières de tous les assureurs et réassureurs réunies ne couvrent pas les dégâts d'un tremblement de terre majeur aux État Unis (≥ 200 Md€). Cette somme correspond à moins de 1% de la capitalisation sur les marchés financiers américains.

La **titrisation** transforme un risque assurantiel en titre négociable, souvent en titres obligataires appelés **Cat-Bonds**. Elle consiste en un échange de principal contre paiement périodique de coupons, dans lequel le paiement des coupons et ou le remboursement du principal sont conditionnés à la survenance d'un événement déclencheur défini *a priori*. Les taux de ces obligations sont majorés en fonction du risque, non pas de défaillance ou de contre partie, mais de la survenance de l'événement (inférieur à 1%). La structure dédiée à cette transformation s'appelle *Special Purpose Vehicle* (SPV).

Le déclencheur peut être liés directement aux résultats de la cédente (Indemnitaire), dépendre d'un indice de sinistralité, d'un paramètre mesurable (somme des excédents de pluie, échelle de Richter, taux de mortalité), ou d'un modèle (RMS & Equecat Storm modelling).

Critère	Indemnitaire	Indice	Paramétrique	Modèle
Transparence	⊖	⊕	⊕	⊕
Risque de base	⊕	⊖	⊖	⊕
Aléa moral	⊖	⊕	⊕	⊕
Universalité des périls	⊕	⊕	⊖	⊕
Délai de déclenchement	⊖	⊖	⊕	⊕

Statistiques des extrêmes et tarification en réassurance

119 – Loi de Pareto

Soit la variable aléatoire X qui suit une loi de Pareto de paramètres (x_m, k) , k est l'indice de Pareto :

$$\mathbb{P}(X > x) = \left(\frac{x}{x_m}\right)^{-k} \text{ avec } x \geq x_m$$

$$f_{k,x_m}(x) = k \frac{x_m^k}{x^{k+1}} \text{ pour } x \geq x_m$$

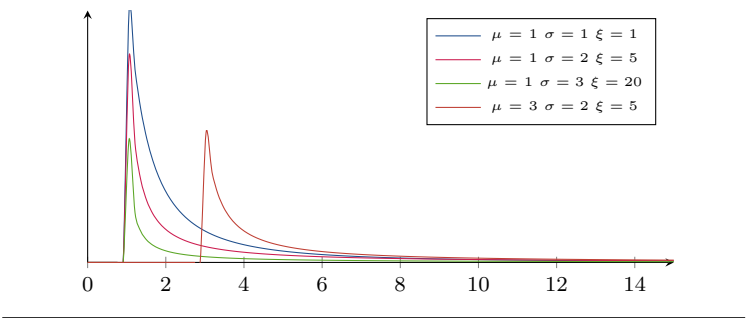
Loi de Pareto généralisée (GPD) a 3 paramètres μ , σ et ξ .

$$F_{\xi,\mu,\sigma}(x) = \begin{cases} 1 - \left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{-1/\xi} & \text{for } \xi \neq 0, \\ 1 - \exp\left(-\frac{x-\mu}{\sigma}\right) & \text{for } \xi = 0. \end{cases}$$

pour $x \geq \mu$ quand $\xi \geq 0$ et $\mu \leq x \leq \mu - \sigma/\xi$ quand $\xi < 0$ et où $\mu \in \mathbb{R}$ est la localisation, $\sigma > 0$ l'échelle et $\xi \in \mathbb{R}$ la forme. No-

tez que certaines références donnent le « paramètre de forme », comme $\kappa = -\xi$.

$$f_{\xi,\mu,\sigma}(x) = \frac{1}{\sigma} \left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{\left(-\frac{1}{\xi}-1\right)} = \frac{\sigma^{\frac{1}{\xi}}}{(\sigma + \xi(x-\mu))^{\frac{1}{\xi}+1}}$$



120 – Loi des valeurs extrême généralisée

La fonction de répartition de la loi des extrêmes généralisée est

$$F_{\mu,\sigma,\xi}(x) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}$$

pour $1 + \xi(x - \mu)/\sigma > 0$, où $\mu \in \mathbb{R}$ est la localisation, $\sigma > 0$ d'échelle et $\xi \in \mathbb{R}$ la forme. Pour $\xi = 0$ l'expression est définie par sa limite en 0.

$$f_{\mu,\sigma,\xi}(x) = \frac{1}{\sigma} \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{(-1/\xi)-1} \times \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}$$

$$f(x; \mu, \sigma, 0) = \frac{1}{\sigma} \exp \left(-\frac{x - \mu}{\sigma} \right) \exp \left[-\exp \left(-\frac{x - \mu}{\sigma} \right) \right]$$

121 – Loi de Gumbel

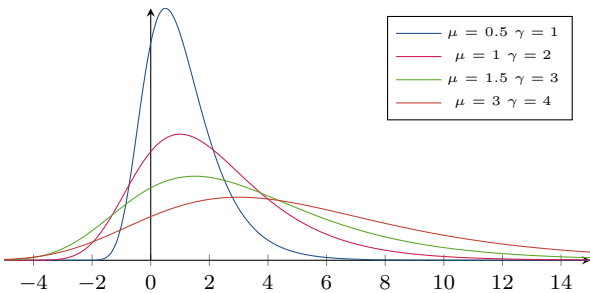
La fonction de répartition de la loi de Gumbel est :

$$F_{\mu,\sigma}(x) = e^{-e^{-(x-\mu)/\sigma}}.$$

Pour $\mu = 0$ et $\sigma = 1$, on obtient la loi standard de Gumbel. La loi de Gumbel est un cas particulier de la GEV (avec $\xi = 0$).

Sa densité :

$$f_{\mu,\sigma}(x) = \frac{1}{\sigma} e^{\left(\frac{x-\mu}{\sigma} - e^{-(x-\mu)/\sigma} \right)}$$



122 – Loi de Weibull

La loi de Weibulla pour fonction de répartition est définie par :

$$F_{\alpha,\mu,\sigma}(x) = 1 - e^{-((x-\mu)/\sigma)^\alpha}$$

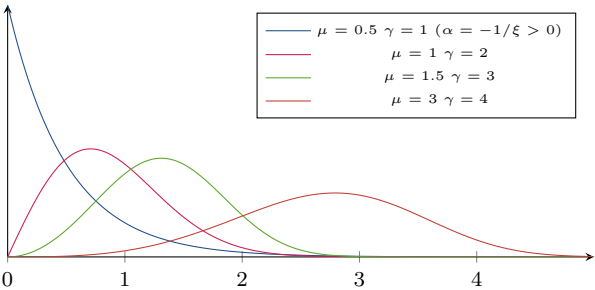
où $x > \mu$. Sa densité de probabilité est :

$$f_{\alpha,\mu,\sigma}(x) = (\alpha/\sigma) ((x - \mu)/\sigma)^{(\alpha-1)} e^{-((x-\mu)/\sigma)^\alpha}$$

où $\mu \in \mathbb{R}$ est la localisation, $\sigma > 0$ d'échelle et $\alpha = -1/\xi > 0$ la forme.

La distribution de Weibull est souvent utilisée dans le domaine de l'analyse de la durée de vie. C'est un cas particulier de la GEV lorsque $\xi < 0$.

Si le taux de pannes diminue au cours du temps alors, $\alpha < 1$. Si le taux de panne est constant dans le temps alors, $\alpha = 1$. Si le taux de panne augmente avec le temps alors, $\alpha > 1$. La compréhension du taux de pannes peut fournir une indication au sujet de la cause des pannes.



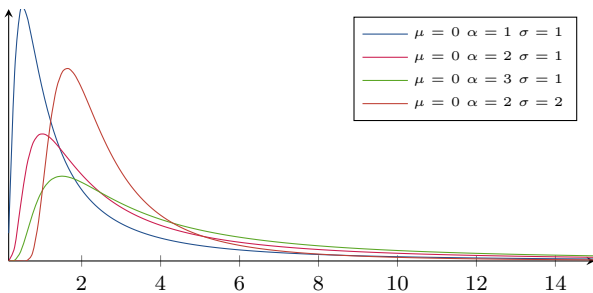
123 – Loi de Fréchet

Sa fonction de répartition de la loi de Fréchet est donnée par

$$F_{\alpha,\mu,\sigma}(x) = \mathbb{P}(X \leq x) = \begin{cases} e^{-\left(\frac{x-\mu}{\sigma}\right)^{-\alpha}} & \text{si } x > \mu \\ 0 & \text{sinon.} \end{cases}$$

où $\mu \in \mathbb{R}$ est la localisation, $\sigma > 0$ l'échelle et $\alpha = 1/\xi > 0$ la forme. C'est un cas particulier de la GEV lorsque $\xi > 0$.

$$f_{\alpha,\mu,\sigma}(x) = \frac{\alpha}{\sigma} \left(\frac{x - \mu}{\sigma} \right)^{-1-\alpha} e^{-\left(\frac{x-\mu}{\sigma}\right)^{-\alpha}}$$



124 – Lien entre GEV, Gumbel, Fréchet et Weibull

Le paramètre de forme ξ gouverne le comportement de la queue de distribution. Les sous-familles définies par $\xi = 0$, $\xi > 0$ et $\xi < 0$ correspondent respectivement aux familles de Gumbel, Fréchet et Weibull :

- Gumbel ou loi des valeurs extrêmes de type I
- Fréchet ou loi des valeurs extrêmes de type II, si $\xi = \alpha^{-1}$ avec $\alpha > 0$,
- Reversed Weibull (\bar{F}) ou loi des valeurs extrêmes de type III, si $\xi = -\alpha^{-1}$, avec $\alpha > 0$.

125 – Théorème général des valeurs extrêmes

Soit X_1, \dots, X_n iid, X de fonction de répartition F_X et soit $M_n = \max(X_1, \dots, X_n)$.

La théorie donne la distribution exacte du maximum :

$$\begin{aligned} \mathcal{P}(M_n \leq z) &= \Pr(X_1 \leq z, \dots, X_n \leq z) \\ &= \mathcal{P}(X_1 \leq z) \cdots \mathcal{P}(X_n \leq z) = (F_X(z))^n. \end{aligned}$$

S'il existe une séquence de paire de nombres réels (a_n, b_n) de telle sorte que $a_n > 0$ et $\lim_{n \rightarrow \infty} \mathcal{P} \left(\frac{M_n - b_n}{a_n} \leq x \right) = F_X(x)$, où F_X est une fonction de répartition non dégénérée, alors la limite de la fonction F_X appartient à la famille des lois GEV.

126 – Densité sous-exponentielle

Cas des puissances

Si $\bar{F}_X(x) = \mathbb{P}(X > x) \sim c x^{-\alpha}$ quand $x \rightarrow \infty$ pour un $\alpha > 0$ et une constante $c > 0$ alors la loi de X est sous-exponentielle.

Si F_X est une fonction de répartition continue d'espérance $\mathbb{E}[X]$ finie, on appelle l'indice des grands risques par

$$D_{F_X}(p) = \frac{1}{\mathbb{E}[X]} \int_{1-p}^1 F_X^{-1}(t) dt, \quad p \in [0, 1]$$

Cette distribution en excès décroît moins vite que n'importe quelle distribution exponentielle. Il est possible de considérer cette statistique :

$$T_n(p) = \frac{X_{(1:n)} + X_{(2:n)} + \dots + X_{(np:n)}}{\sum_{1 \leq i \leq n} (X_i)} \text{ où } \frac{1}{n} \leq p \leq 1$$

$X_{(i:n)}$ désigne le i^e max des X_i .

127 – Théorème de Pickands–Balkema–de Haan (loi des excès)

Soit X de distribution F_X , et soit u un seuil élevé. Alors, pour une large classe de lois F_X , la loi conditionnelle des excès

$$X_u := X - u \mid X > u$$

est approximable, pour u suffisamment grand, par une loi de Pareto généralisée (GPD) :

$$\mathbb{P}(X - u \leq y \mid X > u) \approx G_{\xi, \sigma, \mu=0}(y) := 1 - \left(1 + \frac{\xi(x - \mu)}{\sigma}\right)^{-1/\xi}$$

$y \geq 0$. Autrement dit, pour $u \rightarrow x_F := \sup\{x : F(x) < 1\}$,

$$\sup_{0 \leq y < x_F - u} |\mathbb{P}(X - u \leq y \mid X > u) - G_{\xi, \sigma, \mu=0}(y)| \rightarrow 0.$$

Ce théorème justifie l'utilisation de la *loi de Pareto (généralisée)* pour modéliser les excès au-delà d'un seuil, ce qui est précisément le cadre des traités de réassurance en *excess of loss*.

128 – Les données en réassurance

Comme la réassurance indemnise des agrégations de sinistre ou des sinistres extrêmes, elle utilise souvent des historiques qui devront être utilisés avec prudence :

- l'actualisation des données (impact de l'inflation monétaire).
- la revalorisation prend en compte l'évolution du risque :
 - l'évolution des taux de prime, garanties et modalités des contrats,
 - l'évolution des coûts des sinistres (indice des coûts de la construction, indices des coûts de réparation automobile, ...)
 - l'évolution de l'environnement juridique.
- le redressement de la statistique pour prendre en compte l'évolution de la base portefeuille :
 - profil des polices (nombre, capitaux, ...),
 - natures des garanties (évolution des franchises, des exclusions, ...)

Après ces corrections, les données sont dites « as if » (en économie, on utilise l'expression contre-factuel)

129 – La prime *Burning Cost*

x_i^j désigne le i^e sinistre de l'année j « as if » actualisé, revalorisé et redressé, n^j le nombre de sinistres l'année j , c^j la charge de l'assureur c . Le taux pur par la méthode de **Burning Cost** est donné par la formule :

$$BC_{pur} = \frac{1}{s} \sum_{j=1}^n \frac{c^j}{a_j}$$

Le Burning Cost n'est qu'une moyenne des ratios S/P croisés : les sinistres à la charge du réassureur sur les primes reçues par la cédante. La prime Burning Cost est alors : $P_{pure} = BC_{pur} \times a_{s+1}$.

Dans le cas d'un p XS f),

$$c^j = \sum_{i=1}^{n^j} \max\left(\left(x_i^j - f\right), p\right) \mathbb{1}_{x_i^j \geq f}$$

Si l'assurance vie calcule des taux de prime en référence au capital, l'assurance non vie utilise comme référence à la valeur assurée, la réassurance prend elle comme référence le total des primes de la cédante, appelée **assiette**. On note a_j désigne l'assiette de prime à l'année j et a_{s+1}^* désigne l'assiette estimée de l'année à venir et où s désigne le nombre d'années d'historique.

130 – Le modèle Poisson-Pareto

[Prime de l'XS ou de l'XL] Soit p et f respectivement la portée et la priorité (franchise) de l'XS, avec la limite $l = p + f$ (p XS f).

La prime XS correspond à :

$$\mathbb{E}[S_N] = \mathbb{E}\left[\sum_{i=1}^N Y_i\right] = \mathbb{E}[N] \times \mathbb{E}[Y]$$

où

$$\mathbb{E}[Y] = l\mathbb{P}[X > l] - f \times \mathbb{P}[X \geq f] + \mathbb{E}[X \mid f \leq X \leq l]$$

Si $l = \infty$ et $\alpha \neq 1$:

$$\mathbb{E}[S_N] = \lambda \frac{x_m^\alpha}{\alpha - 1} f^{1-\alpha}$$

si $l = \infty$ et $\alpha = 1$ il n'y a pas de solution.

Si $l < \infty$ et $\alpha \neq 1$:

$$\mathbb{E}[S_N] = \lambda \frac{x_m^\alpha}{\alpha - 1} (f^{1-\alpha} - l^{1-\alpha})$$

Si $l < \infty$ et $\alpha = 1$:

$$\mathbb{E}[S_N] = \lambda x_m \ln\left(\frac{1}{f}\right)$$

131 – Le modèle Poisson-LogNormal

Si X suit une $LN(x_m, \mu, \sigma)$ alors $X - x_m$ suit une $LN(\mu, \sigma)$ Il vient :

$$\mathbb{P}[X > f] = \mathbb{P}[X - x_m > f - x_m] = 1 - \Phi\left(\frac{\ln(f - x_m) - \mu}{\sigma}\right)$$

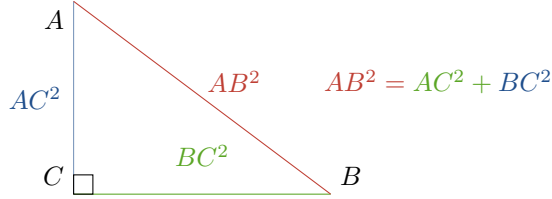
$$\begin{aligned} \mathbb{E}[X \mid X > f] &= \mathbb{E}[X - x_m \mid X - x_m > f - x_m] + x_m \mathbb{P}[X > f] \\ &= e^{m+\sigma^2/2} \left[1 - \Phi\left(\frac{\ln(f - x_m) - (\mu + \sigma^2)}{\sigma}\right) \right] \\ &\quad + x_m \left(1 - \Phi\left(\frac{\ln(f - x_m) - \mu}{\sigma}\right) \right) \end{aligned}$$

Avec franchise et sans limite :

$$\begin{aligned} \mathbb{E}[S_N] &= \lambda (\mathbb{E}[X - x_m \mid X - x_m > f - x_m] + x_m \mathbb{P}[X > f] - f \mathbb{P}[X > f]) \\ &= \lambda \left(e^{m+\sigma^2/2} \left[1 - \Phi\left(\frac{\ln(f - x_m) - (\mu + \sigma^2)}{\sigma}\right) \right] \right) \\ &\quad + \lambda (x_m - l) \left(1 - \Phi\left(\frac{\ln(f - x_m) - \mu}{\sigma}\right) \right) \end{aligned}$$

132 – Pythagore

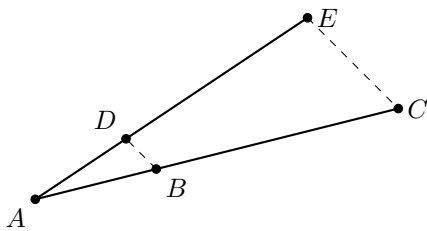
Dans un triangle rectangle, le carré de l'hypoténuse est égal à la somme des carrés des deux autres côtés. Si ABC est rectangle en C , alors



133 – Thalès

Soient deux droites **sécantes en un point** A , et soient deux droites (BC) et (DE) **parallèles**, coupant les deux droites en B, D et C, E , alors :

$$\frac{AB}{AD} = \frac{AC}{AE} = \frac{BC}{DE}$$



134 – Équation du second degré

$$ax^2 + bx + c = 0$$

Le discriminant est défini par :

$$\Delta = b^2 - 4ac$$

— Si $\Delta > 0$, l'équation a deux solutions distinctes :

$$x_1 = \frac{-b + \sqrt{\Delta}}{2a}, \quad x_2 = \frac{-b - \sqrt{\Delta}}{2a}$$

— Si $\Delta = 0$, l'équation a une solution double :

$$x = \frac{-b}{2a}$$

— Si $\Delta < 0$, l'équation a une solution dans les imaginaires

$$x_1 = \frac{-b + i\sqrt{\Delta}}{2a}, \quad x_2 = \frac{-b - i\sqrt{\Delta}}{2a}$$

135 – Fonctions factorielle, dénombrement et Gamma

La fonction **factorielle** (de \mathbb{N} dans \mathbb{N}) est définie par $0! = 1$ et $n! = n \times (n-1) \times \dots \times 2 \times 1 =$ permutations de n éléments

$C_n^k = \binom{n}{k} = \frac{n!}{k!(n-k)!} =$ choix de k éléments parmi n les C_n^k se calculent aussi par le triangle de Pascal et vérifient :

$$C_n^k = C_n^{n-k}, C_n^k + C_n^{k+1} = C_{n+1}^{k+1}.$$

Soit E un ensemble de cardinal $\text{Card}(E)$ et de parties $\mathcal{P}(E)$:

$$\text{Card}(\mathcal{P}(E)) = 2^{\text{Card}(E)}$$

$$\text{Card}(A \times B) = \text{Card}(A) \times \text{Card}(B)$$

$$\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B)$$

$$\Gamma(n) = \int_0^\infty t^{n-1} e^{-t} dt$$

La fonction Γ peut être vue comme le prolongement de la factorielle : $\Gamma(n+1) = n!$.

136 – Développement binomial

Pour un entier positif n ,

$$(x+y)^n = \sum_{i=0}^n \binom{n}{i} x^i y^{n-i}$$

137 – Suites

Les suites arithmétiques de raison r

$$\begin{cases} u_{n+1} = u_n + r \\ u_0 \in \mathbb{R} \end{cases} \Rightarrow \begin{cases} u_n = nr + u_0 \\ \sum_{k=0}^n u_k = \frac{(n+1)(2u_0 + nr)}{2} \end{cases}$$

suites géométriques de raison q $\begin{cases} u_{n+1} = q \times u_n \\ u_0 \in \mathbb{R} \end{cases}$

$$\Rightarrow \begin{cases} u_n = u_0 \times q^n \\ \sum_{k=0}^n u_k = \begin{cases} (n+1)u_0 & \text{si } q = 1 \\ u_0 \frac{1-q^{n+1}}{1-q} & \text{sinon} \end{cases} \end{cases}$$

138 – Exponentielle et Logarithme

La fonction exponentielle e^x peut être définie par le développement en série entière suivant :

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

Cette série converge pour tout $x \in \mathbb{R}$ et permet de définir l'exponentielle comme une somme infinie.

La fonction logarithme naturel $\ln(x)$ est définie comme la primitive de la fonction $\frac{1}{x}$. Autrement dit :

$$\frac{d}{dx} \ln(x) = \frac{1}{x}$$

avec la condition $\ln(1) = 0$. Cette définition permet d'établir le lien entre l'exponentielle et le logarithme via l'inversion : $e^{\ln(x)} = x$ pour $x > 0$.

139 – Relation de congruence

Soit $m > 0$. On dit que deux réels a et b sont congrus modulo m s'il existe un entier relatif $k \in \mathbb{Z}$ tel que :

$$a = b + km.$$

On note $a \equiv b \pmod{m}$.

En trigonométrie, on choisit souvent $m = 2\pi$ ou $m = \pi$.

Soient $m > 0$ et $a, b, c, d \in \mathbb{R}$. Alors :

- **Réflexivité** : $a \equiv a \pmod{m}$.
- **Symétrie** : $a \equiv b \pmod{m} \iff b \equiv a \pmod{m}$.
- **Transitivité** : si $a \equiv b \pmod{m}$ et $b \equiv c \pmod{m}$, alors $a \equiv c \pmod{m}$.
- **Additivité** : si $a \equiv b \pmod{m}$ et $c \equiv d \pmod{m}$, alors $a + c \equiv b + d \pmod{m}$.

142 – Dérivées et primitives

f continue en $x_m \iff \lim_{x \rightarrow x_m} f(x) = f(x_m)$

f dérivable en $x_m \iff \exists \lim_{h \rightarrow 0} \frac{f(x_m + h) - f(x_m)}{h} =: f'(x_m)$

L'**intégrale de Riemann** d'une fonction $f(x)$ sur un intervalle $[a, b]$ est la limite, si elle existe, de la somme des aires des rectangles approchant l'aire sous la courbe, donnée par :

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i^*) \Delta x_i,$$

- $[x_{i-1}, x_i]$ est une subdivision de $[a, b]$,
- $\Delta x_i = x_i - x_{i-1}$ est la largeur du sous-intervalle,
- $x_i^* \in [x_{i-1}, x_i]$ est un point choisi arbitrairement dans chaque sous-intervalle.

Exemple d'intégrale de Riemann (upper)*

L'**intégrale de Lebesgue** d'une fonction $f(x)$ sur un ensemble E est définie en mesurant l'aire sous la courbe en fonction des valeurs prises par f , donnée par :

$$\int_E f d\mu = \int_0^\infty \mu(\{x \in E : f(x) > t\}) dt,$$

- μ est une mesure (souvent la mesure de Lebesgue),
- $\{x \in E : f(x) > t\}$ représente l'ensemble des points où $f(x)$ dépasse t .

À la différence de Riemann, Lebesgue regroupe les points selon leurs valeurs plutôt que selon leur position.

fonction ($n \in \mathbb{R}$)	dérivée	primitive
x	1	$\frac{x^2}{2} + C$
x^2	$2x$	$\frac{x^3}{3} + C$
$1/x$	$-1/x^2$	$\ln(x) + C$
$\sqrt{x} = x^{1/2}$	$\frac{1}{2\sqrt{x}}$	$\frac{2}{3}x^{3/2} + C$
$x^n, n \neq -1$	nx^{n-1}	$\frac{x^{n+1}}{n+1} + C$
$\ln(x)$	$1/x$	$x \ln(x) - x + C$
e^x	e^x	$e^x + C$
$a^x = e^{x \ln(a)}$	$\ln(a) \times a^x$	$a^x / \ln(a) + C$
$\sin(x)$	$\cos(x)$	$-\cos(x) + C$
$\cos(x)$	$-\sin(x)$	$\sin(x) + C$
$\tan(x)$	$1 + \tan(x)$	$-\ln(\cos(x)) + C$
$1/(1+x^2)$	$-2x/(1+x^2)^2$	$\arctan(x) + C$

$$\begin{array}{ll}
(u+v)' = u' + v' & \left(\frac{1}{u}\right)' = -\frac{u'}{u^2} \\
(ku)' = ku' & (\ln(u))' = \frac{u'}{u} \\
(u \times v)' = u'v + uv' & (\exp(u))' = \exp(u) \times u' \\
\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2} & (f(u))' = f'(u) \times u' \\
(u^n)' = nu^{n-1} \times u' & (f \circ u)' = (f' \circ u) \times u'
\end{array}$$

143 – Intégration par parties

Soit $u(x)$ et $v(x)$ deux fonctions continûment dérivables sur l'intervalle $[a, b]$, alors

$$\int_a^b u(x)v'(x) dx = [u(x)v(x)]_a^b - \int_a^b u'(x)v(x) dx$$

où :

- $u(x)$ est une fonction dont on connaît la dérivée $u'(x)$,
- $v'(x)$ est une fonction dont on connaît la primitive $v(x)$.

144 – Intégration avec changement de variable

Soit $f(x)$ une fonction continue et $x = \phi(t)$ un changement de variable, où ϕ est une fonction dérivable. Alors :

$$\int_a^b f(x) dx = \int_{\phi^{-1}(a)}^{\phi^{-1}(b)} f(\phi(t))\phi'(t) dt$$

où :

- $x = \phi(t)$ représente le changement de variable,
- $\phi'(t)$ est la dérivée de $\phi(t)$,
- les bornes de l'intégrale sont ajustées en fonction du changement de variable.

145 – Formule de Taylor

Soit $f(x)$ une fonction n - fois dérivable en un point a . Le développement de Taylor de $f(x)$ autour de a est donné par :

$$\begin{aligned}
f(x) = & f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 \\
& + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \mathcal{O}_n(x)
\end{aligned}$$

où :

- $f^{(n)}(a)$ est la n^e dérivée de f évaluée en a ,
- $\mathcal{O}_n(x)$ est le reste du développement de Taylor, représentant l'erreur d'approximation lorsque l'on tronque la série après le terme d'ordre n , avec

$$\lim_{x \rightarrow 0} \frac{\mathcal{O}_n(x)}{x^n} \Rightarrow 0$$

146 – Théorème des valeurs intermédiaires

Soit f une fonction continue sur un intervalle fermé $[a, b]$ et $f(a) \neq f(b)$. Le théorème des valeurs intermédiaires affirme que pour tout réel c compris entre $f(a)$ et $f(b)$, il existe un point $x \in [a, b]$ tel que :

$$f(x) = c$$

Autrement dit, si une fonction est continue sur un intervalle, elle prend toutes les valeurs comprises entre $f(a)$ et $f(b)$ au moins une fois.

147 – Matrices et propriétés

Matrices diagonales : Une matrice est dite diagonale si tous les éléments en dehors de la diagonale principale sont nuls. Pour une matrice $A \in \mathbb{R}^{n \times n}$, cela s'écrit :

$$A = \text{diag}(a_1, a_2, \dots, a_n)$$

où a_i sont les éléments diagonaux.

Matrices triangulaires : Une matrice est triangulaire supérieure si tous les éléments en dessous de la diagonale sont nuls, c'est-à-dire $A_{ij} = 0$ pour $i > j$. Inversement, elle est triangulaire inférieure si $A_{ij} = 0$ pour $i < j$.

148 – Déterminant d'une matrice

Le déterminant d'une matrice carrée $A \in \mathbb{R}^{n \times n}$ est un scalaire, noté $\det(A)$:

- Si A est une matrice carrée $n \times n$, alors A est inversible si et seulement si $\det(A) \neq 0$.
- Le déterminant d'une matrice triangulaire (supérieure ou inférieure) ou d'une matrice diagonale :

$$\det(A) = \prod_{i=1}^n A_{ii}$$

—

$$\det(AB) = \det(A) \cdot \det(B), \quad \det(\lambda B) = \lambda \det(B),$$

—

$$\det(A^T) = \det(A)$$

- Si une matrice A contient deux lignes ou colonnes identiques, alors $\det(A) = 0$.

Calcul du déterminant : Le déterminant d'une matrice 2×2 se calcule simplement par :

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

Pour une matrice 3×3 , il est donné par :

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a(ei - fh) - b(di - fg) + c(dh - eg)$$

Pour des matrices de dimension supérieure, le déterminant peut être calculé par cofacteurs ou via une méthode de réduction (par exemple, la méthode de Gauss).

149 – Inversibilité d’une matrice

Une matrice $A \in \mathbb{R}^{n \times n}$ est inversible s’il existe une matrice A^{-1} telle que :

$$AA^{-1} = A^{-1}A = I_n$$

où I_n est la matrice identité. L’inversibilité d’une matrice est garantie par $\det(A) \neq 0$.

Trace : La trace d’une matrice carrée A , notée $\text{Tr}(A)$, est la somme de ses éléments diagonaux :

$$\text{Tr}(A) = \sum_{i=1}^n A_{ii}$$

Elle représente souvent des grandeurs liées à la somme des valeurs propres d’une matrice.

Décomposition de Cholesky : La décomposition de Cholesky est applicable aux matrices symétriques définies positives. Elle permet de factoriser une matrice $A \in \mathbb{R}^{n \times n}$ en un produit de la forme :

$$A = LL^T$$

où L est une matrice triangulaire inférieure. Cette décomposition est utile dans les calculs numériques et les algorithmes d’optimisation.

150 – Gradient et matrice Hessienne

Pour une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^2 , on définit :

- Le **gradient** $\nabla f(x)$ comme le vecteur des dérivées par-

tielles premières :

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}$$

- La **matrice Hessienne** $\nabla^2 f(x)$ comme la matrice symétrique des dérivées secondes :

$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}$$

151 – Théorème des fonctions implicites

Soit $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction de classe C^1 , et supposons que $F(a^*, b) = 0$ pour un certain couple $(a^*, b) \in \mathbb{R}^2$. Si

$$\frac{\partial F}{\partial y}(a^*, b) \neq 0,$$



alors il existe un réel $h > 0$ et une unique fonction φ , définie sur un voisinage $(a^* - h, a^* + h)$, telle que

$$\varphi(a^*) = b \quad \text{et} \quad \forall x \in (a^* - h, a^* + h), \quad F(x, \varphi(x)) = 0$$

De plus, la fonction implicite φ est de classe C^1 et sa dérivée est donnée par :

$$\varphi'(x) = - \frac{\partial F / \partial x}{\partial F / \partial y} \Big|_{y=\varphi(x)}$$

À propos des auteurs

Martial Phéllippé-Guinvarc’h @   est Actuaire et Maître de conférences à Le Mans Université depuis 2012 et enseigne également à l’EURIA depuis 2003. Il est actuaire (EURIA, 2006), membre de l’institut des actuaires (IA) et membre de l’association des économistes agricoles européens (EAAE). Martial Phéllippé-Guinvarc’h enseigne les statistiques, l’analyse de données, la modélisation actuarielle, la programmation, en particulier en SAS et SAS IML, les mathématiques financières, les marchés des matières premières, la finance de marché et des dérivés, l’actuariat non vie, l’actuariat vie, le provisionnement, la solvabilité, la réassurance, l’économie de l’assurance, la gestion des risques, la gestion actifs passif. Il est SAS Protor, et a obtenu pour le Master MBFA de Le Mans Université le SAS Join certificate **Insurance and Economic Analytics**.
Macelino Moreno Porrás est