Credit Risk, CDOs and Copulas (Draft)

Pierre Brugière

University Paris 9 Dauphine

pierre.brugiere@dauphine.fr

April 2, 2021

Overview

Structural and Reduced Form Models

In this course we consider mainly credit risk analysis for structured financial products such as CDOs but credit risk arises in many different forms

Credit risk problematics are embedded in all types of activities:

- retail banking
- private banking
- commercial and investment banking with corporates
- lending between banks

Credit risk is embedded in a wide variety of financing instruments:

- Bonds, Commercial Papers
- Loans (syndicated, bilateral), Credit Lines
- Project Finance, Structured Finance
- Specialised finance, Equity Financing, Stock loan..etc

Financial Analysis for Credit Risk

Two issues to analyse "Credit Risk" in traditional lending:

- capacity to pay financial flows
- capacity to reimburse debt at maturity or to refinance

To measure the capacity to pay interests people will look usually, amongst other things, at the interest coverage ratio:

$$\textbf{Interest coverage ratio} = \frac{\text{interest expenses on debt}}{\text{Earnings Before Interests and Taxes}}$$

- EBIT: Earnings Before Interests and Taxes
- EBITDA: Earnings Before Interests, Taxes, Depreciation and Amortization

In more structured financing a collateral/security can also be taken into account

It seems Altman was the first to introduce statistical models to predict bankruptcy and to quantify this risk

Altman's (1968) financial **score** based on some financial ratios:

- X_1 Working Capital / Total Assets.
- X_2 Retained Earnings / Total Assets
- X_3 Earnings Before Interest and Taxes / Total Assets
- X_4 Market Value / Book Value of Total Debt
- X_5 Sales / Total Assets (industry dependent)

Reminders:
Working Capital = Current Assets - Current Liabilities
Current Assets = Cash + Account Receivables + Inventories

Altman's Z-Score: with various revisions in 1983 and 1993

Based on historical studies of bankruptcies the Z-score was defined as:
$Z = 0.012X_1 + 0.014X_2 + 0.033X_3 + 0.006X_4 + 0.999X_5$
with the following predictions associated:
if $Z < 1.81$ default within 1 year is predicted
if $1.81 \leq Z \leq 2.67$ no prediction
if $Z > 2.67$ prediction of no default within 1 year

The following results were obtained on the samples tested:

- 90.9% success rate in predicting bankruptcy
- 97% success rate in predicting non-bankruptcy

Some refinements have been done to the model in particular to define a probability of default.

Ohlson (1980) proposed the **LOGIT Model**:
$P(Default|X_1, X_2, \cdots, X_n) = \frac{1}{1+exp(-\beta_1 X_1 - \beta_2 X_2 - \cdots - \beta_n X_n)}$
where the parameters β_i are estimated on a sample by maximizing the likelihood.

Remarks:

- $P(Default|X) = \frac{1}{1+exp(-\langle \beta' X \rangle)} \iff ln(\frac{P(Default|X)}{1-P(Default|X)}) = \langle \beta' X \rangle$
- $f(p) = ln(\frac{p}{1-p})$ is called the logit function
- $g(x) = \frac{1}{1+e^{-x}}$ is called the sigmoid function
- $f = g^{-1}$
- $f(0.5) = 0$

Remarks: Likelihood of a binomial model $\mathcal{B}(p)$:
For a single observation $y_i \in \{0,1\}$ $L(y_i) = p^{y_i}(1-p)^{1-y_i}$

The Likelihood for the defaults or not of p companies according to Ohlson's model is : $\prod_{i=1}^{i=n} p_i^{y_i}(1-p_i)^{1-y_i}$ (with $y_i \in \{0,1\}$)

- the factors X_i which impact the probability of defaults p_i are observed for all these companies
- the Maximum Likelihood method consists in finding the β^i which maximise the likelihood (calibration of the model)
- when considering a new company, the probability of default for this company will be calculated by observing the variables X_i for this company an using the coefficients β^i calculated previously

Altman and Ohlson's models were developed principally for financial analysts with an estimation of the parameters based on bankruptcy historicals.

Since then other models have been developed which:

- are better suited for trading purposes
- are calibrated on corporate bonds and credit derivatives market prices instead of observed bankruptcies
- modelize the credit dynamics

Structural and Reduced Form Models

For these new models two principal types:

- Structural Models: (Black-Scholes(1973), Merton(1974), Leland(1994), Schaeffer(2004))
- Reduced Form Models: (Jarrow-Turnbull(1997), Duffie-Singleton (1999))

Definition: Structural Models - KMV

- model the dynamics of the Assets A_t and Liabilities L_t
- assumption that the default happens iff $A_t < L_t$
- in their simplest form: $A_t = A_0 exp(\sigma W_t - \frac{\sigma^2}{2} t)$ and $L_t = L_0$
- approach adopted by Moody's KMV

Definition: Distance to Default (DTD)

The Distance to Default is the number of standard deviations between the *ln* of the current value of the company's assets (assumed to be normally distributed) and the log of its liabilities.

Example: We assume

- Assets: $A_0 = EUR100$ and $A_T = A_0 exp(rT) exp(\sigma W_T - \frac{\sigma^2}{2} T)$
- volatility of the assets: $\sigma = 25\%$, $r = 0$
- Liabilities: $L_T = L_0 = EUR60$

then: $A_T < L_T \iff \frac{W_T}{\sqrt{T}} > \frac{1}{\sigma\sqrt{T}}[ln(\frac{A_0}{L_0}) - \frac{\sigma^2}{2} T] = DTD + \frac{\sigma}{2}\sqrt{T}$
so here the 1 year Distance to Default is 2.

Reduced Form models

Definition: Reduced Form models

- the instant of default τ is modelised by an **Exponential law** of intensity λ, $\mathcal{E}(\lambda)$, under the risk neutral probability P
- when the intensity is constant :
  $P(\tau > t) = exp(-\lambda t)$
- when the intensity is deterministic but time dependent :
  $P(\tau > t) = exp(-\int_0^t \lambda_s ds)$
- when the intensity is stochastic τ is a **Cox process** :
  $P(\tau > t) = E[exp(-\int_0^t \lambda(s, X_s) ds)]$ where $(X_s)_{s \geq 0}$ is a stochastic process

Remark:
(X_s) can be taken as the short term interest rate if it is considered that monetary policy has a significant impact on the economy

Proposition : variational definitions of exponential laws

Exponential laws can be defined equivalently if the following ways:

- $\tau \sim \mathcal{E}(\lambda) \Longleftrightarrow P(t < \tau < t + dt | \tau > t) = \lambda dt$
- $\tau \sim \mathcal{E}(\lambda_t) \Longleftrightarrow P(t < \tau < t + dt | \tau > t) = \lambda_t dt$
- $\tau \sim \mathcal{E}(\lambda(t, X_t)) \Longleftrightarrow P(t < \tau < t + dt | \tau > t, \{X_s, s \in [0, T]\}) = \lambda(t, X_t) dt$ for all $t < T$

Notation We note $\mathcal{X}_t = \{X_s, s \in [0, t]\}$

Demonstration:

We just need to show the stochastic case as the other cases are particular cases. We note $F(t) = P(\tau > t|\mathcal{X}_T)$.

$P(t < \tau < t + dt|\tau > t, \mathcal{X}_T) = \lambda(t, X_t)dt$

$\Longrightarrow \frac{P(t<\tau<t+dt|\mathcal{X}_T)}{P(\tau>t|\mathcal{X}_T)} = \lambda(t, X_t)dt$

$\Longrightarrow \frac{F(t)-F(t+dt)}{F(t)} = \lambda(t, X_t)dt$

$\Longrightarrow dlnF(t) = -\lambda(t, X_t)dt$

$\Longrightarrow F(t) = F(0)exp(-\int\limits_0^t \lambda(s, X_s)ds)$

but $F(0) = 1$ and $F(t) = E(1_{\tau>t}|\mathcal{X}_T)$

so, $P(\tau > t) = E(1_{\tau>t}) = E\left[E(1_{\tau>t}|\mathcal{X}_T)\right] = E\left[F(t)\right]$

$= E\left[exp(-\int\limits_0^t \lambda(s, X_s)ds)\right]$. Q.E.D.

Reduced Form models: Stochastic Intensity and Interest Rate

In an economy where the instantaneous short term interest rate depends on the factors $(X_s)_{s\geq 0}$ if we note $\beta_t = exp(-\int\limits_0^t r(X_s)ds)$ the actualisation factor then the price of a zero coupon bond of maturity T and nominal 1 with credit risk and zero recovery rate is $E[\beta_T 1_{\tau > T}]$.

Proposition

$$E[\beta_T 1_{\tau > T}] = E\left[exp(-\int_0^T (r(X_s) + \lambda(s, X_s))ds)\right]$$

so, $\lambda(s, X_s)$ is the "instantaneous spread" at time s

Demonstration:

$$E[\beta_T 1_{\tau_\lambda > T}] = E\left[E[\beta_T 1_{\tau_\lambda > T} | \mathcal{X}_T]\right]$$

$= E\left[\beta_T E[1_{\tau_\lambda > T} | \mathcal{X}_T]\right]$ (as β_T is known if $\mathcal{X}_T$ is known)

$$= E\left[\beta_T exp(-\int_0^T \lambda(s, X_s)ds)\right]$$

$$= E\left[exp(-\int_0^T r(X_s)ds) exp(-\int_0^T \lambda(s, X_s)ds)\right]$$

$= E\left[exp(-\int_0^T (r(X_s) + \lambda(s, X_s))ds)\right]$ Q.E.D.

Corollary

If the short term interest rate r and the intensity of default λ are constant then the price of a zero coupon bond of maturity T and nominal 1 with credit risk and zero recovery rate is $E[\beta_T 1_{\tau > T}] = exp(-(r + \lambda)T)$.
As a consequence λ can be inferred from the price of a risky bond.

Exemple :
If we assume that a one year zero coupon government bond is worth 100.10 % and that a one year zero coupon bond issued by risky issuer Zco is worth 99.50% then the intensity of default λ for the risky issuer is $ln(\frac{100.10}{99.50}) = 0.6\%$

Remarks:

- a time dependent λ_t enables to calibrate a model to a term structure of spreads
- as we will show later a stochastic $\lambda(t, X_t)$ enables **to modelize correlation between bonds**
- if λ is small then $P(\tau < 1) \sim \lambda$. So if $\lambda = 2\%$ the probability of default within one year is approx 2%
- exponential laws are memoryless i.e : $P(t < \tau < t + \delta | \tau > t) = P(\tau < \delta)$
- exponential laws and normal laws are the two "benchmarks" in finance

Pricing a new Issuance

To price a new issuance several methods can be considered at this stage, amongst them:

- if a rating already exists for the company and if the bond is vanilla price the spread based on this rating, the type of industry and the comparables
- analyse the fundamentals of the company, find a "comparable company" having a similar bond already issued and price by comparison
- analyse the fundamentals of the company and use for example Ohlson's model to calculate a probability of default and from there derive a price for the bond

Pricing a new Issuance

If the bond is complex because :

- there are some conditional payouts
- there are some collaterals which guarantees it
- there are some specific optionalities embedded
- there are some hybrid issues involved

Then it it may be necessary to start with a full modelisation of all the stochastic elements involved before being able to be able to come up with a price.

The example below show how **credit risk modelisation can be embedded in a classic "Black and Scholes" modelisation framework** (which is based on the notion of non arbitrage possibility and risk neutral probability).

Example: Construction of P and λ by arbitrage

Exemple: we consider a two period economy $0, T$ with:

- a risk free asset, a company's stock and bond
- the stock is worth 100 with possible future values 130 and 0 (default)
- the risk free asset has a return of 5%

we assume that the company's bond will be worth at maturity:

- 106 if the company's stock is worth 130
- 84.8 if the company defaults (Recovery Rate 80%)

If there is no arbitrage, we have the following results:
a) the bond can be replicated by investing in the stock and risk-free bond
b) the value of the risky bond is 97.10 today
c) the risk neutral probability verifies $p = 80.77\%$ (probability no default)
d) we have $\lambda = 21.35\%$

Example: Construction of P and λ by arbitrage

Demonstration:

a) we search for a and b such that:
$a105 + b130 = 106$ and $a105 + b \times 0 = 84.8$
So, $a = 0.808$ is the number of risk-free bonds to purchase and $b = 0.163$ is the number of stocks to purchase to replicate the corporate bond.
b) if there is no arbitrage, the price of the risky bond is then
$0.808 \times 100 + 0.163 \times 100 = 97.10$
c) the corresponding risk neutral probability is such that
$\frac{130}{1.05} p + \frac{0}{1.05}(1 - p_1) = 100$ so $p = 80.77\%$.
d) $P(\tau > 1) = p_1 \iff e^{-\lambda} = 19.23\% \iff \lambda = 21.35\%$

Remarks: For the risky-Bond we have $97.10 = \frac{106}{1.05} \times e^{-3.89\%}$ so the return of the bond will be 3.89% higher than the return of the risk-free bond if the bond does not default. This excess return is called the spread of the bond (calculated as a continuous rate).

Remarks:

- when calibrating a Reduced Form model the risk free rate and the price of the risky bond are observed and from there λ can be deducted
- Duffie and others have compared the "implied" λ (under the risk neutral probability) for corporate Bonds derived from their prices and compared them to the "realized" λ (under the "real probability") derived from the defaults over the subsequent periods and found that $\lambda_{implied} \sim 2 \times \lambda_{realized}$
- discrepencies between $\lambda_{implied}$ under the risk neutral possibility and $\lambda_{realized}$ under historical probability can be seen as similar issues to the discrepencies between "implied volatility" and "realized volatility"

Recovery Rate

Theorem and Definition : Recovery Rate R and Spread

The Recovery Rate R is the fraction of the amount due recovered if the counterparty defaults.
In practice R depends on the type of debt issued by the company (senior, junior, secured...)

- if R is the recovery rate of a zero coupon of maturity T
- if r is the risk-free rate for the same maturity
- if S is the spread of the risky bond of maturity T
- if λ is the constant default rate (under the risk neutral probability)

then: $S \sim (1 - R)\lambda$

Demonstration:
Pricing the zero coupon with the risk neutral probability we have:
$e^{-(r+S)T} = e^{-rT}(e^{-\lambda T} + R(1 - e^{-\lambda T})) \Longrightarrow e^{-ST} = (1 - R)e^{-\lambda T} + R$.
Developing to the first order we get the result.

Remark:
In the previous example we have $S = 3.89\%$, $\lambda = 21.35\%$ and the recovery is $R = 80\%$ so $(1 - R)\lambda = 4.27\%$. Here the (first order) approximation of S is not very good because λ is taking a quite large value

Exercise: Construction of a Cox process (and proof of existence)
Let $\lambda(.)$ be positive on $\mathbb{R} \times \mathbb{R}^d$ and
$X = (X_s)_{s \geq 0}$ be a stochastic process of $\mathbb{R}^d$ and $\mathcal{X}_t = \{X_s, s \in [0, t]\}$
Let $\mathcal{E}_1$ be an exponential law of parameter 1 independent from X

Let τ_λ de defined by $\tau_\lambda(\omega) = \inf\{t, \int_0^t \lambda(s, X_s)(\omega) ds \geq \mathcal{E}_1(\omega)\}$

Show that $\tau_\lambda \sim \mathcal{E}(\lambda(s, X_s))$

Demonstration:

$$P(\tau_\lambda > t) = E[1_{\tau_\lambda > t}] = E\Big(E[1_{\tau_\lambda > t}|X]\Big)$$

$$E[1_{\tau_\lambda > t}|X] = P\Big(\mathcal{E}_1 > \int_0^t \lambda(s, X_s)(\omega) ds | X\Big)$$

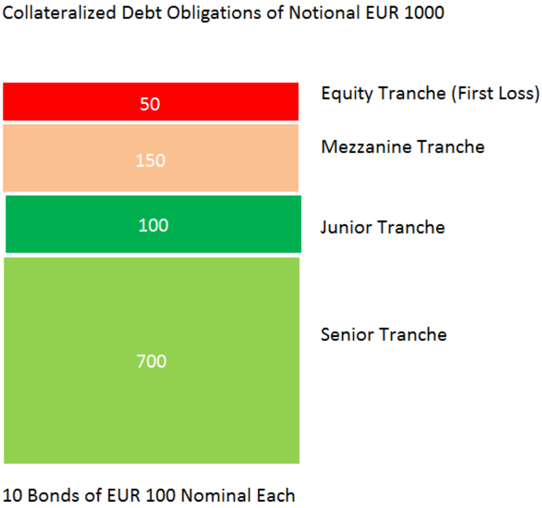
as X and τ_1 are independent

$$P\Big(\tau_1 > \int_0^t \lambda(s, X_s)(\omega) ds | X\Big) = exp\Big(-\int_0^t \lambda(s, X_s)(\omega) ds\Big)$$

so, $P(\tau_\lambda > t) = E\Big[exp\Big(-\int_0^t \lambda(s, X_s)(\omega) ds\Big)\Big]$ Q.E.D.

Collateralized Debt Obligations

Collateralized Debt Obligations



Collateralized Debt Obligations

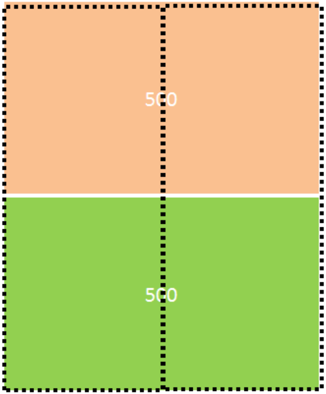
Rationale of the transaction:

- the risk is repackaged to be able to sell it better
- different type of investors can choose between different type of risks
- potentially Rating/ Pricing Arbitrage (up to 2008 too many senior tranches rated AAA)
- in the past potentially regulatory arbitrage (for keeping the risk on the equity tranche and deconsolidating)
- technology of packaging and tranching which can be applied to cash or synthetic underlyings

$\Longrightarrow$ **Important to notice the importance of the "correlation" when pricing a CDO's tranche**

Collateralized Debt Obligations

Example: two Tranches CDO made of two bonds



Bond 1
probability of default p

Bond 2
probability of default q

Case studies:
correl = 1

correl = -1

2 Bonds of EUR 500 Notional Each

Collateralized Debt Obligations

Example: we consider 2 Bonds, with zero Recovery rate, of EUR 500 Nominal each, packaged in a EUR 1000 Notional CDO and note $Z_i = 1$ if the Bond i defaults before maturity and otherwise $Z_i = 0$

a) if we assume that $Z_1 = Z_2$ then:

- either the two bonds default together, resulting in a payout of zero for both tranches or
- none of the bonds defaults, resulting in a payout for both tranches of EUR 500

In this case, both tranches are the same, the senior tranche is not safer than the junior tranche and the correlation between the defaults is 100%.

b) if we assume that $Z_2 = 1 - Z_1$ then:
there is always one bond which defaults so

- the junior tranche has always a payout of zero
- the senior tranche has always a payout of EUR 500

In this case the correlation between the defaults is -100% and the two tranches have extremely different behaviours

$\Longrightarrow$ Note that in this extreme example, the pricing of the two tranches does not depend on the probabilities of default (which nevertheless have to add up to 100% here) but only on the correlation !

Remarks:

- a low correlation between the bonds is good for senior tranche holders and bad for junior tranche holders
- a high correlation between the bonds is good for junior tranche holders and bad for senior tranche holders
- the impact of correlation is less clear for mezzanine tranches holders

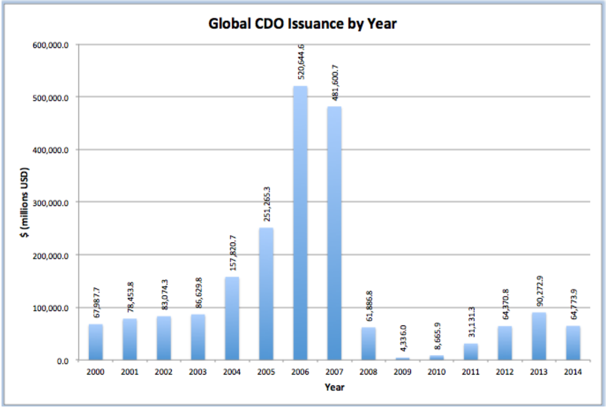
To price CDOs we will need to simulate Bernouilli variables which are correlated

CDOs and CLOs

- CLOs package together high-risk corporate debt and are then sold to institutional investors seeking potentially substantial returns
- CDOs (collateralized debt obligations), were comprised mostly of subprime mortgages and were blamed for the financial meltdown a decade ago
- The global CLO market is in 2019 between 1.4 and 2 trillion. In 2007, CDOs were 1.2 to 2.4 trillion
- Today's CLOs usually comprise corporate loans across a diversified set of industries

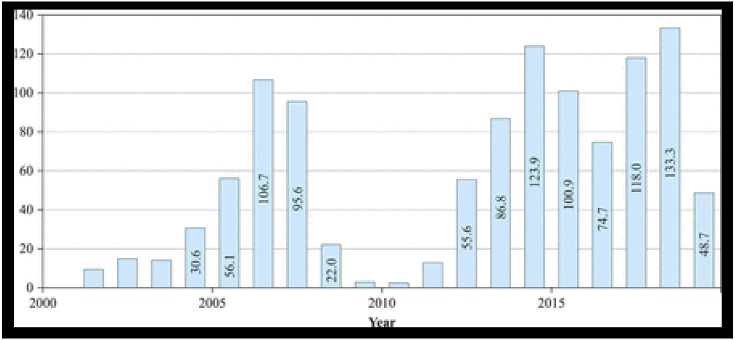
source *https://www.intralinks.com/blog/2019/11/clos-should-they-stay-or-should-they-go*

Collateralized Debt Obligations



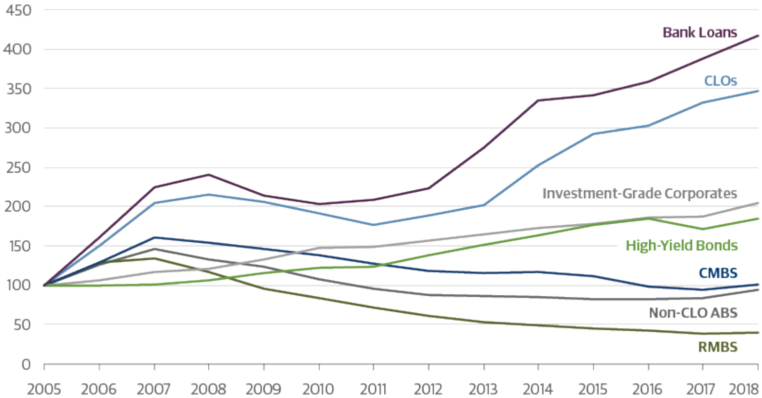
Source: Securities Industry and Financial Markets Association (SIFMA)

Collateralized Loan Obligations



US CLO Issuance, Source: The Journal of Structured Finance Summer 2019

Collateralized Loan Obligations

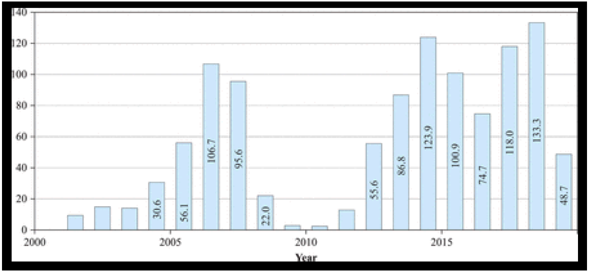


Fixed-Income instruments issuance; Index 2005 = 100, Source: Guggenheim May 2019
https://www.guggenheiminvestments.com/perspectives/portfolio-strategy/collateralized-loan-obligations-clo

From CDOs to CLOs

Source:https://seekingalpha.com/article/4307316-downgrades-in-clo-market-leading-to-higher-default-rates-in-2020

US CLO issuance



Source: https://www.spglobal.com/marketintelligence

Spreads US CLOs 2019-2020

US CLO average coupon across the stack and weighted average cost of capital (bps)

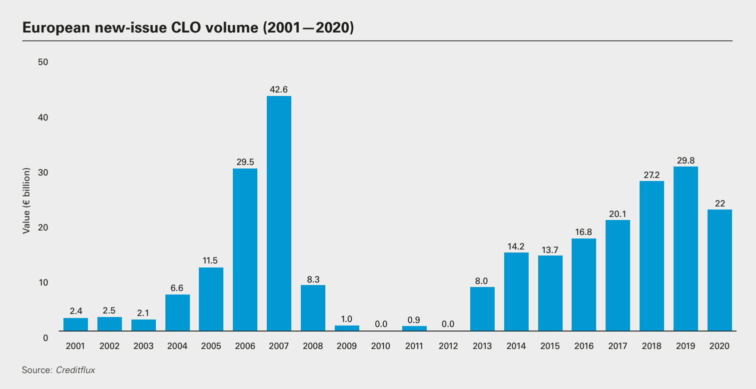
Time frame	AAA	AA	A	BBB	BB	WACC
2Q19 (L+)	135	193	268	378	672	201
3Q19 (L+)	135	190	278	398	704	206
4Q19 (L+)	134	196	277	406	734	203
1Q20 (L+)	124	174	234	349	675	182
2Q20 (L+)	193	266	338	463	690	241
3Q20 (L+)	160	213	279	417	757	217
4Q20 (L+)	138	184	254	396	738	195
Change from 3Q20	-22	-29	-24	-21	-19	-22
Change from a year ago	4	-12	-23	-10	4	-8

Data through Dec. 14, 2020.

Source: LCD, an offering of S&P Global Market Intelligence

Source: https://www.spglobal.com/marketintelligence

European CLO issuance 2001-2020



Source: White and Case
https://www.whitecase.com/publications/insight/european-levfin-2021/european-clos

Simulating Correlated Binomials

We construct here Bernouilli variables with the same parameter p which are correlated. The correlation is created through the default parameter in the following way.

Theorem: Simulation of Correlated Bernouilli Variables

Let $(Z_i)_{i\in[\![1,n]\!]}$ be independent variables of uniform law in $[\![0,1]\!]$
Let $\tilde{p}$ be a random variable in $[\![0,1]\!]$ with density f
Let $(X_i)_{i\in[\![1,n]\!]}$ be Bernouilli variables defined by $X_i = 1 \iff Z_i < \tilde{p}$
Then:
a) the $(X_i)_{i\in[\![1,n]\!]}$ are Bernouilli variables of parameters $\bar{p} = E[\tilde{p}]$
b) $\forall i \neq j$, $\rho(X_i, X_j) = \frac{Var(\tilde{p})}{\bar{p}(1-\bar{p})}$

Demonstration :

a) $E(X_i) = E\Big(E(X_i|\tilde{p})\Big) = E\Big(E(1_{Z_i<\tilde{p}}|\tilde{p})\Big) = E(\tilde{p})$

b) $E(X_iX_j) = E\Big(E(X_iX_j|\tilde{p})\Big) = E(\tilde{p}^2)$ as X_i and X_j are independent conditionnally on $\tilde{p}$

so, $Cov(X_iX_j) = E(\tilde{p}^2) - E(\tilde{p})^2 = Var(\tilde{p})$ and we know that for Bernouilli $Var(X_i) = Var(X_j) = E(\tilde{p})(1 - E(\tilde{p}))$ Q.E.D.

We consider now CDOs composed of bonds of the same notional with the same probabilities of default and same correlations and we are interested in calculating the law of the number of Bonds which default and therefore the law of $D_n = \sum_{i=1}^{i=n} X_i$

Remarks :

- the limitation of the model is that the resulting correlation between two bounds is always positive
- is $\tilde{p}$ is constant the correlation between the bonds is zero
- if $P(\tilde{p} = 0) = \frac{1}{2}$ and $P(\tilde{p} = 1) = \frac{1}{2}$ the correlation between the bonds is 100% as $var[\tilde{p}] = \frac{1}{4}$ and $\bar{p}(1 - \bar{p}) = \frac{1}{4}$

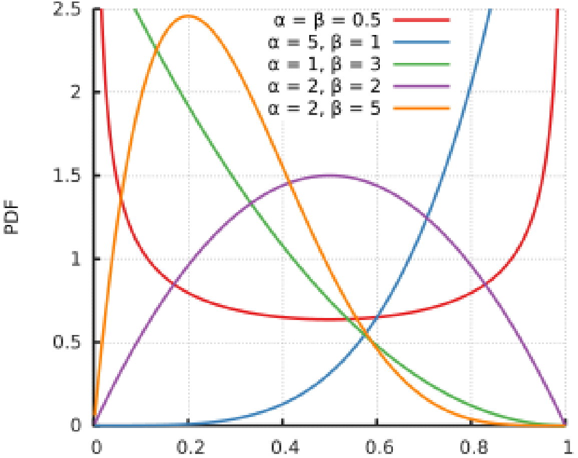
exercise 1 :
Show that $\forall X$ random variable in $[\![0, 1]\!]$, $Var[X] \leq \frac{1}{4}$

Hint : $Var[X] = E[(X - E(X))^2] = E[\left((X - \frac{1}{2}) + (\frac{1}{2} - E(X)\right)^2]$
$= E[(X - \frac{1}{2})^2] + E[(\frac{1}{2} - E(X))^2] + 2E[(X - \frac{1}{2})(\frac{1}{2} - E(X))]$
$= E[(X - \frac{1}{2})^2] + (\frac{1}{2} - E(X))^2 - 2(\frac{1}{2} - E(X))^2$
$= E[(X - \frac{1}{2})^2] - (\frac{1}{2} - E(X))^2 \leq E[(X - \frac{1}{2})^2] \leq \frac{1}{4}$ and the minimum is attained iff $\forall \omega$, $|X(\omega) - \frac{1}{2}| = \frac{1}{2}$

exercise 2 :
Often in simulations $\tilde{p} \sim B(\alpha, \beta)$ (beta law of parameters $\alpha > 0$ and $\beta > 0$) where the density is given by $f_{\alpha,\beta}(x) \propto x^{\alpha-1}(1-x)^{\beta-1}1_{x \in [\![0,1]\!]}$
Show that:
a) $E[\tilde{p}] = \frac{\alpha}{\alpha+\beta}$ noted $(\bar{p})$
b) $Var[\tilde{p}] = \frac{\bar{p}(1-\bar{p})}{\alpha+\beta+1}$
c) simulating with $\tilde{p}$ we have $\forall i \neq j$, $\rho(X_i, X_j) = \frac{1}{\alpha+\beta+1}$
d) show that $\forall p, \rho \in]\!]0,1[\![$, $\exists \alpha > 0, \beta > 0$, $\frac{\alpha}{\alpha+\beta} = \bar{p}$ and $\frac{1}{\alpha+\beta+1} = \rho$

Remarks: The Beta law is quite useful for the simulation of correlated Bernouilli variables as it is possible to choose α and β to obtain any possible probability of default and (positive) correlation wanted in the model.

Beta distribution



Beta distribution - source wikipedia

Simulating Correlated Binomials

Theorem: Law of $\frac{D_n}{n}$

a) $E\left(\frac{D_n}{n}\right) = \bar{p}$

b) $Var\left(\frac{D_n}{n}\right) = \frac{\bar{p}(1-\bar{p})}{n} + \frac{n-1}{n} Var[\tilde{p}]$

c) $\frac{D_n}{n} \longrightarrow \mathcal{L}(\tilde{p})$ (convergence in law)

so, in practice the probability that less than k bonds over n default is approximated by $P(\tilde{p} < \frac{k}{n})$

demonstration

a) $E\left(\frac{D_n}{n}\right) = E\left(\frac{1}{n}\sum_{i=1}^{i=n} X_i\right) = \frac{1}{n}\sum_{i=1}^{i=n} E\left(X_i\right) = E[\tilde{p}]$

b) $Var\left(\frac{D_n}{n}\right) = \frac{1}{n^2}\sum_{i=1}^{i=n} Var(X_i) + \frac{1}{n^2}\sum_{i \neq j} Cov(X_i, X_j)$

$= \frac{1}{n^2} \times n \times \bar{p}(1-\bar{p}) + \frac{1}{n^2} \times n(n-1) \times var[\tilde{p}]$

c) to show the convergence in law we show the convergence of the distribution functions

$$\lim_{n \longrightarrow +\infty} P(\frac{D_n}{n} < t) = \lim_{n \longrightarrow +\infty} E(1_{\frac{D_n}{n} < t})$$
$$= \lim_{n \longrightarrow +\infty} E(E(1_{\frac{D_n}{n} < t} | \tilde{p})) = E(E(\lim_{n \longrightarrow +\infty} 1_{\frac{D_n}{n} < t} | \tilde{p}))$$

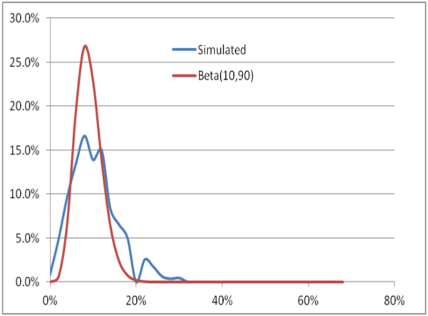
but when $\tilde{p}$ is known $\frac{D_n}{n} \longrightarrow \tilde{p}$ almost surely. so

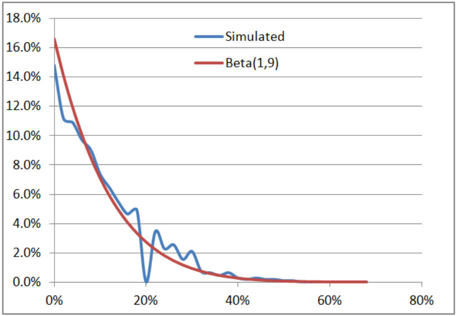
$E(\lim_{n \longrightarrow +\infty} 1_{\frac{D_n}{n} < t} | \tilde{p}) = 1_{\tilde{p} < t}$ so,

$$\lim_{n \longrightarrow +\infty} P(\frac{D_n}{n} < t) = E[1_{\tilde{p} < t}] = P(\tilde{p} < t) \text{ Q.E.D.}$$

Remarks: If the variables were not correlated in c) we would have convergence towards a single number, the mean, according to the Law of Large Numbers, instead of a convergence to a distribution

Histograms for $\frac{D_n}{n}$ for a CDO of 50 Bonds (1000 simulations)





Histograms are plotted by joining the values obtained for each 2% bucket
$\tilde{p} \sim Beta(10, 90) \Longrightarrow E[\tilde{p}] = 10\%$ and $Var[\tilde{p}] = 0.99\%$
$\tilde{p} \sim Beta(1, 9) \Longrightarrow E[\tilde{p}] = 10\%$ and $Var[\tilde{p}] = 9.09\%$

Beta Law for $\tilde{p}$ and CDO Pricing

Example: we consider a CDO made of 50 Bonds of equal Notional 100 each. We assume that the probabilities of default of the Bonds is 10% and note ρ the correlation of default between the bonds. We assume that the CDO has three tranches: Equity tranche (First 10% Loss), Junior Tranche (next 20% Loss), Senior Tranche (last 70% Loss). To calculate the price of the three tranches we use the approximation in Law $\mathcal{L}(\frac{D_n}{n}) \sim Beta(\alpha, \beta)$:

Table: Pricing as a function of ρ

	i.i.d Bernouilli	Beta(10,90)	Beta(1,9)
$E[\tilde{p}]$	10%	10%	10%
ρ	0	0.99%	9.09%
Senior	100%	100%	99.45%
Junior	91.68%	89.85%	82.93%
Equity	16.64%	20.32%	38.08%

Simulating Correlated Binomials

Definition: Diversity Score (Moody's)

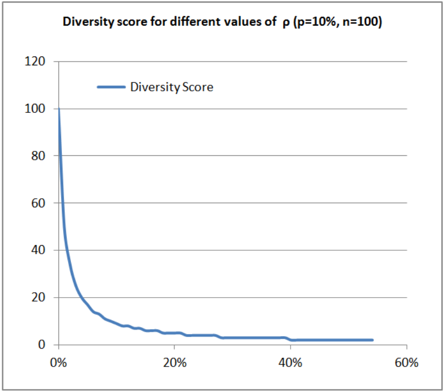
The Diversity Score is the number of uncorrelated bonds with the same probability of default $\bar{p}$ for which the variance of the proportion of losses would be the closest to $Var(\frac{D_n}{n})$

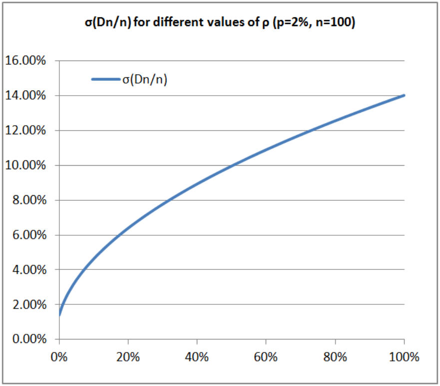
Remark: The diversity score summarizes the real diversification effect created by Bonds which are correlated.

Example: for n bonds with probability of default p and correlation ρ
$Var\left(\frac{D_n}{n}\right) = \frac{\bar{p}(1-\bar{p})}{n} + \frac{n-1}{n} Var[\tilde{p}]$ so we are searching for m such that
$\frac{\bar{p}(1-\bar{p})}{m} = \frac{\bar{p}(1-\bar{p})}{n} + \frac{n-1}{n} Var[\tilde{p}]$

N.A: for $n = 100$, $p = 2\%$ and $\rho = 20\%$, $\sigma_{20\%}(\frac{D_{100}}{100}) = 6.38\%$
with 5 independent assets $\sigma_{0\%}(\frac{D_5}{5}) = 6.26\%$ and with 4 independent assets $\sigma_{0\%}(\frac{D_4}{4}) = 7.00\%$. So we will take 5 as the Diversity Score.

Diversity Score





Diversity Score and Standard Deviation for different values of the correlation

Structural Models for $\tilde{p}$

We analyse here how Structural Models enable to create correlation in the modellisation (instead of creating it ex-nihilo in the model)
We make the following assumptions :

- bond i is in default at time T iif $A^i_T < D^i$ where:
  $dA^i_t = rA^i_t dt + \sigma^i A^i_t dW^i_t$
- σ^i is the same for all companies and is noted σ
- the distance to default is the same for all companies and we note
  $c = \frac{1}{\sigma\sqrt{T}}[ln(\frac{D^i}{A^i_0}) - rT + \frac{\sigma^2}{2}T]$
- we assume that the Brownian motions W^i_t verify
  $dW^i_t = \rho dW_t + \sqrt{1-\rho^2} dB^i_t$ where the B^i_t are brownian motions which are independent between them and independent from W_t

Remarks :
With the model $\forall i \neq j$, $\rho(W^i_t, W^j_t) = \rho$, and W_t is the common factor which creates correlation between the A^i_t and the default of the bonds.

Structural Models for $\tilde{p}$

Proposition

Let Z_i be the Bernouilli random variable with value 1 if the company i defaults and 0 otherwise. Then $Z_i = 1 \iff \frac{B^i_T}{\sqrt{T}} < \frac{c}{\sqrt{1-\rho^2}} - \frac{\rho}{\sqrt{1-\rho^2}} \frac{W_T}{\sqrt{T}}$

Demonstration simple

Remark 1 : Let Φ be the repartition fonction of a normal law $\mathcal{N}(0,1)$.
$\frac{B^i_T}{\sqrt{T}} \sim \mathcal{N}(0,1) \Longrightarrow \Phi(\frac{B^i_T}{\sqrt{T}}) \sim \mathcal{U}(0,1)$
so $X_i = 1 \iff \Phi\left(\frac{W^i_T}{\sqrt{T}}\right) < \Phi\left(\frac{c}{\sqrt{1-\rho^2}} - \frac{\rho}{\sqrt{1-\rho^2}} \frac{W_T}{\sqrt{T}}\right)$
so we end up simulating (as previously) correlated Bernouilli variables with the function $\tilde{p}$ having a law $\tilde{p} \sim \Phi\left(\frac{c}{\sqrt{1-\rho^2}} - \frac{\rho}{\sqrt{1-\rho^2}} \frac{W_T}{\sqrt{T}}\right)$

Remark 2 : We have different alternatives for $\tilde{p}$ to generate correlated binomials:

- to use a beta distribution $B(\alpha, \beta)$ (as seen previoulsy)
- to use the distribution of $\Phi(\alpha + \beta Z)$ (where $Z \sim \mathcal{N}(0,1)$)

In both cases:

- first we solve for α and β to match the desired value for $\bar{p}$ and ρ
- then to price the CDO we approximate the law of $\frac{D_n}{n}$ by the law of $\tilde{p}$

Proposition

If $\tilde{p} \sim \Phi(\alpha + \beta Z)$ (where $Z \sim \mathcal{N}(0,1)$) then

a) $E[\tilde{p}] = \Phi\left(\frac{\alpha}{\sqrt{1+\beta^2}}\right)$ (that we note also $\bar{p}$)

b) $E[\tilde{p}^2] = \Phi_{2,\frac{\beta^2}{1+\beta^2}}\left(\frac{\alpha}{\sqrt{1+\beta^2}}, \frac{\alpha}{\sqrt{1+\beta^2}}\right)$

c) $P(\tilde{p} < t) = \Phi\left(\frac{1}{\beta}[\Phi^{-1}(t) - \sqrt{1+\beta^2}\Phi^{-1}(\bar{p})]\right)$

Structural Models for $\tilde{p}$

a) $E[\tilde{p}] = E[\Phi(\alpha + \beta Z)] = E[E(1_{Z_0<\alpha+\beta Z}|Z)]$
(with $Z_0 \sim \mathcal{N}(0,1)$ independent from Z)
$= E[1_{Z_0-\beta Z<\alpha}] = E[1_{\frac{Z_0-\beta Z}{\sqrt{1+\beta^2}}<\frac{\alpha}{\sqrt{1+\beta^2}}}] = \Phi(\frac{\alpha}{\sqrt{1+\beta^2}})$

b) $E(\tilde{p}^2) = E[\Phi(\alpha + \beta Z)^2]$
and $\Phi(\alpha + \beta Z)^2 = E[1_{Z_0<\alpha+\beta Z}1_{Z_1<\alpha+\beta Z}|Z]$
(with Z_0, Z_1,Z independent $\mathcal{N}(0,1)$)
so, $E[\Phi(\alpha + \beta Z)^2] = E[1_{Z_0<\alpha+\beta Z}1_{Z_1<\alpha+\beta Z}]$
$= E[1_{\frac{Z_0-\beta Z}{\sqrt{1+\beta^2}}<\frac{\alpha}{\sqrt{1+\beta^2}}}1_{\frac{Z_1-\beta Z}{\sqrt{1+\beta^2}}<\frac{\alpha}{\sqrt{1+\beta^2}}}]$
$= \Phi_{2,\frac{\beta^2}{1+\beta^2}}(\frac{\alpha}{\sqrt{1+\beta^2}}, \frac{\alpha}{\sqrt{1+\beta^2}})$

with $\Phi_{2,\gamma}$ repartition function of a bivariate normal variable $\mathcal{N}\begin{pmatrix} 1 & \gamma \\ \gamma & 1 \end{pmatrix}$

c) $P(\tilde{p} < t) = P(\Phi(\alpha + \beta Z) < t) = \Phi(\frac{\Phi^{-1}(t) - \alpha}{\beta})$
as $\Phi(\frac{\alpha}{\sqrt{1+\beta^2}}) = \bar{p}$ we have $\alpha = \Phi^{-1}(\bar{p})\sqrt{1 + \beta^2}$ so
$P(\tilde{p} < t) = \Phi\left(\frac{1}{\beta}[\Phi^{-1}(t) - \sqrt{1 + \beta^2}\Phi^{-1}(\bar{p})]\right)$. Q.E.D.

Example we consider a CDO with 100 Bonds of the same Notional and recovery rate of zero. The default of the bonds are modelized by Bernouilli variables X_i of parameter p and correlations ρ. We consider a junior tranche for the CDO which is exposed to the losses between above 10% and up to 30%. Price this junior tranche assuming $\bar{p} = 2\%$ and $\rho = 10\%$ with the two previous models:
a) assuming $\tilde{p} \sim B(\alpha, \beta)$
b) assuming $\tilde{p} \sim \Phi(\alpha + \beta Z)$ where $Z \sim \mathcal{N}(0, 1)$

Results:

a) we solve and find $\alpha = (\frac{1}{\rho} - 1)\bar{p}$ and $\beta = (\frac{1}{\rho} - 1)(1 - \bar{p})$. So, here $\tilde{p} \sim B(0.18, 8.82)$. Taking a risk free rate of zero we price the junior tranche in % of face value as $\frac{1}{20}\sum_{i=11}^{i=30} P(\frac{D_{100}}{100} < \frac{i}{100})$ that we approximate by $\frac{1}{20}\sum_{i=11}^{i=30} P(\tilde{p} < \frac{i}{100}) = 98.12\%$

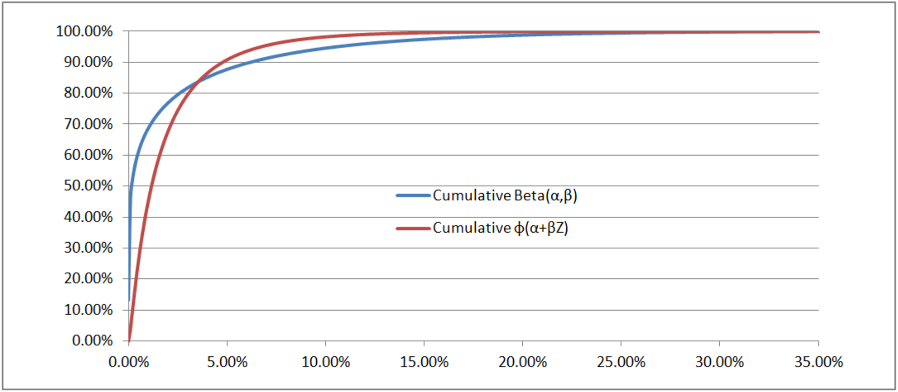
b) we solve $\frac{\alpha}{\sqrt{1+\beta^2}} = -2.05375$ and (using a program to calculate the bivariate normal) $\frac{\beta^2}{1+\beta^2} = 0.18$. This implies $\alpha = -2.2678$ and $\beta^2 = 0.2195$. The sign of β is not determined as both Z and $-Z$ are $\mathcal{N}(0,1)$, we will take $\beta = 0.468521$. Now,

$$\frac{1}{20}\sum_{i=11}^{i=30} P(\Phi(\alpha + \beta Z) < \frac{i}{100}) = \frac{1}{20}\sum_{i=11}^{i=30} \Phi(\frac{1}{\beta}[\Phi^{-1}(\frac{i}{100}) - \alpha]) = 99.66\%$$

Remarks:

- the pricings for a) and b) are not exactly the same as the two laws used for $\tilde{p}$ produce the same expectations and correlations between the default events (the Bernouilli variables) but not exactly the same joint distributions. Also they do not generate the same laws for $\frac{D_n}{n}$ and therefore not the same pricing.
- the fact that the two laws generated for $\frac{D_n}{n}$ are different is also put in evidence by the fact that (as it has been demonstrated previoulsy) $\frac{D_n}{n}$ converge here towards two different distributions which are the two distinct laws of $\tilde{p}$ that we use.
- the choice of the distribution $\tilde{p}$ used to create the correlation structure is therefore important and it is exactly the aim of the study of copulas to create adequate correlation structures. Transformations of normal variables $\Phi(\alpha + \beta Z)$ to create correlation structures have been criticised after 2008.

Calibration of the two Models for $\tilde{p}$



Modelizing $\tilde{p}$ with Beta(α, β) or $\Phi(\alpha + \beta Z)$

Proposition

We assume here that:

- the percentage lost for a bond which defaults (i.e $1 - R$) is $f(\tilde{p})$
- $f(p)$ is an increasing function of p

If we note $L_n^f = \frac{1}{n} \sum_{i=1}^{i=n} f(\tilde{p}) 1_{Z_i < \tilde{p}}$ the loss in percentage for the CDO we have : $L_n^f \longrightarrow \mathcal{L}(\tilde{p} f(\tilde{p}))$ (convergence in law).

Demonstration :

to show the convergence in law we show the convergence of the distribution function

Generalization to None Zero Recovery Rate

$\lim_{n \longrightarrow +\infty} P(L_n^f < t) = \lim_{n \longrightarrow +\infty} E(1_{L_n^f < t}) = \lim_{n \longrightarrow +\infty} E(E(1_{L_n^f < t}|\tilde{p}))$
$= E(E(\lim_{n \longrightarrow +\infty} 1_{L_n^f < t}|\tilde{p}))$
when $\tilde{p}$ is known then according to the law of large numbers
$L_n^f \longrightarrow E[f(\tilde{p})1_{Z_i < \tilde{p}}] = f(\tilde{p})\tilde{p}$ and so $1_{L_n^f < t} \longrightarrow 1_{f(\tilde{p})\tilde{p} < t}$
from there $\lim_{n \longrightarrow +\infty} P(L_n^f < t) = E(E(1_{\tilde{p}f(\tilde{p}) < t}|\tilde{p})) = P(\tilde{p}f(\tilde{p}) < t)$ Q.E.D.

Remarks: if $R = 0$ then $f(\tilde{p})$ is always 1 and we find the result we already demonstrated that $\mathcal{L}(\tilde{p}f(\tilde{p})) \sim \mathcal{L}(\tilde{p})$

Remark: In the Cox model a default occurs for company i iff

$$\mathcal{E}_i < \int_0^t \lambda(X_s)ds$$

where the $\mathcal{E}_i$ are independent and independent from X.
Let $F(t) = P(\mathcal{E}_i \leq t) = 1 - exp(-t)$. Then the $Z_i = F^{-1}(\mathcal{E}_i)$ are independent $\mathcal{U}([0,1])$ and a default occurs for company i iff

$$Z_i < F^{-1}(\int_0^t \lambda(X_s)ds)$$

$$\Longleftrightarrow Z_i < \tilde{p}$$

where $\tilde{p} = 1 - exp(-\int_0^t \lambda(X_s)ds)$.

Definition: Infection Models

Let $(Z_i)_{i\in[\![1,n]\!]}$ and $(Y_{i,j})_{i\neq j\in[\![1,n]\!]}$ be independent variables, we assume $Z_i \sim \mathcal{B}(p)$ and $\forall i \neq j,\ Y_{i,j} \sim \mathcal{B}(q)$
Then we define in a contagion model the variables $(X_i)_{i\in[\![1,n]\!]}$ by :
$X_i = Z_i + (1 - Z_i)[1 - \prod_{j\neq i}(1 - Z_j Y_{j,i})]$

Remark:
The only possible values for X_i are 1 and 0.
$X_i = 1 \iff Z_i = 1$ or $\exists i \neq j,\ Z_j = 1$ and $Y_{j,i} = 1$ (i.e contamination)

We are now going to study the law of the X_i and their correlations.

Proposition

$X_i \sim \mathcal{B}(1-(1-p)(1-pq)^{n-1})$

Demonstration : because of independence

$E(X_1) = E(Z_1) + (1-E(Z_1))[1 - \prod_{j\neq 1}(1-E(Z_j)E(Y_{j,1}))]$

$= p + (1-p)[1-(1-pq)^{n-1}]$

$= p + 1 - p - (1-p)(1-pq)^{n-1}$

$= 1-(1-p)(1-pq)^{n-1}$ Q.E.D.

Remarks :

$\mathcal{L}(X_1) \xrightarrow[n \longrightarrow +\infty]{} 1$

Proposition

$E[X_1X_2] = 1 - 2(1-p)(1-pq)^{n-1} + (1-p)^2(1-2pq+pq^2)^{n-2}$

Demonstration :

$E[X_1X_2]$

$= E\left[\left(Z_1+(1-Z_1)[1-\prod_{j\neq 1}(1-Z_jY_{j,1})]\right)\left(Z_2+(1-Z_2)[1-\prod_{j\neq 2}(1-Z_jY_{j,2})]\right)\right]$

we have 3 different type of terms:

a) $E[Z_1Z_2] = p^2$ (because Z_1 and Z_2 are independent)

b) $E\left(Z_1(1-Z_2)[1-\prod_{j\neq 2}(1-Z_jY_{j,2})]\right)$ (this value will appear two times)

$= E\left(Z_1(1-Z_2)\right) - E\left(Z_1(1-Z_1Y_{1,2})(1-Z_2)\prod_{j\notin\{1,2\}}(1-Z_jY_{j,2})\right)$

$= p(1-p) - (p-pq)(1-p)(1-pq)^{n-2}$

$= p(1-p)[1-(1-q)(1-pq)^{n-2}]$

c) $E\left[(1-Z_1)(1-Z_2)[1-\prod_{j\neq 1}(1-Z_jY_{j,1})][1-\prod_{j\neq 2}(1-Z_jY_{j,2})]\right]$

multiplying first the two terms on the right we get 3 different type of terms:

∘ $E[(1-Z_1)(1-Z_2)] = (1-p)^2$

$\circ -E\left[(1-Z_1)(1-Z_2)\prod_{j\neq 2}(1-Z_jY_{j,2})\right]$ (this value will appear two times)

$$= -E\left[(1-Z_1)(1-Z_1Y_{1,2})(1-Z_2)\prod_{j\notin 1,2}(1-Z_jY_{j,2})\right]$$
$$= -(1-pq-p+pq)(1-p)(1-pq)^{n-2}$$
$$= -(1-p)^2(1-pq)^{n-2}$$

$\circ\ E\left[(1-Z_1)(1-Z_2)\prod_{j\neq 1}(1-Z_jY_{j,1})\prod_{j\neq 2}(1-Z_jY_{j,2})\right]$

$$= E\left[(1-Z_1)(1-Z_1Y_{1,2})(1-Z_2)(1-Z_2Y_{2,1})\prod_{j\notin 1,2}[(1-Z_jY_{j,1})(1-Z_jY_{j,2})]\right]$$
$$= (1-pq-p+pq)^2(1-pq-pq+pq^2)^{n-2}$$
$$= (1-p)^2(1-2pq+pq^2)^{n-2}$$

so at the end we obtain:

$$p^2+2p(1-p)[1-(1-q)(1-pq)^{n-2}]+(1-p)^2$$
$$-2(1-p)^2(1-pq)^{n-2}+(1-p)^2(1-2pq+pq^2)^{n-2}$$
$$= 1-2(1-p)(1-pq)^{n-1}+(1-p)^2(1-2pq+pq^2)^{n-2}$$ Q.E.D.

exercise :

Let $D_n = \sum_{i=1}^{i=n} X_i$, calculate as a function of p and q :

a) $E[D_n]$ and

b) $Var[D_n]$

Hint :

a) $E[D_n] = \sum_{i=1}^{i=n} E[X_i] = nE[X_1]$

b) $Var[D_n] = E\left[(\sum_{i=1}^{i=n} X_i)^2\right] - (E[D_n])^2 = \sum_{i=1}^{i=n}\sum_{j=1}^{j=n} E\left[X_i X_j\right] - (E[D_n])^2$

$= n(n-1)E[X_1X_2] + nE[X_1] - n^2E[X_1]^2$

So we know how to calculate the first two moments of D_n as a function of p and q but in fact we can also calculate the law of D_n

Proposition :

$\forall k \in [\![1, n]\!]$,

$$P(D_n = k) = C_n^k \sum_{i=1}^{i=k} C_k^i p^i (1-p)^{(n-i)} (1-q)^{i(n-k)} [1-(1-q)^i]^{(k-i)}$$

Demonstration :

$P(D_n = k) = C_n^k P(X_1 = 1, X_2 = 1, \cdots, X_k = 1, X_{k+1} = 0, \cdots, X_n = 0)$
and $\{X_1 = 1, X_2 = 1, \cdots, X_k = 1\}$ can de decomposed in $k-1$ cases depending on the number i of "direct" defaults. So

$P(X_1 = 1, X_2 = 1, \cdots, X_k = 1, X_{k+1} = 0, \cdots, X_n = 0)$

$= \sum_{i=1}^{i=k} C_k^i P(Z_1 = 1, Z_2 = 1, \cdots, Z_i = 1, (Z_{i+1} = 0, X_{i+1} = 1), \cdots, (Z_k = 0, X_k = 1), X_{k+1} = 0, \cdots, X_n = 0)$

we can write each event as the intersection of three events

$\circ$ $\{Z_1 = 1, Z_2 = 1, \cdots, Z_i = 1, Z_{i+1} = 0, \cdots, Z_n = 0\}$
$\circ$ $\{\exists j \in [\![1, i]\!], Y_{j,i+1} = 1, \cdots, \exists j \in [\![1, i]\!], Y_{j,k} = 1\}$
$\circ$ $\{\forall j \in [\![1, i]\!], Y_{j,k+1} = 0, \cdots, \forall j \in [\![1, i]\!], Y_{j,n} = 0\}$
the three events are independent.
$\circ$ the probability of the first one is $p^i(1-p)^{(n-i)}$
$\circ$ the probability of the second one is $[1-(1-q)^i]^{(k-i)}$
$\circ$ the probability of the third one is $(1-q)^{i(n-k)}$ Q.E.D.

Example :

Table: Infection Models for $n = 30$

$p = P(Z_i = 1)$	1%	1%	1%	1%	1%
q	0%	10%	20%	50%	100%
$p^* = P(X_i = 1)$	1%	3.83%	6.58%	14.39%	26.03%
Correlation	0%	12%	21%	50%	100%
Diversity Score	30	6.7	4.1	2	1

Remarks : if $q = 100\%$
$P(X_i = 1) = 1 - P(X_i = 0) = 1 - (P(Z_1 = 0))^{30} = 1 - (1 - p)^{30} = 26.03\%$

Copulas

Definition : Copulas

$C : [0,1]^d \longrightarrow [0,1]$ is a copula iff C is a multivariate cumulative distribution for a random vector of $[0,1]^d$ i.e
$\exists (U_1, U_2, \cdots, U_d)$ r.v $(\Omega, P) \longrightarrow [0,1]^d$ such that:

- $\forall i \in [\![1, d]\!],\ U_i \sim \mathcal{U}([0,1])$
- $C(u_1, u_2, \cdots, u_d) = P(U_1 \leq u_1, U_2 \leq u_2, \cdots, U_d \leq u_d)$

Notation : we note
$F_U(u_1, u_2, \cdots, u_d) = P(U_1 \leq u_1, U_2 \leq u_2, \cdots, U_d \leq u_d)$ the multidimensional cumulative distribution function of U.
By definition for any copula C there is $U = (U_1, U_2, \cdots, U_d)$ with $U_i \sim \mathcal{U}([0,1])$ such that $C = F_U$

Exemples : Let $U = (U_1, U_2)$ where U_1 and $U_2 \sim \mathcal{U}([0,1])$
a) if U_1 and U_2 are independent then $F_U(u,v) = uv$
b) if $U_2 = U_1$ then $F_U(u,v) = min(u,v)$
c) if $U_2 = 1 - U_1$ then $F_U(u,v) = max(u+v-1,0)$

Demonstration : Let's show c)
$P(U_1 \leq u, 1 - U_1 \leq v) = P(U_1 \leq u, U_1 \geq 1 - v) = P(1 - v \leq U_1 \leq u)$
$= max(u+v-1,0)$ Q.E.D.

Theorem : Frechet-Hoeffding Bounds

Let $U = (U_1, U_2)$ be a r.v with $U_i \sim \mathcal{U}([0,1])$ then
$\forall u, v \in [0,1]$, $max(u+v-1,0) \leq F_U(u,v) \leq min(u,v)$
so the cases $U_2 = U_1$ and $U_2 = 1 - U_1$ represents the two extreme "correlation-structures".

Demonstration :
$P(U_1 \leq u, U_2 \leq v) \leq P(U_1 \leq u)$ and $P(U_1 \leq u, U_2 \leq v) \leq P(U_2 \leq v)$
implies $P(U_1 \leq u, U_2 \leq v) \leq min(P(U_1 \leq u), P(U_2 \leq v))$
$P(\{U_1 \leq u\} \cup \{U_2 \leq v\}) = P(U_1 \leq u) + P(U_2 \leq v) - P(U_1 \leq u, U_2 \leq v)$
implies $P(U_1 \leq u, U_2 \leq v) \geq P(U_1 \leq u) + P(U_2 \leq v) - 1$ Q.E.D.

Definition : Quantile (or Pseudo-Inverse)

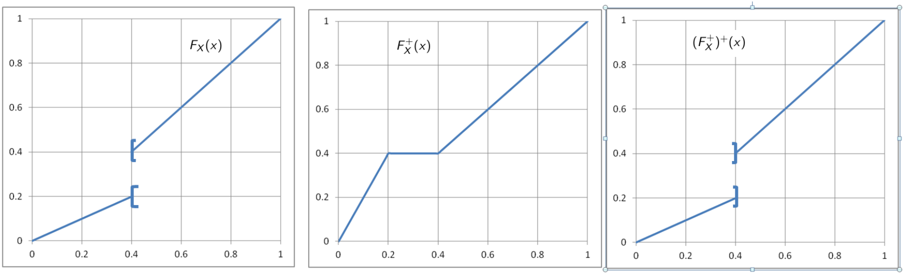
We define $F_X^+ : [0,1] \longrightarrow \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ by
$F_X^+(y) = \inf_{x \in \mathbb{R}} \{P(X \leq x) \geq y\}$

Definition

f is strictly increasing at x iff $\forall x_1 < x < x_2$, $f(x_1) < f(x) < f(x_2)$

General Properties

- F_X is increasing and right-continuous
- F_X^+ is increasing and left continuous
- F_X is continuous at $x \iff P(X = x) = 0$
- $F_X(x) \geq y \iff x \geq F_X^+(y)$
- $F_X^+(F_X(x)) \leq x$
- $F_X(F_X^+(y)) \geq y$
- F_X^+ continuous at $F_X(x) \iff F_X$ strictly increasing at x
- F_X^+ continuous at $F_X(x) \iff F_X^+(F_X(x)) = x$
- F_X continuous at $F_X^+(y) \iff F_X^+$ strictly increasing at y
- F_X continuous at $F_X^+(y) \iff F_X(F_X^+(y)) = y$
- F_X and F_X^+ continuous $\iff F_X$ invertible and $F_X^{-1} = F_X^+$



Calculation of the Pseudo Inverse

Demonstration : Let as an exercise

Proposition and Definition. Copula C_X of a Random Vector

If X r.v taking values $\in \mathbb{R}$
a) F_X continuous $\Longrightarrow F_X(X) \sim \mathcal{U}([0,1])$

if $X = (X_1, \cdots, X_d)$ with cdfs $F_{X_1}, \cdots, F_{X_1}$ continuous and
$U = (F_{X_1}(X_1), \cdots, F_{X_d}(X_d))$ then
b) $\forall i \in [\![1, d]\!]$, $F_{X_i}(X_i) \sim \mathcal{U}([0,1])$

We call copula of X and note C_X the function F_U
(that we can also note C_U)

Demonstration:

a) let $y \in]0,1[$

$P(F_X(X) < y) = 1 - P(F_X(X) \geq y)$

$= 1 - P(X \geq F_X^{+}(y))$ (according to the general properties)

$= P(X < F_X^{+}(y)) = P(X \leq F_X^{+}(y))$ (because F_X is continuous)

$= F_X(F_X^{+}(y)) = y$ (according to the proposition as F_X is continuous)

so $F_X(X) \sim \mathcal{U}([0,1])$ Q.E.D.

b) direct consequence of a)

Remarks :

if $U = (U_1, U_2, \cdots, U_d)$ with $U_i \sim \mathcal{U}([0,1])$ then $C_U = F_U$.

Sklar's theorem enable to build a random vector with given continuous cdf marginals and copula.

Sklar's Theorem: Multivariate with given Marginals and Copula

Let $U = (U_1, U_2, \cdots, U_d)$ with $U_i \sim \mathcal{U}([0,1])$
Let $F_1, F_2, \cdots, F_d$ be continuous cumulative distribution functions.
Let $X = (F_1^{+}(U_1), F_2^{+}(U_2), \cdots, F_d^{+}(U_d))$
Then,

- $F_{X_i} = F_i$ and
- $C_X = C_U$

Demonstration :
a) $P(X_i \leq x) = P(F_i^+(U) \leq x)$
$= P(U \leq F_i(x))$ (according to the general properties)
$= F_i(x)$ Q.E.D.

b) $C_X(u_1, u_2, \cdots, u_d)$
$= P(F_1(X_1) \leq u_1, \cdots, F_d(X_d) \leq u_d)$ (by definition)
$= P((F_1 \circ F_1^+)(U_1) \leq u_1, , \cdots, (F_d \circ F_d^+)(U_d) \leq u_d)$
but F_i continuous $\Longrightarrow F_i \circ F_i^+ = Id$ (according to the general properties)
so,
$= P(U_1 \leq u_1, \cdots, U_d \leq u_d)$
$= C_U(u_1, \cdots, u_d)$ Q.E.D.

Remark 1 : According to Sklar's theorem:

- for any copula C and
- for any continuous cdfs $(F_i)_{i\in[\![1,d]\!]}$

we can find a multivariate random variable X such that:

- the F_i are the marginal cdfs of X
- $C_X = C$

we will have $F_X(x_1, x_2, \cdots, x_d) = C(F_1(x_1), F_2(x_2), \cdots, F_d(x_d))$

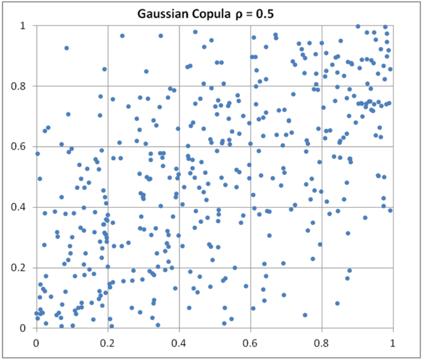
Remark 2 : It is easy to simulate a Gaussian vector Z with a given correlation matrix and therefore easy to simulate variables U with marginals $\mathcal{U}([0,1])$ and with $C_U = C_Z$ by calculating for each value of Z the vector $U = (F_{Z_1}(Z_1), \cdots, F_{Z_d}(Z_d))$

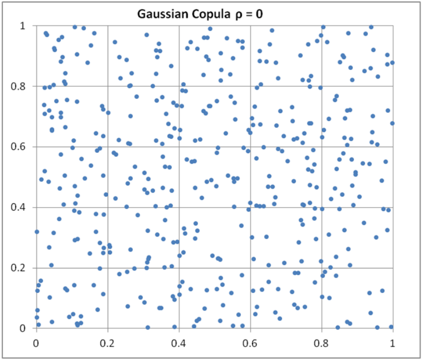
Exercise: Let Z be a Gaussian vector with a given correlation matrix and C_Z be its copula. Let $F_1, \cdots, F_d$ be continuous cdfs.
Show that we can simulate a r.v X of copula C_Z and marginals F_i by

- simulating Z and
- calculating for each value of Z the vector $X = ((F_1^+ \circ F_{Z_1})(Z_1), \cdots, (F_d^+ \circ F_{Z_d})(Z_d))$

Solution:
$F_{Z_i}(Z_i) \sim \mathcal{U}([0,1]) \Longrightarrow F_i^+(F_{Z_i}(Z_i))$ has for cdf F_i (according to Sklar's theorem). So, we get the right marginals for X
$C_X(u_1, \cdots, u_d) = P(F_1(X_1) \leq u_1, \cdots, F_d(X_d) \leq u_d)$
$= P(F_1 \circ (F_1^+ \circ F_{Z_1})(Z_1) \leq u_1, \cdots, F_d \circ (F_d^+ \circ F_{Z_d})(Z_d) \leq u_d)$
$= P(F_{Z_1}(Z_1) \leq u_1, \cdots, F_{Z_d}(Z_d) \leq u_d) = C_Z(u_1, \cdots, u_d)$
so we get the right copula for X. Q.E.D.

Copula





Simulations Gaussian Copula for various values of ρ

Remarks:

In the SRM model $X_i = 1 \iff \frac{B_T^i}{\sqrt{T}} < \frac{c}{\sqrt{1-\rho^2}} - \frac{\rho}{\sqrt{1-\rho^2}} \frac{W_T}{\sqrt{T}}$

but $Z^i < \alpha + \beta Z \iff \Phi\left(\frac{Z^i - \beta Z}{\sqrt{1+\beta^2}}\right) < \Phi\left(\frac{\alpha}{\sqrt{1+\beta^2}}\right)$

Let

$U = (U_1, \cdots, U_d)$ with $U_i = \Phi\left(\frac{Z^i - \beta Z}{\sqrt{1+\beta^2}}\right)$ and

$G = (\frac{Z^1 - \beta Z}{\sqrt{1+\beta^2}}, \cdots, \frac{Z^i d - \beta Z}{\sqrt{1+\beta^2}})$ then

$U_i \sim \mathcal{N}(0,1)$ and $C_U = C_G$ and therefore C_U is a Gaussian Copula

The fact that the defaults $(U_i < \Phi\left(\frac{\alpha}{\sqrt{1+\beta^2}}\right))$ are correlated implies, when simulating U, the formation of a cluster of points (density of points higher than the average) in the region near 0 as the probability for this event is higher than what would be expected for a product of the marginals.

Proposition : Invariance Properties of the Copulas

Let $X = (X_1, X_2, \cdots, X_d)$ with continuous marginal cdfs F_i
Let $T_1, T_2, \cdots, T_d$ be strictly increasing real functions
Let $Y = (Y_1, Y_2, \cdots, Y_d)$ with $Y_i = T_i(X_i)$
then $C_Y = C_X$
So, the Copula, which measures the association between the variables, is invariant by change of variables under strictly increasing functions (which is not the case for the linear correlation).

Demonstration : $C_Y(u_1, \cdots u_d) = P(F_{Y_1}(Y_1) \leq u_1, \cdots, F_{Y_d}(Y_d) \leq u_d)$
but $F_{Y_i}(y) = P(Y_i \leq y) = P(T_i(X_i) \leq y)$ so
$F_{Y_i}(T_i(x)) = P(T_i(X_i) \leq T_i(x)) = P(X_i \leq x)$ (as T_i is strictly increasing)
so, $F_{Y_i}(T_i(x)) = F_{X_i}(x)$ and in particular $F_{Y_i}(Y_i) = F_{Y_i}(T_i(X_i)) = F_{X_i}(X_i)$
so,

$$C_Y(u_1, \cdots u_d) = P(F_{X_1}(X_1) \leq u_1, \cdots, F_{X_d}(X_d) \leq u_d) = C_X(u_1, \cdots u_d).$$

Proposition : Copulas for Normalized Gaussian Vectors

Let X be a "normalized" Gaussian vector $\mathcal{N}(0,\Sigma)$ with components $X_i \sim \mathcal{N}(0,1)$ and correlation matrix Σ invertible.
Let C_X the copula of X and c_X its density. Then:

- $C_X(x) = F_X(\Phi^{-1}(x_1), \Phi^{-1}(x_2), \cdots, \Phi^{-1}(x_d))$
- $c_X(x) = \frac{1}{det(\Sigma)^{\frac{1}{2}}} exp(-\frac{1}{2}x'(\Sigma^{-1} - I_d)x)$

where Φ is the cdf of a $\mathcal{N}(0,1)$

Demonstration :

$F_X(x_1, x_2, \cdots, x_d) = P(X_1 \leq x_1, \cdots, X_d \leq x_d)$
$= P(\Phi(X_1) \leq \Phi(x_1), \cdots, \Phi(X_d) \leq \Phi(x_d))$ (as Φ is strictly increasing)
$= C_X(\Phi(x_1), \cdots, \Phi(x_d))$ so,
$C_X(x_1, x_2, \cdots, x_d) = F_X(\Phi^{-1}(x_1), \Phi^{-1}(x_2), \cdots, \Phi^{-1}(x_d)) =$ Q.E.D.

Applying $\frac{\partial}{\partial x_1 \partial x_2 \cdots \partial x_d}$ to $F_X(x_1, x_2, \cdots, x_d) = C_X(\Phi(x_1), \Phi(x_2), \cdots, \Phi(x_d))$ we get

$f_X(x_1, x_2, \cdots, x_d) = c_X(\Phi(x_1), \Phi(x_2), \cdots, \Phi(x_d))\phi(x_1)\phi(x_2)\cdots\phi(x_d)$
the density $f_X(x)$ equals $(\frac{1}{\sqrt{2\pi}})^d exp(-\frac{1}{2}x'\Sigma^{-1}x)$ and
$\phi(x_1)\phi(x_2)\cdots\phi(x_d) = (\frac{1}{\sqrt{2\pi}})^d exp(-\frac{1}{2}x'x)$ Q.E.D.

Proposition

The Copula of a Gaussian vector depends only on its correlation matrix Λ

Demonstration : if X is a Gaussian vector of correlation matrix Σ we know (from the invariance property) that the normalized Gaussian vector Y where $Y_i = T_i(X_i) = \frac{X_i - \mu_i}{\sigma_i}$ has the same copula as X and that $Y \sim \mathcal{N}(0, \Lambda)$ where Λ is the correlation matrix of X. Q.E.D.

Remark: We can create a correlation structure on d binomial variables $X_i \sim \mathcal{B}(p_i)$ by choosing a copula C and $(\alpha_1, \alpha_2, \cdots, \alpha_d)$ such that:
$\forall i \in [\![1, d]\!]$, $C(1, \cdots, \alpha_i, 1 \cdots) = p_i$
Example: Here $d = 3$ and $p_1 = 1\%$, $p_2 = 2\%$ and $p_3 = 3\%$.
$Z = (Z_1, Z_2, Z_3)$ is a Gaussian vector with correlation $\rho = 50\%$ between two variables and we assume that $X_i = 1$ (i defaults) $\Longleftrightarrow Z_i \leq \alpha_i$
Solving $P(Z_i \leq \alpha_i) = p_i$ we find: $\alpha_1 = -2.326$ $\alpha_2 = -2.054$ $\alpha_3 = -1.881$
Here we do not try to calibrate a correlation matrix for Z to match some input correlations between the X_i but calculate the correlations between the defaults induced by the correlation matrix of Z.
Here we get $E(X_1X_2) - E(X_1)E(X_2) = P(Z_1 < \alpha_1, Z_2 < \alpha_2) - p_1p_2$ and
$\rho(X_1, X_2) = \frac{cov(X_1,X_2)}{\sigma(X_1)\sigma(X_2)} = 13.32\%$
and in the same way $\rho(X_1, X_3) = 13.89\%$ and $\rho(X_2, X_3) = 16.16\%$.
To simulate Z we simulate X and then calculate $(1_{X_1<\alpha_1}, 1_{X_2<\alpha_2}, 1_{X_3<\alpha_3})$

Rosenblatt's Theorem

Let $X = (X_1, X_2, \cdots, X_d)$ be a random vector of $\mathbb{R}^d$
we assume that the law of X has a density $f_X(x_1, x_2, \cdots, x_d)$ strictly positive
Let $T : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ be defined by $T(x) = y$ with
$y_1 = P(X_1 \leq x_1)$
$y_2 = P(X_2 \leq x_2 | X_1 = x_1), \cdots$
$y_d = P(X_d \leq x_d | X_1 = x_1, X_2 = x_2, \cdots, X_{d-1} = x_{d-1})$
Then, $T(X) \sim \mathcal{U}([0,1]^d)$

Demonstration:

Let h be a measurable function from $\mathbb{R}^d$ to $\mathbb{R}$

$E[h(Y)] = E[h(T(X))] = \int_{\mathbb{R}^d} h(T(x))f_X(x)dx$

We consider the change of variable $y = T(x)$.

The Jacobian matrix $[\frac{dy}{dx}]$ is triangular and the diagonal elements are:

- $\frac{\partial}{\partial x_1} P(X_1 \leq x_1) = f_{X_1}(x_1)$
- $\frac{\partial}{\partial x_2} P(X_2 \leq x_2 | X_1 = x_1) = f_{X_2|X_1=x_1}(x_2) \cdots$
- $\frac{\partial}{\partial x_d} P(X_d \leq x_d | X_{d-1} = x_{d-1} \cdots, X_1 = x_1) = f_{X_d|(X_{d-1}=x_{d-1}\cdots)}(x_d)$

so, the determinant of the Jacobian Matrix equals $f_X(x_1, x_2, \cdots, x_d)$

so after the change of variable : $E[h(Y)] = \int_{T(\mathbb{R}^d)} h(y)dy$

as the y_i are probabilities $T(\mathbb{R}^d) \subset [0,1]^d$ and by mass conservation

$T(\mathbb{R}^d) = [0,1]^d$

so $\forall h, E[h(Y)] = \int_{[0,1]^d} h(y)dy \Longrightarrow Y \sim \mathcal{U}([0,1]^d)$ Q.E.D.

In some situations a Copula C is defined analytically as any function satisfying the properties of a cdf of a variable taking its values in $[0,1]^d$ and whose marginals are $\mathcal{U}([0,1])$

Exemples of Copula :

- Clayton $C(u,v) = max(u^{-\theta} + v^{-\theta} - 1, 0)^{-\frac{1}{\theta}}$ with $\theta > -1$ and $\theta \neq 0$
- Gumbel-Hougaard $C(u) = exp\Big(\big[-\sum_{i=1}^{i=d}(-ln(u_i))^{\theta}\big]^{\frac{1}{\theta}}\Big)$
  with $\theta > 1$ and the conventions $ln(0) = -\infty$ and $exp(-\infty) = 0$
- Archimedean $C(u) = \psi(\sum_{i=1}^{i=d} \psi^{-1}(u_i))$
  with $\psi : [0,\infty] \longrightarrow [0,1]$ satisfying (among other things) $\psi(0) = 1$ and $\psi(\infty) = 0$

Remarks

- there are some conditions C must satisfy to be a copula. For a two dimension copula we need at least $C(u,0) = C(0,v) = 0$ and $C(u,1) = u$ and $C(1,v) = v$
- there are some conditions the function ψ must satisfy for the Archimedean expression to be a copula
- the Clayton and Gumbel-Hougaard copulas are two particular cases of Archimedean copulas. For the Clayton Copula $\psi(\theta) = \frac{1}{\theta}(u^{-\theta} - 1)$

Archimedean Copulas

Theorem (admitted)

If Z is a random variable with $Z > 0$ and if
$\psi_Z(s) = E[exp(-sZ)]$ for $s \in [0, \infty]$ (Laplace transform)
Then,
the Archimedean function defined by ψ_Z is a copula

exercise:
Show that if $Z \sim Gamma(\frac{1}{\theta}, 1)$ with $0 < \theta < +\infty$ then the Archimedean Copula generated by ψ_Z is the Clayton Copula of parameter θ

Solution:

$$\psi_Z(s) = \int_0^{+\infty} \frac{z^{\frac{1}{\theta}-1} e^{-z}}{\Gamma(\frac{1}{\theta})} e^{-sz} dz = (1+z)^{-\frac{1}{\theta}}$$

Remarks: Archimedean Copulas can be created "on demand" by calculating the Laplace transform of any arbitrary random variable Z

Simulating Copulas

Even if the simulation of Gaussian Copulas is easy it may be more complicated to simulate arbitrary copulas.
The Rosenblatt Theorem provides an easy way to simulate copulas in dimension 2.

Proposition

let $U = (U_1, U_2)$ of copula C with $U_i \sim \mathcal{U}([0,1])$
If T is the Rosenblatt's transformation

- $T(U) = (U_1, \frac{\partial C}{\partial u_1}(U_1, U_2))$ and the components are i.i.d $\mathcal{U}([0,1])$

So, if we take $W \sim \mathcal{U}([0,1])$ independent from U_1 then

- $T^{-1}(U_1, W) \sim U$

Remark: in practice for each simulation (u_1, w) we find u_2 the solution of $\frac{\partial C}{\partial u_1}(u_1, u_2) = w$ and by doing so we simulate U of Copula C.

Background : Pearson's linear correlation $\rho(X,Y) = \frac{cov(X,Y)}{\sigma(X)\sigma(X)}$ measures only the affine relationship between variables and presents some imperfections to measure the "link" between two variables. For example:

- if $X \sim \mathcal{N}(0,1)$ and $Y = X^2$ then $\rho(X,Y) = 0$ while there is a strong link between Y and X (we can indeed predict Y perfectly from X)
- if f and g are increasing in general $cov(X,Y) \neq cov(f(X), g(Y))$

Definition : Kendall's tau

Let (X,Y) be a random variable. Let (X_1,Y_1), (X_2,Y_2) be independent with the same law as (X,Y). We call Kendall's tau and note $\tau(X,Y)$ the quantity $P((X_1 - X_2)(Y_1 - Y_2) \geq 0) - P((X_1 - X_2)(Y_1 - Y_2) < 0)$

Properties Kendall's τ

- $-1 \leq \tau(X, Y) \leq 1$
- if f and g are strictly increasing $\tau(f(X), g(Y)) = \tau(X, Y)$
- if f_X and f_Y are strictly positive, $\tau(F_X(X), F_Y(Y)) = \tau(X, Y)$
- $U \sim \mathcal{U}([0, 1) \Longrightarrow \tau(U, U) = 1$ and $\tau(U, 1 - U) = -1$
- If (X, Y) has C for Copula then $\tau(X, Y) = -1 + 4 \int_{[0,1]^2} C(u, v) \frac{\partial^2 C}{\partial u \partial v} dudv$

Demonstration: let's show the last point
as $\tau(F_X(X), F_Y(Y)) = \tau(X, Y)$ we can show it for X and $Y \sim \mathcal{U}([0,1])$
$\tau(X, Y) = P((X_1 - X_2)(Y_1 - Y_2) \geq 0) - (1 - P((X_1 - X_2)(Y_1 - Y_2) \geq 0))$
$= -1 + 2P((X_1 - X_2)(Y_1 - Y_2) \geq 0)$
$= -1 + 2(P(X_1 - X_2 \leq 0, Y_1 - Y_2 \leq 0) + P(X_2 - X_1 \leq 0, Y_2 - Y_1 \leq 0))$
as (X_1, Y_1) and (X_2, Y_2) have the same law, so we just need to calculate the first probability.
$P(X_1 - X_2 \leq 0, Y_1 - Y_2 \leq 0) = E(E(1_{X_1 \leq X_2} 1_{Y_1 \leq Y_2}) | X_2, Y_2))$ and
$E(1_{X_1 \leq x_2} 1_{Y_1 \leq y_2}) = P(X_1 \leq x_2, Y_1 \leq y_2) = C_X(x_2, y_2)$
so we have to calculate $E(C(X_2, Y_2)) = \int_{[0,1]^2} C(u, v) \frac{\partial^2 C}{\partial u \partial v} dudv$
so $\tau(X, Y) = -1 + 4 \int_{[0,1]^2} C(u, v) \frac{\partial^2 C}{\partial u \partial v} dudv$ Q.E.D.

Remark Kendall's tau:
Based on the observations $(x_i, y_i)_{i \in [\![1,n]\!]}$ the Kendall's tau is estimated by the quantity

$$\frac{2}{n(n-1)} \sum_{i<j} sgn(x_i - x_j) sgn(y_i - y_j)$$

Copulas : Measures of Association between Variables

Definition : Spearman's correlation

If (X, Y) is a random variable with marginal laws F_X and F_Y then the Spearman's correlation ρ_S is defined by $\rho_S(X, Y) = \rho(F_X(X), F_Y(Y))$

Properties Spearman's correlation

- $-1 \leq \rho_S(X, Y) \leq 1$
- if f and g are strictly increasing $\rho_S(f(X), g(Y)) = \rho_S(X, Y)$
- $\tau(F_X(X), F_Y(Y)) = \rho_S(X, Y)$
- $U \sim \mathcal{U}([0, 1) \Longrightarrow \rho_S(U, U) = 1$ and $\rho_S(U, 1 - U) = -1$
- If (X, Y) has C for Copula then $\rho_S(X, Y) = -3 + 12 \int_{[0,1]^2} C(u, v) dudv$

Demonstration :

Let C be the copula of (X, Y) i.e the cdf of (U, V) where $U = F_1(X)$ and $V = F_2(Y)$, let $\rho_S(X, Y) = \frac{E(UV) - E(U)E(V)}{E(U)E(V)}$

$$E(UV) = \int_0^1 \int_0^1 uv \frac{\partial^2 C}{\partial u \partial v} du dv = \int_0^1 u \Big(\int_0^1 v \frac{\partial^2 C}{\partial u \partial v} dv \Big) du$$

$$= \int_0^1 u \Big([v \frac{\partial C}{\partial u}]_0^1 - \int_0^1 \frac{\partial C}{\partial u} dv \Big) du = \int_0^1 u \Big(f_U(u) - \int_0^1 \frac{\partial C}{\partial u} dv \Big) du$$

$$= E(U) - \int_0^1 \Big(\int_0^1 u \frac{\partial C}{\partial u} du \Big) dv = E(U) - \int_0^1 \Big([uC(u, v)]_0^1 - \int_0^1 C(u, v) du \Big) dv$$

$$= E(U) - \int_0^1 P(V \leq v) dv + \int_0^1 \int_0^1 C(u, v) du dv$$

$$= E(U) - E(V) + \int_0^1 \int_0^1 C(u, v) du dv = \int_0^1 \int_0^1 C(u, v) du dv$$

and $E(U)E(V) = \frac{1}{4}$ and $Var(U) = Var(V) = \frac{1}{12}$ Q.E.D.

Remark: Spearman's correlation.
Based on the observations $(x_i, y_i)_{i \in [\![1,n]\!]}$ the Spearman's correlation is estimated by calculating the correlations of the

$$(F_{x[n]}(x_i), F_{y[n]}(y_i))$$

where $F_{x[n]}(x_i)$ is the quantile for x_i amongst $x_1, x_2, \cdots, x_n$ and $F_{y[n]}(y_i)$ is the quantile for y_i amongst $y_1, y_2, \cdots, y_n$.

Appendix : Risk Neutral Probability and Utility functions

Background: We consider an economy with two instants $\{0, 1\}$ where there are d assets whose vector of prices X is represented today by the vector $X_0 = (X_0^1, X_0^2, \cdots, X_0^d)'$. We assume that at instant 1 there are n possible states for the economy and for each state $i \in [\![1, n]\!]$ the vector of the prices of the assets is $X_i = (X_i^1, X_i^2, \cdots, X_i^n)'$. We assume that prices are all strictly positive.

Definition: Absence of Arbitrage (AOA)

We say that there is no arbitrage in the economy iff:
$\{w \in \mathbb{R}^d, \forall i \in [\![1, n]\!] \langle w, X_i \rangle \geq 0\} \subset \{w \in \mathbb{R}^d, \langle w, X_0 \rangle \geq 0\}$

Remarks :

The definition means that it is not possible to receive money today to build a strategy which has positive values tomorrow in all cases.

Risk Neutral Probability (discrete case)

Theorem and Definition: Risk Neutral Probability

a)the two following propositions are equivalent:

- there is no arbitrage in the economy
- we can find $(\lambda_i)_{i\in[\![1,n]\!]}, \lambda_i \geq 0$ such that $X_0 = \sum_{i=1}^{i=n} \lambda_i X_i$

b) if there is a risk-free asset in the economy of return r over $[0, 1]$ then:

- $\sum_{i=1}^{i=n} \lambda_i = \frac{1}{1+r}$
- if we define a probability π, over the n possible values of X at time 1, by $\pi_i = \lambda_i(1+r)$ then $X_0 = \frac{1}{1+r} E_\pi[X]$ and π is called the risk neutral probability for the economy.

Risk Neutral Probability (discrete case)

Demonstration :

a) one of the implications is obvious.

We assume now that there is no arbitrage and define the cone

$\mathcal{C} = \{\sum_{i=1}^{i=n} \lambda_i X_i, \forall i \in [\![1, n]\!] \lambda_i \geq 0\}$.

Then $\mathcal{C}$ is convex and if $X \notin \mathcal{C}$ we can separate X from $\mathcal{C}$ by an hyperplane and find $w \in \mathbb{R}^d$ such that: $\langle w, X_0 \rangle < 0$ and for all X_i in $\mathcal{C}$ $\langle w, X_i \rangle > 0$ but this would contradict the AOA hypothesis, so $X_0 \in \mathcal{C}$. Q.E.D.

b) if we assume that the risk-free asset is component j then:

$X_0^j = \sum_{i=1}^{i=n} \lambda_i X_i^j$ and $X_i^j = X_0^j(1+r) \Longrightarrow \sum_{i=1}^{i=n} \lambda_i = \frac{1}{1+r}$ Q.E.D.

Exercise : we assume that there are 3 assets, of prices today $X_0 = (1, 5, 10)'$ and 3 possible states of the economy tomorrow defined by the 3 vector of prices for the assets:
$X_1 = (1.03, 5, 11)'$, $X_2 = (1.03, 5, 10)'$, $X_3 = (1.03, 6, 10)'$.
a) explain why the risk-free rate is 3%
b) show that $\pi = (0.30, 0.55, 0.15)'$
c) explain why there is no arbitrage in this economy.

Remark : If we assume that the vector of prices today is $X_0 = (1, 5, 10)'$ and that the 3 vectors of prices for tomorrow are :
$X_1 = (1.03, 5, 10)'$, $X_2 = (1.03, 6, 12)'$, $X_3 = (1.03, 6, 13)'$ then
a) $\pi = (0.85, 0.15, 0)'$ is the risk neutral probability
b) according to a) there is no arbitrage
c) the strategy $w = (0, -2, 1)'$ costs today zero and the possible outcomes tomorrow are 0 for the first two states and 1 for state 3, so it seems attractive to play it (as there is only upside) but strictly speaking this is not an arbitrage according to our definition.

Risk Neutral Probability (continuous case)

Background:
We consider a probability space $(\Omega, \mathcal{F}, P)$ with $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ where $\mathcal{F}_t$ represents the information available at time t.
We assume that there are d financial assets following the equations:
$dX_s^i = \mu_s^i X_s^i ds + \sigma_s^i X_s^i dW_s^i$ where $W_s = (W_s^1, W_s^2, \cdots, W_s^d)$ is a d-dimensional Brownian motion.

Theorem and Definition : Risk Neutral Probability

We can find a probability Q on $(\Omega, \mathcal{F})$ such that:
W^* defined by: $dW_s^{i*} = (dW_s^i + \frac{\mu_s^i - r_s}{\sigma_s^i} ds)$ is a Brownian motion under Q
We can then re-write the model:
$dX_s^i = r_s X_s^i ds + \sigma_s^i X_s^i dW_s^{*i}$ where W^* is a Brownian motion under Q and Q is called the risk neutral probability.

Risk Neutral Probability (continuous case)

Lemma and Definition

If $(Z_s)_{s\geq 0}$ is a martingale under P with $Z_s \geq 0$ and $Z_0 = 1$ and if we define Q for any random variable Y_t $\mathcal{F}_t$-measurable by $E_Q[Y_t] = E_P[Y_t Z_t]$ then:

- Q is a probability on $(\Omega, \mathcal{F})$
- $E_Q[Y_T|\mathcal{F}_t] = E_P[Y_T \frac{Z_T}{Z_t}|\mathcal{F}_t]$

Usually we note $(\frac{dQ}{dP})_t = Z_t$ and so we write $E_Q[Y_t] = E_P[Y_t(\frac{dQ}{dP})_t]$

Demonstration Lemma : easy

Demonstration Theorem (hint) :

We note $\Delta_s^i = \frac{\mu_s^i - r_s}{\sigma_s^i}$ and $\Delta_s = (\Delta_s^1, \Delta_s^2, \cdots, \Delta_s^d)'$.

Risk Neutral Probability (continuous case)

We search for a probability Q under which $(W_s^*)_{s\geq 0}$ is a Brownian motion
For that we need $E_Q[dW_t^*|\mathcal{F}_t]=0$ and $E_Q[dW_t^*(dW_t^*)'|\mathcal{F}_t]=Id_{\mathbb{R}^d}dt$
We note $(\frac{dQ}{dP})_t=Z_t$ and so search for Z_t.
$E_Q[dW_t^*|\mathcal{F}_t]=E_P[dW_t^*\frac{Z_{t+dt}}{Z_t}|\mathcal{F}_t]=E_P[(dW_t+\Delta_t dt)(1+\frac{dZ_t}{Z_t})|\mathcal{F}_t]$
$=E_P[dW_t|\mathcal{F}_t]+\frac{1}{Z_t}E_P[dW_t dZ_t|\mathcal{F}_t]+\Delta_t dt=\frac{1}{Z_t}E_P[dW_t dZ_t|\mathcal{F}_t]+\Delta_t dt$
If we search Z of the form $dZ_s=\langle B_s, dW_s\rangle$ with $B_s\in\mathbb{R}^d$ (no drift term as martingale) then:
$E_P[dW_t dZ_t|\mathcal{F}_t]=E_P[dW_t\langle dW_t, B_t\rangle|\mathcal{F}_t]=E_P[dW_t(dW_t)'B_t|\mathcal{F}_t]$
$=E_P[dW_t(dW_t)'|\mathcal{F}_t]B_t=B_t dt$ so, $E_Q[dW_t^*|\mathcal{F}_t]=0\iff B_t=-\Delta_t Z_t$
Solving $dZ_s=\langle -\Delta_t, dW_s\rangle Z_s$ and $Z_0=1$ we get:
$Z_t=exp\left(\int\limits_0^t -\langle\Delta_s, dW_s\rangle-\frac{1}{2}\int\limits_0^t ||\Delta_s||^2 ds\right)$ (we do not discuss here the conditions on Δ_s for integrability that can be found in Girsanov's theorem)
The condition $E_Q[dW_t^*(dW_t^*)'|\mathcal{F}_t]=Id_{\mathbb{R}^d}dt$ is easy to verify Q.E.D.

Remark 1: for $d = 1$ we get that:

- $X_T = X_0 e^{\mu T} e^{\sigma W_T - \frac{1}{2}\sigma^2 T}$ where $(W_s)_{s \geq 0}$ is a Brownian under P
- $X_T = X_0 e^{rT} e^{\sigma W_T^* - \frac{1}{2}\sigma^2 T}$ where $(W_s^*)_{s \geq 0}$ is the Brownian under Q defined by $W_T^* = W_T + \frac{\mu - r}{\sigma} T$
- $(\frac{dQ}{dP})_T = exp(\frac{r-\mu}{\sigma} W_T - \frac{1}{2}(\frac{r-\mu}{\sigma})^2 T)$
- for any function h, $E_Q[h(X_T)] = E_P[h(X_T)(\frac{dQ}{dP})_T]$

Exercise: verify by calculations for $d = 1$ that
$E_Q[h(rT + \sigma W_T^*)] = E_P[h(\mu T + \sigma W_T)(\frac{dQ}{dP})_T]$
Solution: $E_Q[h(rT + \sigma W_T^*)] = E_{Q^{W_T^*}}[h(rT + \sigma z)]$
$= \int h(rT + \sigma z)\frac{1}{2\pi\sqrt{T}} exp(-\frac{z^2}{2T})dz$

Risk Neutral Probability (continuous case)

If we take the new variable u such that $\mu T + \sigma u = rT + \sigma z$ we get :

$\circ$ $h(rT + \sigma z) = h(\mu T + \sigma u)$

$\circ$ $exp(-\frac{z^2}{2T}) = exp(-\frac{(u+\frac{\mu-r}{\sigma}T)^2}{2T}) = exp(-\frac{u^2}{2T})exp(-\frac{\mu-r}{\sigma}u - \frac{1}{2}(\frac{\mu-r}{\sigma})^2 T)$

so,

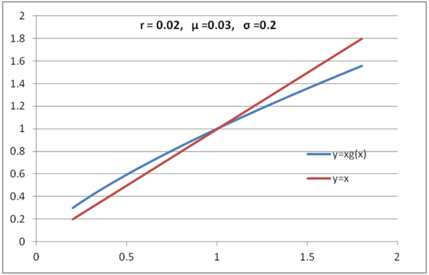
$\int h(rT + \sigma z)\frac{1}{2\pi\sqrt{T}}exp(-\frac{z^2}{2T})dz$

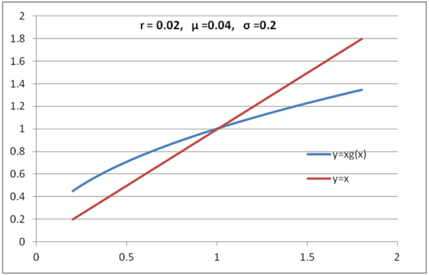
$= \int h(\mu T + \sigma u)exp(-\frac{\mu-r}{\sigma}u - \frac{1}{2}(\frac{\mu-r}{\sigma})^2 T)\frac{1}{2\pi\sqrt{T}}exp(-\frac{u^2}{2T})du$

$= E_P[h(\mu T + \sigma W_T)exp(-\frac{\mu-r}{\sigma}W_T - \frac{1}{2}(\frac{\mu-r}{\sigma})^2 T)]$ Q.E.D.

Remark 2: we call u a utility function compatible with the price of asset X, P the "real" probability and "Q" the risk neutral probability.
As $X_0 = E_P[e^{-rT}u(X_T)]$ and $X_0 = E_Q[e^{-rT}X_T]$ we have under the previous assumptions concerning the law of X under P :
$E_P[e^{-rT}u(X_T)] = E_P[e^{-rT}X_T f(W_T)]$ with
$f(w) = exp(\frac{r-\mu}{\sigma}w - \frac{1}{2}(\frac{r-\mu}{\sigma})^2 T)$
if we define $g(x) = f(\frac{1}{\sigma}[ln(\frac{x}{X_0}) + (\frac{\sigma^2}{2} - \mu)T])$ then
$f(W_T) = g(X_T)$ and $E_P[e^{-rT}u(X_T)] = E_P[e^{-rT}X_T g(X_T)]$
so, $u(x) = xg(x)$ is an adequate utility function for this modelisation of X.
In the following graph we represent $xg(x)$ for various values of r, μ and σ with $x_0 = 1$ and $T = 1$.
Note that depending on the value of the parameters $xg(x)$ is not always increasing (which shows its limits in terms of admissible utility function..)

Risk Neutral Probability (continuous case)





Utility functions derived from Girsanov's Theorem

References

David Lando
Credit Risk Modeling
Princeton Series in Finance 2004, pp.310

David Lando
On Cox Processes and Credit Risky Securities
Review of Derivatives Research 1998, pp.22

Mark Davis, Violet Lo
Infectious Defaults
Tokyo-Mitsubushi International plc 1999, pp.12

Mark Davis, Violet Lo
Moody's Correlated Binomial Default Distribution
Moody's Investors Services August 10, 2004, pp.12

Jeremy Graveline, Michael Kokalari
Credit Risk
The Research Foundation of CFA Institute 2010, pp.22

Donald MacKenzie and Taylor Spears
The Formula that Killed Wall Street

References

Johan Segers
Copulas: An Introduction
Columbia University (9-11 Oct 2013), pp.1-74

Johan Segers
Copulas: An Introduction Part II: Models
Columbia University (9-11 Oct 2013), pp.1-65

Erik Bolviken
Copulas
University of Oslo (29 Sep 2010), pp.1-11