… solutions for robust engineering

Probability theory basics

Gaëtan Blondet, Phimeca Engineering SA

‘HPC and Uncertainty Treatment – Examples with Open TURNS and Uranie’

EDF – Phimeca – Airbus Group – IMACS – CEA

PRACE Advanced Training Center – May, 27-29 2019

MAISON DE LA SIMULATION

Motivation

- Uncertainty is here defined in a broad sense. It is meant to include variability, uncertainty and lack-of-knowledge.

  - *Aleatory* uncertainty: intrinsic randomness, variance of a phenomenon
    - Due to lack of control over environmental variability and test settings (temperature, humidity, etc.), and to errors made during testing.
    - Can be better characterized but cannot be reduced by taking more measurements or performing more simulations.

  - *Epistemic* uncertainty: lack-of-knowledge, ambiguity, haziness.
    - Due to lack-of-knowledge about materials, loads, initial conditions, etc. and to assumptions made during testing and modeling.
    - Can be reduced by collecting more information and evidence.

These sources of uncertainty are modeled and propagated by means of *probability theory*

Note: *Other theories have been developed to represent epistemic uncertainty such as Imprecise Theory (IP), Possibility theory, Fuzzy sets and fuzzy logic.*

Outline

General definitions

- Random experiment – Event
- Measurable space
- Kolmogorov axioms
- Bayes' theorem

Random variables

- Definitions
- Cumulative distribution function (CDF) and probability density function (PDF)
- Discrete / continuous random variables
- Statistical moments
- Confidence intervals (CI)

Random vectors

- Definitions
- Moments
- Copulas

Outline

General definitions

- Random experiment – Event
- Measurable space
- Kolmogorov axioms
- Bayes' theorem

Random variables

- Definitions
- Cumulative distribution function (CDF) and probability density function (PDF)
- Discrete / continuous random variables
- Statistical moments
- Confidence intervals (CI)

Random vectors

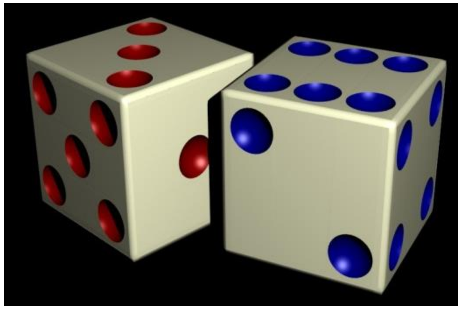
- Definitions
- Moments
- Copulas

Definitions

Random experiment

- A ***random experiment*** is a repeatable procedure that has more than one possible ***outcome***. The result is aleatory.
- A ***realization*** or an ***outcome*** is a possible result of an experiment.
- The ***sample space*** is the set of all possible outcomes of the experiment. It is commonly referred by Ω

Throwing of two six-sided dice



Sample space:

✓ The pairs of faces

$$\Omega = \{(1,1), (1,2), (1,3), \ldots\}$$

✓ The sum of the faces

$$\Omega = \{2,3,4,5,6,7,8,9,10,11,12\}$$

Definitions

Event

An ***event*** is a set of outcomes of an experiment (a subset of the sample space Ω). The set of events is denoted Φ.

Throwing of 2 six-face-dice

A_1 = "Do an even number"

A_2 = "Do more than 2"

Some ***particular events***:

Ω Certain event
The sum of the two dice is less than or equal to 12

$\emptyset$ Impossible event
The sum of the two dice is strictly less than to 2

$\{\omega\}$ Simple event
The sum of the two dice is equal to 12

Composed event
Any event whose cardinality is strictly more than 1

Definitions

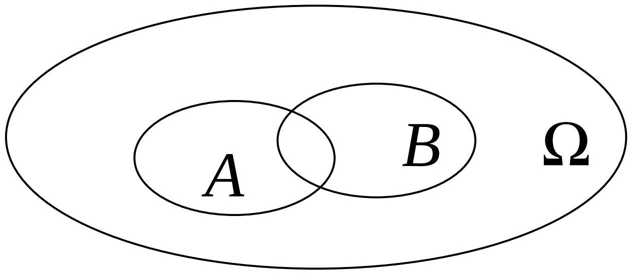
Operations of events

Operators

$A \cup B$: union

$A \cap B$: intersection

$\bar{A}$: complement



Venn diagram

Properties

- Any finite or countable union or intersection of events is an event.
- If $A \cap B = \emptyset$ then A and B are disjoints
- *Commutativity, associativity, distributivity* of intersection over union and De Morgan's laws:

$$A \cup B = B \cup A \quad \text{et} \quad A \cap B = B \cap A \qquad (A \cup B) \cup C = A \cup (B \cup C)$$

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C) \qquad \overline{A \cap B} = \bar{A} \cup \bar{B} \quad \text{et} \quad \overline{A \cup B} = \bar{A} \cap \bar{B}$$

Definitions

Partitions of Ω

- A and B form a *partition* of Ω if and only if they are *mutually exclusive* and *collectively exhaustive*:

$$A \cap B = \emptyset \qquad \text{and} \qquad A \cup B = \Omega$$

Set of measurable spaces (or σ – algebra)

A set of events $\mathcal{F}$ belonging to the set of parts of Ω is measurable if and only if:

- $\emptyset \in \mathcal{F}$ and $\Omega \in \mathcal{F}$
- If A_i is a sequence in $\mathcal{F}$ then $\bigcup_i A_i \in \mathcal{F}$ and $\bigcap_i A_i \in \mathcal{F}$
- If $A \in \mathcal{F}$ then $\overline{A} \in \mathcal{F}$

If $\mathcal{F}$ is measurable in Ω then $(\mathcal{F}, \Omega)$ is a *measurable space.*

Definitions

Kolmogorov axioms

A *probability measure* allows to associate numbers to events, i.e. *their probability of occurrence.*

It is defined as an application $\mathbb{P} : \mathcal{F} \longmapsto [0,1]$ satisfying the *Kolmogorov axioms:*

- $\forall A \in \mathcal{F}, 0 \leq \mathbb{P}(A) \leq 1$
- $\mathbb{P}(\Omega) = 1$
- For any set of finite or countable of disjoints events A_i

$$\mathbb{P}[\cup_i A_i] = \sum_i \mathbb{P}[A_i]$$

The *probability space* thus build is denoted $(\Omega, \mathcal{F}, \mathbb{P})$

Throwing 2 dice

A_1= « Do an even number »

$\mathbb{P}[A_1] = \frac{1}{2}$

A_2= « Do more than 2

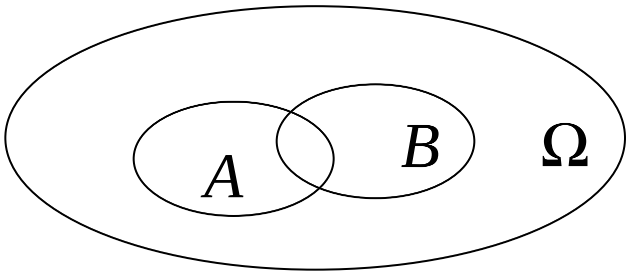
$\mathbb{P}[A_2] = \frac{35}{36}$

Definitions

- From the Kolmogorov axioms, the elementary results hold:

  - $\mathbb{P}[\emptyset] = 0$
  - $\mathbb{P}[\omega] = 1$
  - $\mathbb{P}[\overline{A}] = 1 - \mathbb{P}[A]$
  - $\mathbb{P}[A \setminus B] = \mathbb{P}[A] - \mathbb{P}[A \cap B]$
  - $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]$

  - $A \subseteq B \implies \mathbb{P}[A] \leq \mathbb{P}[B]$



Definitions

- Frequentist interpretation of probabilities

The *probability of an event* is the limit of its *empirical frequency* of occurrence

Throwing 2 dice:

- A random experiment is made N times
- The event A_1 = « Do an even number » is observed
- N_{A_1} is the number of times for which the event A_1 is observed

$$\mathbb{P}[A_1] = \lim_{N \to \infty} \frac{N_{A_1}}{N}$$

Definitions

Conditional probability

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$$

« probability of A given B »

Independence

- Two events A and B are independent when the occurrence of B does not affect the probability of occurrence of A, and vice versa:

$$\mathbb{P}[A|B] = \mathbb{P}[A] \Longrightarrow \mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$$

Definitons

Bayes' theorem

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[B|A]}{\mathbb{P}[B]} \mathbb{P}[A]$$

Initial probability of *A*

Influence of the information contained in B

It is the probability of A that is *updated by the knowledge of the occurrence of B*

- Proof:

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

Outline

- General definitions
  - Random experiment – Event
  - Measurable space
  - Kolmogorov axioms
  - Bayes' theorem
- Random variables
  - Definitions
  - Cumulative distribution function (CDF) and probability density function (PDF)
  - Discrete / continuous random variables
  - Statistical moments
  - Confidence intervals (CI)
- Random vectors
  - Definitions
  - Moments
  - Copulas

Random variables

Definition

A random variable is a *measurable function*:

$$\begin{aligned} X \ : \Omega &\longrightarrow \mathbb{X} \\ \omega &\longmapsto x = X(\omega) \end{aligned}$$

A *discrete random variables* can take either a finite or at most a countably infinite set of discrete values

$$\mathbb{X} \subseteq \mathbb{Z}$$

Examples : sum of two dice, rupture cycles number

Continuous random variables take on values that vary continuously within one or more real intervals

$$\mathbb{X} \subseteq \mathbb{R}$$

Examples : Young modulus of a material, value of loading applied on a structure.

Random variables

Cumulative distribution function

It is the function that relates x to the probability that the random variable X takes on a value less than or equal to x

$$F_X(x) = \mathbb{P}[X \leq x]$$

Probability density function

Discrete case: it is the function that relates x to the probability that the random variable X takes on a given value *equal* to x:

$$p_X(x) = \mathbb{P}[X = x]$$

Continuous case: it is the function that relates x to the probability that the random variable X belongs to the infinitesimal interval $[x, x + \mathrm{dx}]$.

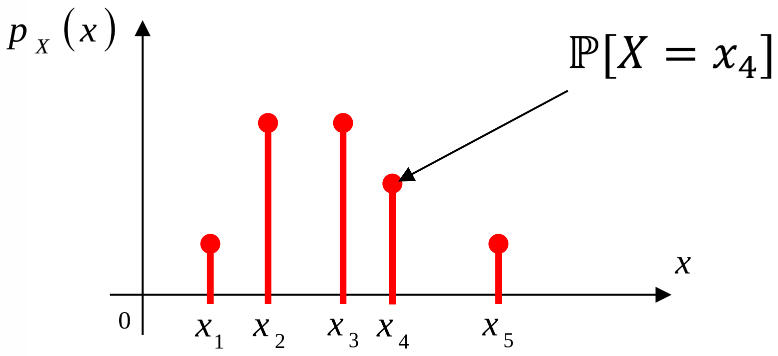
$$f_X(x)\mathrm{dx} = \mathbb{P}(x < X \leq x + \mathrm{dx})$$

And $f_X(x)$ is the *derivative of the cumulative distribution function* :

$$\boxed{f_X(x) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x}}$$

Random variables

Discrete random variable



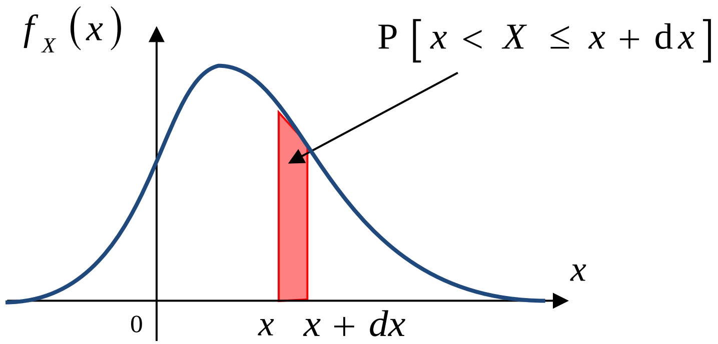
$$\boxed{p_X(x_i) = \mathbb{P}[X = x_i]}$$

p_X = *probability mass function*

$$\forall x_i,\ 0 \leq p_X(x_i) \leq 1$$

$$\sum_{x_i} p_X(x_i) = 1$$

Continuous random variable



$$\boxed{f_X(x)\,\mathrm{d}x = \mathbb{P}[x < X < x + \mathrm{d}x]}$$

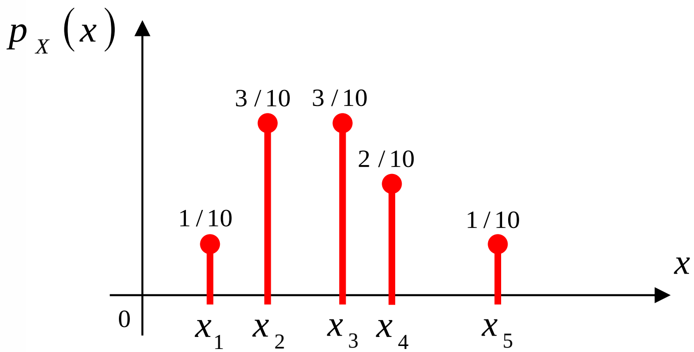
f_X = *probability density function*

$$\forall x,\ f_X(x) \geq 0$$

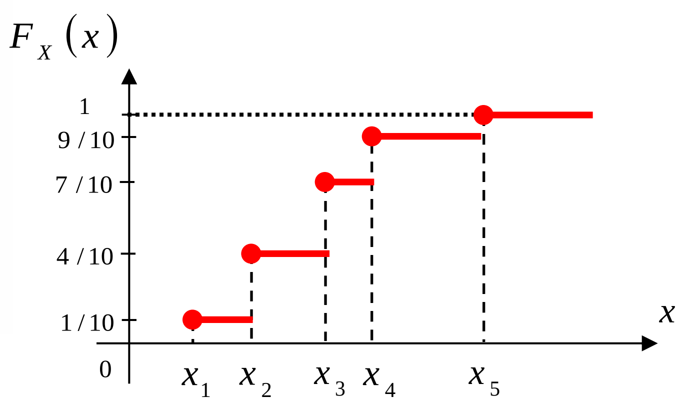
$$\int_{x \in \mathrm{X}} f_X(x)\,\mathrm{d}x = 1$$

Random variables

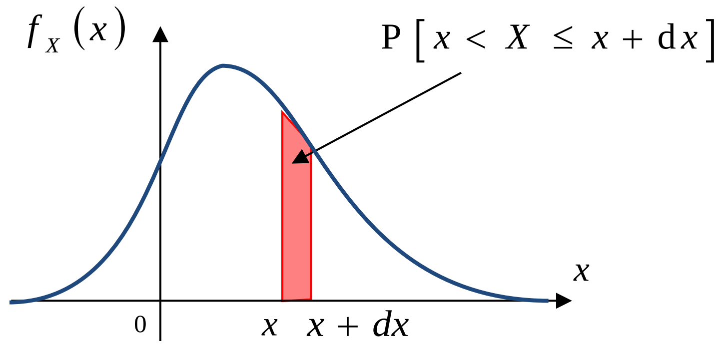
Discrete random variable



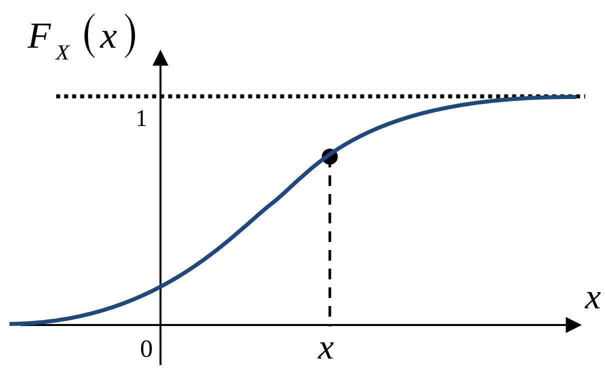
$$F_X(x) = \mathbb{P}[X \leq \mathrm{x}] = \sum_{x \leq x_i} p_X(x_i)$$



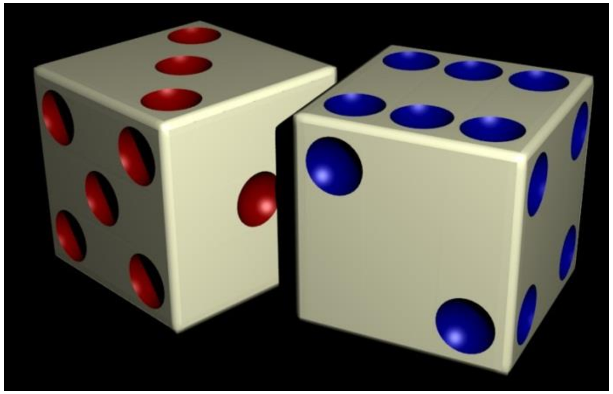
Continuous random variable



$$F_X(x) = \mathbb{P}[X \leq \mathrm{x}] = \int_{-\infty}^{x} f_X(x)\mathrm{dx}$$

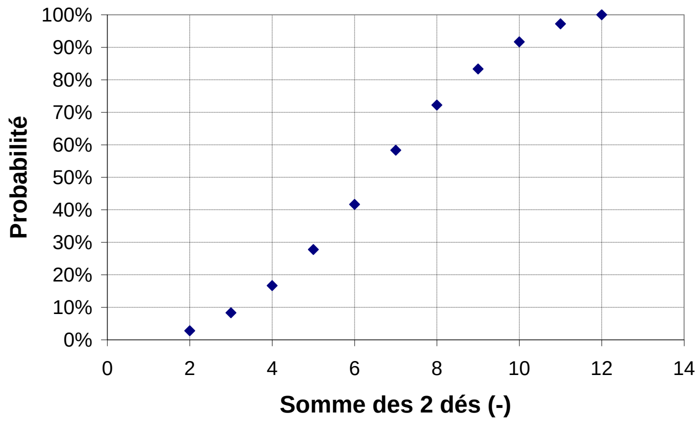


Random variables



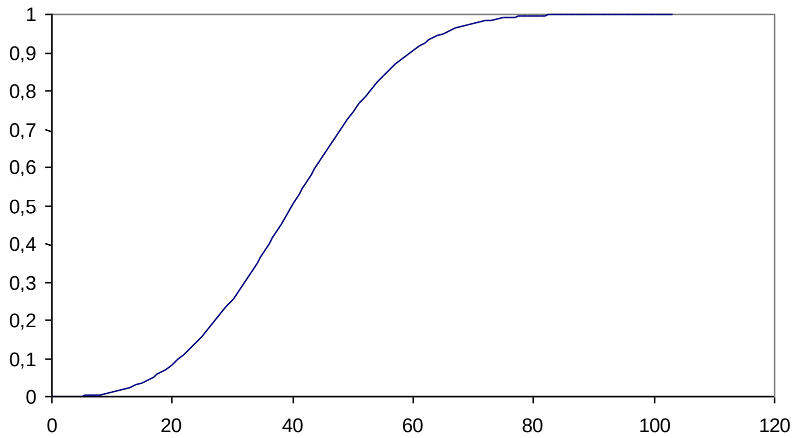
Discrete variable

Sum of 2 dice: $\Omega \rightarrow \{2,\ldots,12\}$



Continuous variable

Wind speed: $\Omega \longrightarrow \mathbb{R}^+$



Random variables

Sum of 2 independent random variables

Given two independent continuous random variables X and Y, the sum S of the two variables is:

$$S = X + Y$$

The probability density function of the sum of the random variable S is the *convolution* of the two separate density functions of X and Y.

$$\boxed{f_S = f_X * f_Y}$$

Where the convolution is defined as:

$$f_S(y) = \int f_X(x) f_Y(x-y) \mathrm{d}x$$

The convolution is *commutative*:

$$f_X * f_Y = f_Y * f_X$$

Random variables

Theorem of « composition of laws »

Let X be a *continuous random variable* and φ *a continuously differentiable strictly monotonic function.* The random variable $Y = \varphi(X)$ has a probability density function:

$$\boxed{f_Y(y) = f_X(x)\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right|}$$

- **Proof:**

Case of strictly increasing function:

From the rules on inequallities: $X \leq x \Rightarrow \varphi(X) \leq \varphi(x) \Rightarrow Y \leq y$

Thus : $\mathrm{P}[Y \leq y] = \mathrm{P}[X \leq x] \Leftrightarrow F_Y(y) = F_X(x)$

By derivating according to y: $\frac{\mathrm{d}F_Y(y)}{\mathrm{d}y} = \frac{\mathrm{d}F_X(x)}{\mathrm{d}y} \Leftrightarrow f_Y(y) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}y} = f_X(x)\frac{\mathrm{d}x}{\mathrm{d}y}$

Case of strictly decreasing function :

From the rules on inequallities: $X > x \Rightarrow \varphi(X) \leq \varphi(x) \Rightarrow Y \leq y$

Thus :$\mathrm{P}[Y \leq y] = \mathrm{P}[X > x] \Leftrightarrow F_Y(y) = 1 - F_X(x)$

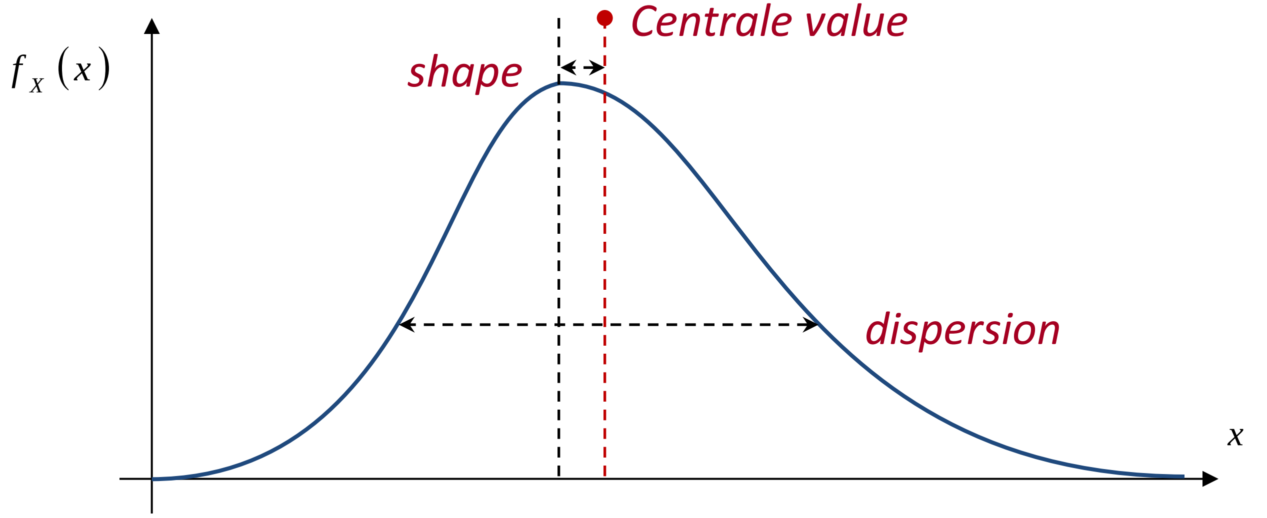
By derivating according to y: $\frac{\mathrm{d}F_Y(y)}{\mathrm{d}y} = -\frac{\mathrm{d}F_X(x)}{\mathrm{d}y} \Leftrightarrow f_Y(y) = -\frac{\mathrm{d}F_X(x)}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}y} = -f_X(x)\frac{\mathrm{d}x}{\mathrm{d}y}$

Random variables

Characterization of a random variable

A probability distribution is characterized by a number of features:

- Its *central value*
- its *dispersion*
- its *shape* (asymmetry, shift, etc.)



Generally, one will define *statistical moments* to characterized some features related to a random variable's probability distribution.

Random variables

Expected value (definition)

To define the statistical moments, one introduces the « *expectation* » operator denoted $\mathbb{E}$ for a random variable (under some conditions).

Case of *discrete random variables:*

$$\mathbb{E}[X] = \sum_{x_i} x_i p_X(x_i)$$

(if the sum converges)

Case of *continuous random variables:*

$$\mathbb{E}[X] = \int_{x \in \mathbb{X}} x f_X(x) \mathrm{d}x$$

(if the integral converges)

Random variables

Expected value (properties)

Given X and Y, two random variables and a and b two reals.

- The expected value operator is *linear*:

$$\boxed{\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]}$$

- Caution, *in the general case:*

$$\boxed{\mathbb{E}[XY] \neq \mathbb{E}[X]\mathbb{E}[Y]}$$

- The equality is true only if X are Y are *independent*.

Statistical moments (centered) (normed) of order $r > 0$

$$\mu_X^r = \mathbb{E}[X^r]$$

$$\mu_{X\,centered}^r = \mathbb{E}[(X - \mu_X)^r] \qquad \mu_{X\,centered\,normed}^r = \mathbb{E}\left[\frac{(X - \mu_X)^r}{\sigma_X^r}\right]$$

Random variables

First statistical moment (mean)

The *mean* refers to one measure of the central tendency of a probability distribution. It informs about the *location* of the probability distribution and it is defined as:

$$\boxed{\mu_X = \mathbb{E}[X^1] = \mathbb{E}[X]}$$

Second central moment (variance)

The variance is the second indicator of the central tendency, it sums up the *variability* of the probability distribution, the variance is given by:

$$\boxed{\sigma_X^2 = \mathrm{Var}\,[X] = \mathbb{E}[(X - \mu_X)^2]}$$

(if it exists)

A random variable with a finite variance is a variable of the *second order* (counter-example : The Cauchy distribution).

An other indicator of dispersion is the coefficient of variation :

$$\mathrm{c.o.v.} = \frac{\sigma_X}{|\mu_X|}, \quad \mu_X \neq 0$$

(σ_X is the *standard deviation* , homogeneous to X and μ_X)

Random variables

Properties of variance

The variance is obviously *non linear* but:

$$Var[aX + b] = a^2 Var[X]$$

$$Var[X+Y] = var[X] + Var[Y] + 2\underbrace{\mathbb{E}[(X-\mu_X)(Y-\mu_Y)]}_{Cov[X,Y]}$$

Another important relation (for hand calculations), is the *König-Huyghens formula* :

$$\begin{aligned} Var[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ &= \mathbb{E}[X^2] - \mu_X^2 \end{aligned}$$

Thus, for instance, it can be shown that if X and Y are *independents* :

$$Var[XY] = Var[X]Var[Y] + Var[X]E[Y]^2 + Var[Y]E[X]^2$$

Random variables

The normalized 3rd central moment (skewness)

The *skewness* is a shape indicator, it measures the (a)symmetry of the distribution:

$$\delta_X = \mathbb{E}\left[\frac{(X-\mu_X)^3}{\sigma_X^3}\right]$$

A symmetric distribution has a zero skewness (example: the normal distribution).

The normalized 3rd central moment (kurtosis)

The *kurtosis is* a shape indicator, measuring the flattening of the probability distribution:

$$\kappa_X = \mathbb{E}\left[\frac{(X-\mu_X)^4}{\sigma_X^4}\right]$$

It is generally compared to the kurtosis of the *normal distribution* ($\kappa_X = 3$) to know if the studied distribution is more or less flattened than the normal one.

Random variables

Quantiles

The *quantile at probability level α, denoted x_α* is determined by the inverse reading of the cumulative distribution function (strictly increasing)

$$F_X(x_\alpha) = \alpha \quad \Rightarrow \quad x_\alpha = F_X^{-1}(\alpha), \quad 0 \le \alpha \le 1$$

The *quantile function* is defined as *the inverse cumulative distribution function.*

The *median* is the 50% quantile. The *first* (resp. *third*) *quartile* is the 25% quantile (resp. 75%).

Confidence intervals

To sum up the variability of a random variable, one can use a confidence interval.

It is bounded by two quantiles *centered on the median*.

The *confidence interval at the probability level of $1-\alpha$* is given by:

$$\left[x_{\alpha/2}; x_{1-\alpha/2}\right] = \left[F_X^{-1}(\alpha/2); F_X^{-1}(1-\alpha/2)\right], \quad 0 \le \alpha \le 1$$

Outline

- General definitions
  - Random experiment – Event
  - Measurable space
  - Kolmogorov axioms
  - Bayes' theorem
- Random variables
  - Definitions
  - Cumulative distribution function (CDF) and probability density function (PDF)
  - Discrete / continuous random variables
  - Statistical moments
  - Confidence intervals (CI)
- Random vectors
  - Definitions
  - Moments
  - Copulas

Random vectors

Definition

A random vector is a *measurable function*:

$$\begin{aligned}\mathbf{X} &: \Omega \to \mathbb{X} \subseteq \mathbb{R}^n \\ &\omega \mapsto \mathbf{x} = \mathbf{X}(\omega) = (X_1(\omega), \dots, X_n(\omega))^t\end{aligned}$$

Where the dimension n of the support space $\mathbb{X}$ is larger than 1.

It is a *multi-dimensional random variable*.

It is defined by:

- *Its joint cumulative distribution function*:

$$F_{\mathbf{X}}(\mathbf{x}) = \mathbb{P}\left[\bigcap_{i=1}^{n} X_i \leq x_i\right]$$

- *Its joint probability density function*:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\mathbb{P}[\bigcap_{i=1}^{n} x_i \leq X_i \leq x_i + dx_i]}{\prod_{i=1}^{n} dx_i} = \frac{\partial F_{\mathbf{X}}(\mathbf{x})}{\partial x_1 \dots \partial x\_n}$$

Random vectors

Complementary definitions

- *The marginal* probability density function is the probability density function of a *sub-vector* of $\mathbf{X}$.
  If $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)^t$, the marginal density of $\mathbf{X}_1$ (in $\mathbf{X}$) is given by:

$$f_{\mathrm{X}_1}(\mathbf{x}_1) = \int_{\mathbf{x}_2 \in \mathrm{X}_2} f_{\mathrm{X}}(\mathbf{x}_1, \mathbf{x}_2)\, \mathrm{d}\mathbf{x}_2$$

- The *conditional density function* is the probability density function of the sub-vector of $\mathbf{X}$ given the occurrence value of the *complementary sub-vector*.
  If $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)^t$ the conditional probability density function of $\mathbf{X}_1$ given $\mathbf{x}_2 = \mathbf{a}$ is:

$$f_{\mathrm{X}_1|\mathrm{X}_2}(\mathbf{x}_1 \mid \mathbf{x}_2 = \mathbf{a}) = \frac{f_{\mathrm{X}}(\mathbf{x}_1, \mathbf{a})}{\int_{\mathbf{x}_1 \in \mathrm{X}_1} f_{\mathrm{X}}(\mathbf{x}_1, \mathbf{a})\, \mathrm{d}\mathbf{x}_1} = \frac{f_{\mathrm{X}}(\mathbf{x}_1, \mathbf{a})}{f_{\mathrm{X}_2}(\mathbf{a})}$$

  According to the *Bayes theorem*.

The associate cumulative distribution functions are obtained thanks to their definition (*i.e.* by integration).

Random vectors

Statistical moments

By definition, the *expected value* of a random vector is the vector of expected values of random variables that compose it :

$$\mathbb{E}[\mathbf{X}] = (\mathbb{E}[X_i], i = 1, \dots, n)^t$$

Its property of *linearity* holds.

The *covariance matrix* is the matrix whose element in the *i*, *j* position is:

$$\sigma_{ij} = \mathrm{Cov}\left[X_i, X_j\right] = \mathbb{E}\left[(X_i - \mu_{X_i})(X_j - \mu_{X_j})\right], \quad i, j = 1, \dots, n$$

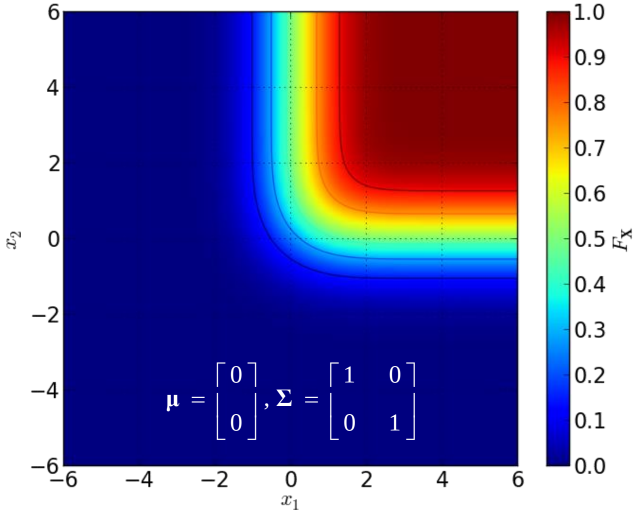
Thus the *variance of the components* are found *on the diagonal* ($\sigma_{ii} = \sigma_i^2$).

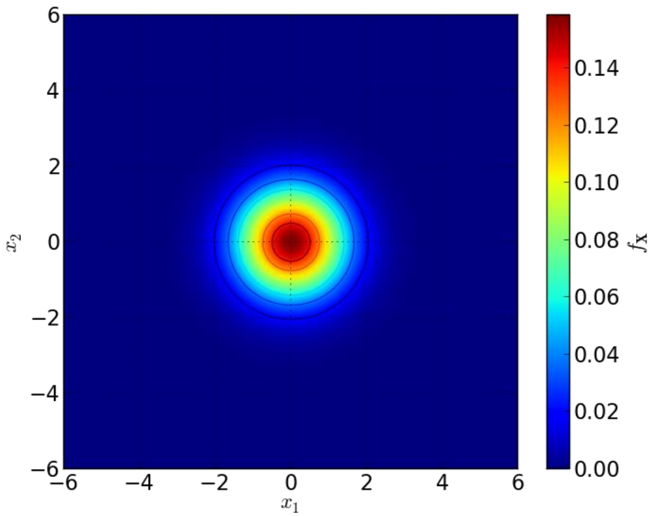
One defines as well the linear *correlation matrix* whose the i-j element is given by:

$$\rho_{ij} = \frac{\mathrm{Cov}\left[X_i, X_j\right]}{\sqrt{\mathrm{Var}\left[X_i\right]\mathrm{Var}\left[X_j\right]}} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}, \quad i, j = 1, \cdots, n$$

Multivariate normal distribution

$\mathbf{X} \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ on $\mathbb{R}^n$





CDF	$\Phi_n(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$
PDF	$\varphi_n(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \dfrac{\exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]}{\det(\boldsymbol{\Sigma})^{1/2}(2\pi)^{n/2}}$
Mean	$\boldsymbol{\mu}$
Covariance	$\boldsymbol{\Sigma}$

By definition, if $\boldsymbol{\Xi}$ is a vector of n independent standard normal random, if $\mathbf{L}$ is solution of $\boldsymbol{\Sigma} = \boldsymbol{L}\,\boldsymbol{L}^T$ (symmetric squared matrix of size n) and $\boldsymbol{\mu}$ is a vector of size n, then:

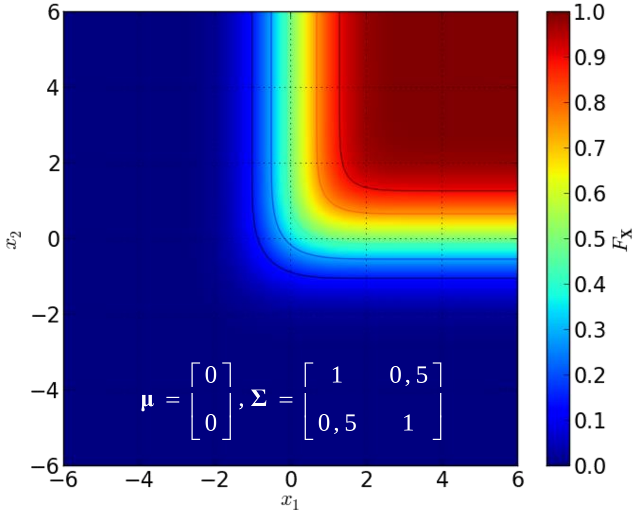
$$\mathbf{X} = \mathbf{L}\,\boldsymbol{\Xi} + \boldsymbol{\mu} \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

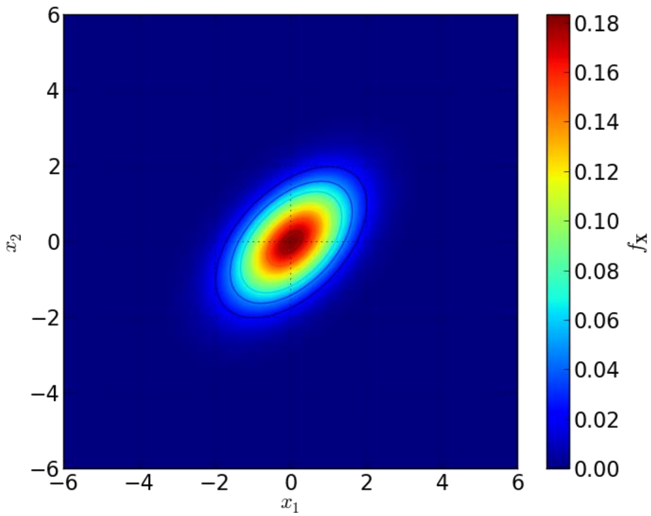
Consequently, any linear combination of Gaussian vectors is Gaussian.

Random vectors

Multivariate normal distribution

$\mathbf{X} \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ on $\mathbb{R}^n$





Let $\mathbf{X}$ be a Gaussian vector defined as :

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} \sim \mathrm{N}_n \left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{12}^{\mathrm{T}} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \right)$$

The sub-vector $\mathbf{X_1}$ (as $\mathbf{X}_2$) is also Gaussian and it is enough to forget the crossed terms of covariance matrix:

$$\mathbf{X}_1 \sim \mathrm{N}_{n_1} (\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$$

Random vectors

Copulas

A *copula* (denoted C) is a joint cumulative distribution function defined on the unit cube $[0\,;\,1]$ with uniform variables (*marginal*). See Sklar's theorem for more details.

Let $\mathbf{X}$ be a random vector of size n, with multivariate cumulative distribution function $F_{\mathbf{X}}$, and with marginal cumulative distribution functions $(F_{X_i}, i = 1, \dots, n)$.

There is a copula C of size n such that:

$$F_{\mathbf{X}}(\mathbf{x}) = C\left(F_{X_1}(x_1), \cdots, F_{X_n}(x_n)\right), \quad \mathbf{x} \in \mathrm{X}$$

If $\mathbf{X}$ is a *continuous random vector*, then the copula is *unique*. If $\mathbf{X}$ is *discrete*, the copula is *defined uniquely on the support* $\mathbb{X}$.

The *copula* is what is remained of a random vector, once the effects of the marginal distributions are removed.
It is the *stochastic dependence structure*.

Random vectors

Synthesis

- A random vector can be defined directly from its *joint distribution* (*e.g. the* multivariate normal distribution).
- Or, it can be defined from a *collection of marginal distributions* and *a stochastic dependence structure* expressed as a copula.
- The copulas formalism allows also to simply express the joint probability density function from its definition:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\partial F_{\mathbf{X}}(\mathbf{x})}{\partial x_1 \cdots \partial x_n} = \left.\frac{\partial C(u_1,\cdots,u_n)}{\partial u_1 \cdots \partial u_n}\right|_{u_i = F_{X_i}(x_i)} \prod_{i=1}^{n} \frac{\partial F_{X_i}(x_i)}{\partial x_i}$$

$$= c\left(F_{X_1}(x_1),\cdots,F_{X_n}(x_n)\right)\prod_{i=1}^{n} f_{X_i}(x_i)$$

Where c is, by definition, the *density function of the copula* C.

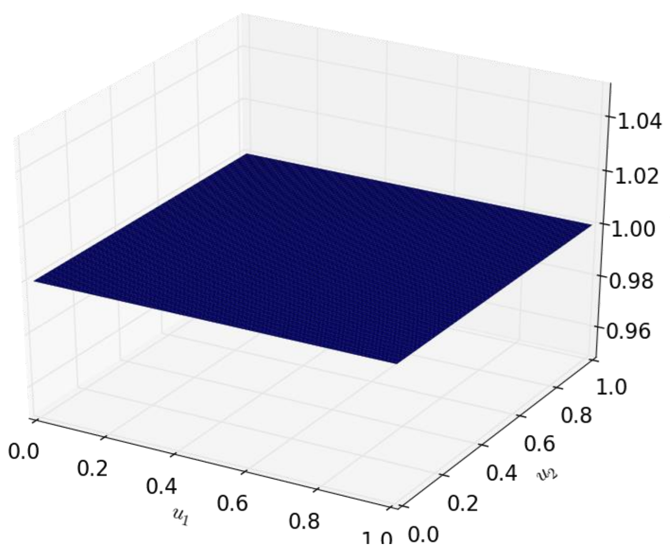
Random vectors

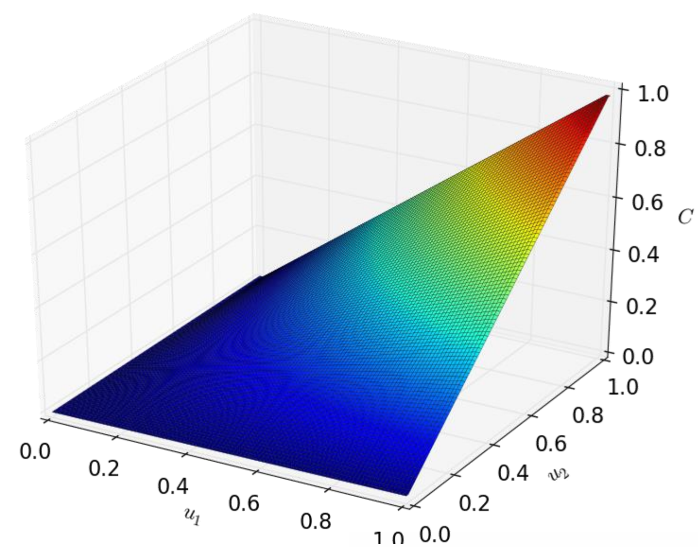
Independent copula $n \geq 2$

CDF	$C(\mathbf{u}) = \prod_{i=1}^{n} u_i$
PDF	$c(\mathbf{u}) = 1, \quad \mathbf{u} \in [0;1]^n$

Thus, the joint cumulative distribution function (resp. density) is reduced to the *product* of the marginal cumulative distribution functions (resp. density):

$$F_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^{n} F_{X_i}(x_i) \qquad f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^{n} f_{X_i}(x_i)$$





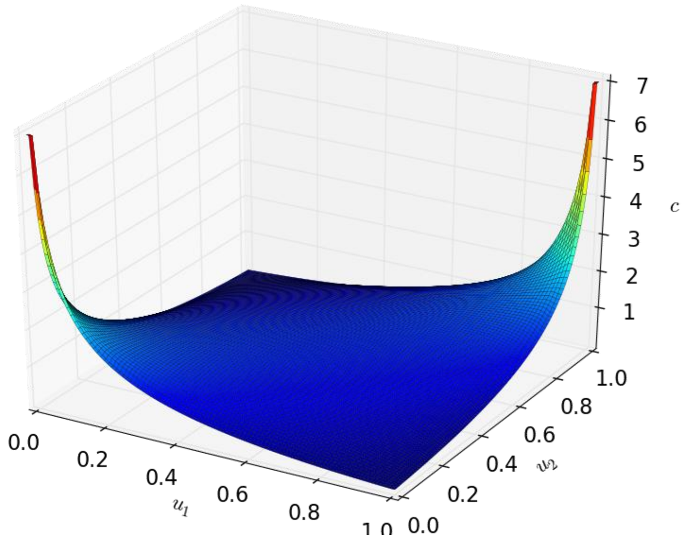
Random vectors

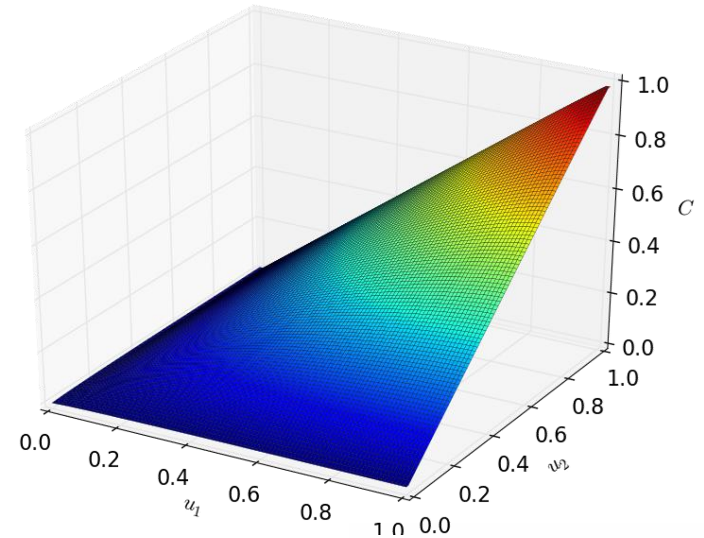
Gaussian copula $n \geq 2$

Family	Elliptic
CDF	$C(\mathbf{u}) = \Phi_n\left(\Phi^{-1}(u_1), \cdots, \Phi^{-1}(u_n); \mathbf{R}_0\right)$
PDF	$c(\mathbf{u}) = \dfrac{\varphi_n\left(\Phi^{-1}(u_1), \cdots, \Phi^{-1}(u_n); \mathbf{R}_0\right)}{\prod_{i=1}^{n} \varphi\left(\Phi^{-1}(u_i)\right)}$

Example: $\mathbf{R}_0 = \begin{bmatrix} 1 & 0,5 \\ 0,5 & 1 \end{bmatrix}$

$\mathbf{R}_0$ is not the linear correlation matrix!





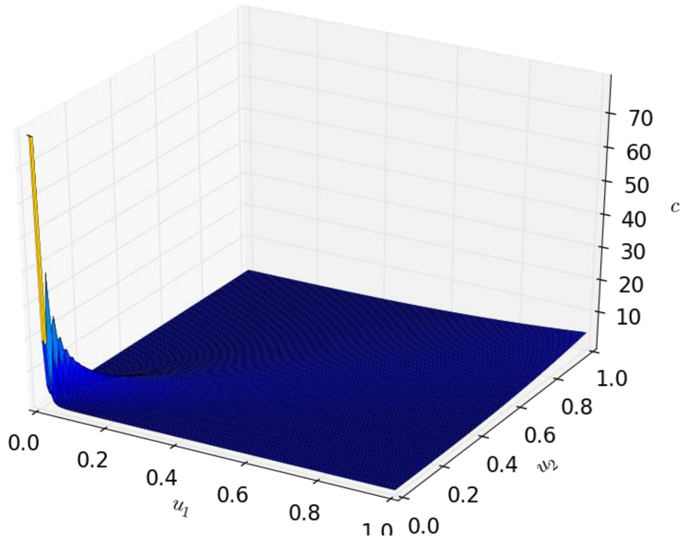
Random vectors

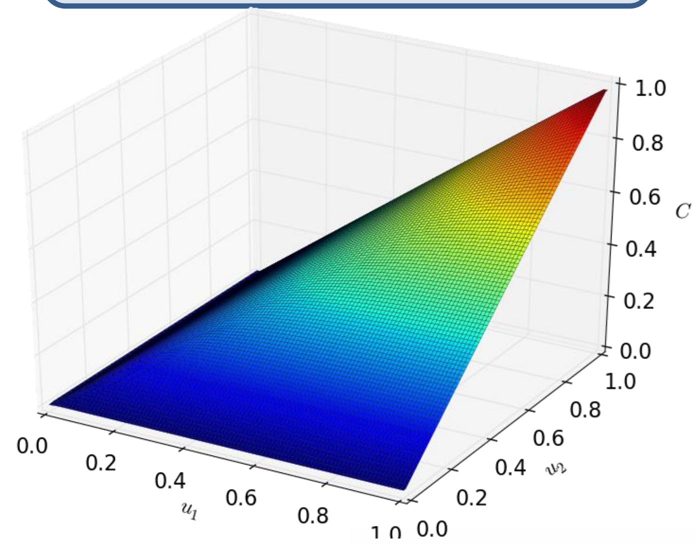
Clayton copula $n = 2$

Family	Archimedean
CDF	$C(u_1, u_2) = \left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta}$
PDF	$c(u_1, u_2) = (\theta + 1)(u_1 u_2)^{-(\theta+1)}\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta - 2}$

Example: $\theta = 3$

Lower tail dependence





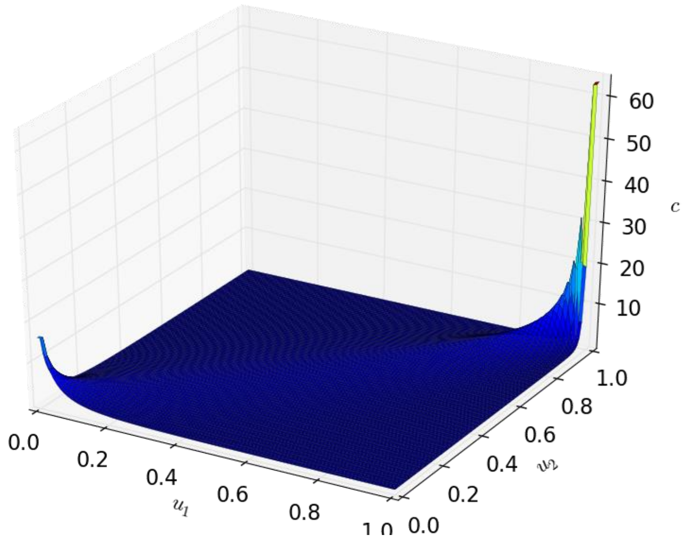
Random vectors

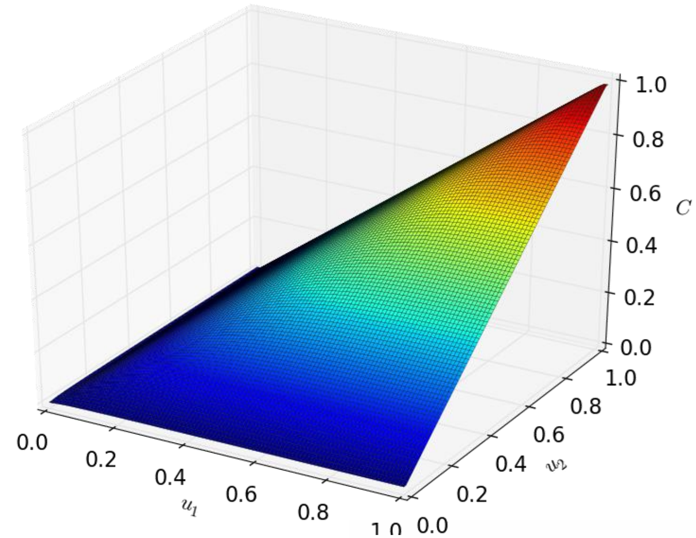
Gumbel copula $n = 2$

Family	Archimedean
CDF	$C(u_1,u_2) = \exp\left[-\left((-\ln u_1)^{\theta} + (-\ln u_2)^{\theta}\right)^{1/\theta}\right]$
PDF	$c(u_1,u_2) = C(u_1,u_2)\dfrac{(-\ln u_1)^{\theta-1}(-\ln u_2)^{\theta-1}\left((-\ln u_1)^{\theta} + (-\ln u_2)^{\theta}\right)^{1/\theta-2}(\theta-1-\ln C(u_1,u_2))}{u_1u_2}$

Example: $\theta = 3$

Upper tail dependence





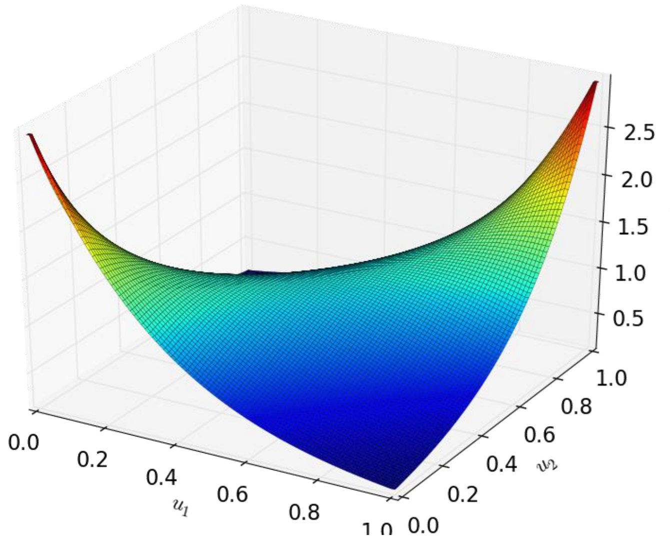
Random vectors

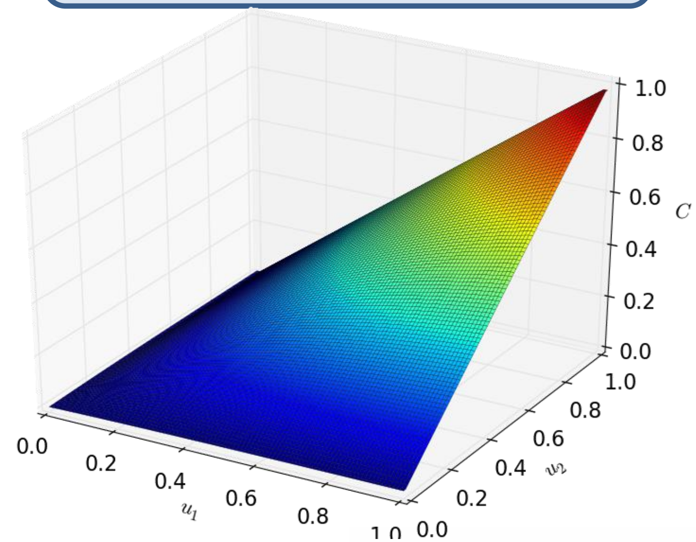
Frank copula $n = 2$

Family	Archimedean
CDF	$C(u_1,u_2) = -\frac{1}{\theta}\ln\left(1+\frac{\left(e^{-\theta u_1}-1\right)\left(e^{-\theta u_2}-1\right)}{\left(e^{-\theta}-1\right)}\right)$
PDF	$c(u_1,u_2) = \frac{\theta\left(1-e^{-\theta}\right)e^{-\theta(u_1+u_2)}}{\left[\left(1-e^{-\theta}\right)-\left(e^{-\theta u_1}-1\right)\left(e^{-\theta u_2}-1\right)\right]^2}$

Example: $\theta = 3$

symmetric dependence

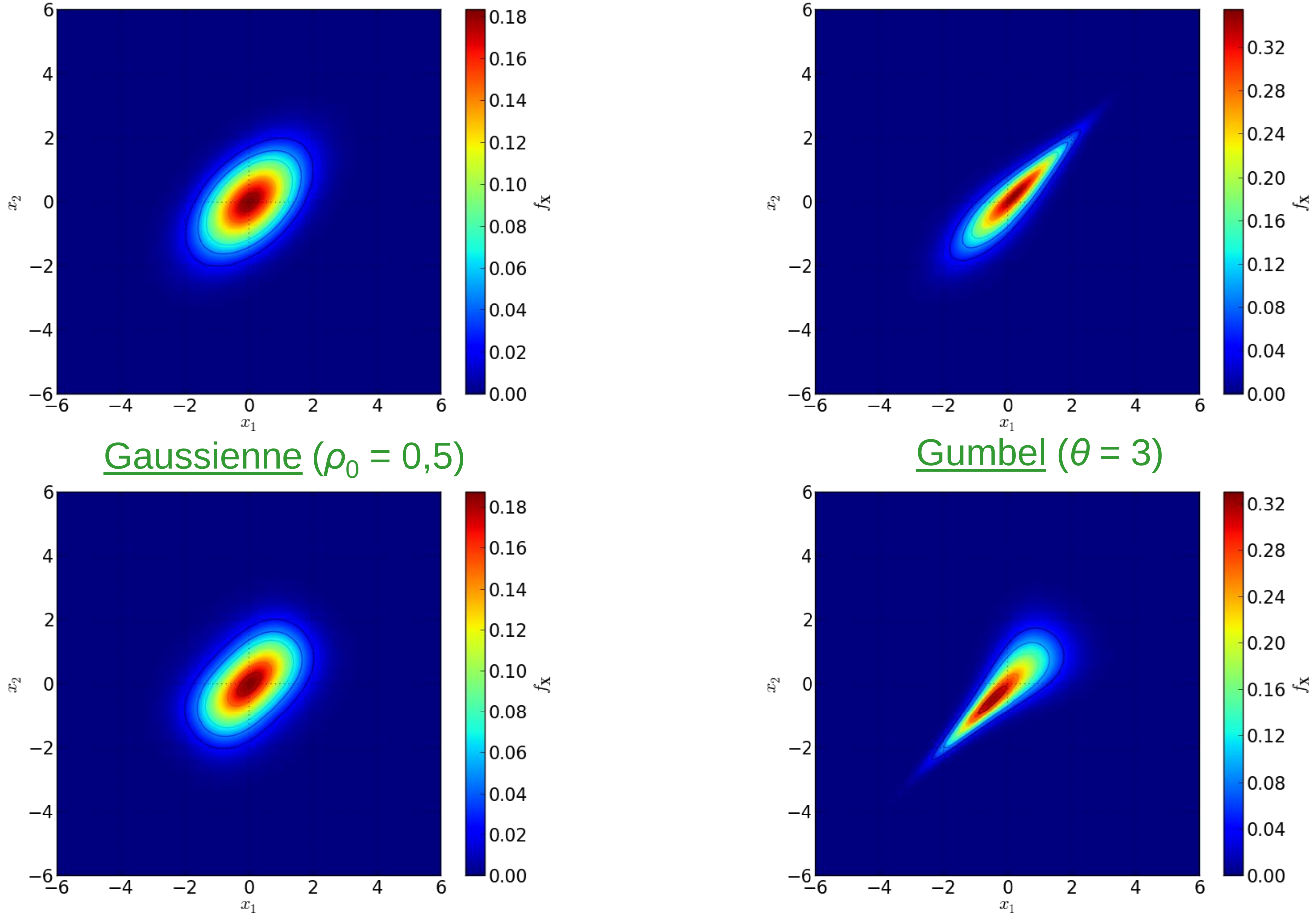




Random vectors

Examples of composed distributions $n = 2$

Two normal standard random variables are linked with different copulas and their corresponding probability density functions are plotted.



Gaussienne (ρ_0 = 0,5) — Gumbel (θ = 3) — Frank (θ = 3) — Clayton (θ = 3)